

**NATIONAL SCIENCE FOUNDATION
SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT
SUBMISSION
NATIONAL SCIENCE FOUNDATION
EDUCATION AND TRAINING APPLICATION PILOT
OMB Clearance Number 3145-0248**

A. Justification

1. Necessity for the Data Collection

This project will pilot test a longitudinal data collection system that the National Science Foundation (NSF) can use to track participants in its human capital development programs. Called the NSF Education and Training Application (ETAP), the system builds upon an existing system developed for the NSF Research Experiences for Undergraduates (REU) program. ETAP will be enhanced based on lessons from the initial REU pilot, revised to allow use by different types of programs, and tested with a broader set of programs before expanding its use among human capital development programs.

ETAP will enable NSF to collect high quality data needed for robust evidence-building activities, including monitoring, targeted research, and rigorous evaluation of its programs. ETAP will bolster Agency capacity to respond to (1) Administration priorities (such as [racial equity](#) and [restoring trust in government](#)), (2) the Foundations for Evidence-Based Policymaking Act of 2018 ([Public Law 115-435](#)), which promotes the collection and use of robust data to generate evidence for decision making, and (3) the America COMPETES Reauthorization Act of 2010, which requires that students in the REU program “be tracked, for employment and continued matriculation in STEM fields, through receipt of the undergraduate degree and for at least three years thereafter” (Section 514[a][6] of [Public Law 111-358](#)).

ETAP will also provide a service to future scientists and the community of NSF Principal Investigators, a service that intentionally promotes equity in participation in NSF programs by:

- Providing a transparent and centralized location of information on training opportunities and reducing burden on respondents (mostly students) who will be able to use a common application to apply to multiple training opportunities funded by NSF
- Lowering barriers to entry into NSF programs for new and aspiring Principal Investigators (PIs) who will be able to leverage a robust and secure data collection system, free of charge, to manage applications to their projects; and reducing administrative costs for existing PIs
- Providing (for PIs and NSF program officers) timely access to data analytics on applicants and participants to inform decision making and support improvement efforts

Participants for the pilot will be recruited among NSF PIs in different programs interested in volunteering. By participating in this study, they will have the opportunity to experience the data collection firsthand and provide feedback to help NSF improve the system before the Agency continues its expansion. Whenever possible, ETAP will leverage extant administrative data to avoid duplication, reduce burden, and encourage use.

This pilot is led by the Evaluation and Assessment Capability (EAC) Section of the Office of Integrative Activities with expert support from a contractor, Mathematica.

Description of the Pilot

This pilot includes:

- a. **Testing a web-based approach** to obtaining basic background and participation information while supporting the collection of additional application materials needed by PIs to make admissions decisions (see Attachment A for system specifications). Two important values informing system design were to ensure that ETAP respected PI autonomy in implementing projects/grants and that the system be scalable to different types of programs. To this end, while adhering to the guidance provided by NSF to grantees under different programs, PIs will continue to (1) choose whether to run applications by prospective participants competitively (for example, an REU Site award recruiting participants nationally) or non-competitively (for example, an REU Supplement award that invites prospective participants without holding an open call for applications), and (2) make admissions/ participation decisions. Data collected from prospective participants will depend on the type of application process selected by PIs.

The web-based system will include:

- **Common Registration.** All applicants will need to register to apply and participate in NSF-funded opportunities included in the pilot. The common registration form collects basic information for ongoing monitoring and future evaluation, such as demographic and contact information. Individuals applying for non-competitive opportunities will only need to complete the registration form. Those applying for competitive opportunities, will need to provide additional application materials (listed in the next bullet) requested by PIs for admissions decisions.
- **Common Application (containing additional application requirements).** Individuals submitting applications for competitive opportunities will be able to use the ETAP to apply to multiple NSF opportunities through a common application. They will first complete the common registration form (described above) and proceed to the common application through which they will submit additional information commonly requested by PIs, such as resumes, transcripts, and contact information of reference writers. The system will allow some level of customization for PIs to include additional requirements.

PIs and their authorized designees will use ETAP to provide information needed by potential applicants, retrieve applicant information, record application decisions and participation status among admitted applicants, and produce reports of data submitted by applicants to their award.

- b. **Gathering information about program experiences and satisfaction.** The pilot will include an exit survey administered to participants to capture program experiences and opinions.
- c. **Obtaining and integrating educational and employment outcomes.** This pilot will test the feasibility of collecting and integrating robust outcomes data. Because it takes time for outcomes to occur (say, for students to graduate, publish, and get jobs), this pilot will leverage the sample of applicants who were part of the REU data system pilot (on which ETAP is based). By 2023, those applicants would be expected to have graduated from college and transitioned to the workforce or further education (such as graduate school). This pilot will follow these individuals to:
- Obtain educational outcomes from administrative data from the National Student Clearinghouse (NSC) that can be purchased at low cost to the Government and no burden to participants or PIs
 - Administer a short survey to obtain information about post-graduation outcomes, such as employment
 - Integrate research productivity outcomes (such as publications and patents) from Web of Science, Scopus, and the United States Patent and Trademark Office (USPTO). These administrative databases are already accessible through NSF systems.
- d. **Conducting usability testing and gathering feedback from different system users** to assess system enhancements, functionality, and analytics capabilities. The pilot will include a short user satisfaction survey (for PIs) as well as one-to-one interviews to gather in-depth feedback.

Draft data collection instruments are included in Attachment B.

Description of Respondents

- **Individuals interested in applying for training opportunities offered by programs in the pilot.** NSF supports individuals at different stages of their careers, as undergraduate students, graduate students, post-docs, teachers, early career faculty, and so on, through a wide range of program offerings. For example, students at two- and four-year colleges and universities are the target population of the REU program, students (at both undergraduate and graduate level) are the target population for the IRES program, and teachers are the target population of the Research Experiences for Teachers (RET). If applying for an opportunity offered through the pilot, these individuals will submit information through the common registration or common application.
- **Reference letter writers.** Individuals applying for competitive opportunities will be asked to provide contact information of reference letter writers. ETAP will send reference writers an automated email requesting reference letters (to be submitted through ETAP).
- **Principal investigators (PIs) or staff designated by the PIs participating in the pilot.** PIs or their designees will submit (1) information useful to prospective applicants (such

as the opening and closing dates of applications) and (2) decisions/participation information (they will indicate which individuals were admitted and which later participated in the program at their site/project). This information will be submitted through a module designed for PIs.

- **Former participants.** After participating in an NSF program through ETAP during this pilot, participants will receive a one-time exit survey to gauge program satisfaction and experiences. Those that participated in the 2019 REU pilot will be included in a survey to obtain post-graduation outcomes, such as employment.

2. Use of Information

The information collected through the pilot will enable NSF to assess an approach to collecting data needed to respond to NSF information needs, Administration priorities, and Congressional requirements, as outlined in section 1. It will also enable PIs to access information they need to make admissions decisions and comply with NSF reporting requirements. The usefulness and success of ETAP will be assessed based on several factors—including data quality, user burden, and user feedback—and compared against the current approach used by NSF to collect data. An important outcome of this project, if it is successful, is to support integration of ETAP with other NSF systems to avoid duplication and enable timely access to relevant data for evidence-based decision making.

3. Improved Information Technology to Reduce Burden

The pilot will create one centralized online infrastructure that will collect information from applicants for NSF opportunities and support PIs seeking to recruit participants, in particular more diverse participants. ETAP will:

- Reduce burden on applicants, as the same application can be submitted to multiple opportunities (for example, at present, most REU Sites require that applicants follow each Site's application procedures)
- Reduce burden on participants, as demographic and educational outcomes data will be obtained from existing sources (instead of through surveys of former participants)
- Reduce burden on PIs and administrators who will no longer need to maintain their own applications and can leverage the system to comply with NSF reporting requirements
- Lowers barriers to participation in NSF programs for new/prospective PIs who can rely on ETAP instead of developing their own application systems
- Increases efficiency in program administration as PIs can quickly launch, and easily monitor, applications

- Reduces burden on the Government, as ETAP provides easy access to robust data on participants and applicants that would otherwise be obtained at higher burden and lower quality

4. Efforts to Identify Duplication

A feasibility study conducted by the Science and Technology Policy Institute (STPI) identified the necessity of developing a system to fill gaps in information. ETAP responds to this need and was designed to avoid duplication. Specifically, ETAP (1) will test using NSF's Single ID for account authentication to avoid duplication (of efforts and of NSF accounts) and foster integration across NSF systems and (2) upload data from other NSF systems, such as Fastlane for award data (so PIs do not need to provide the same information again). Results of the pilot will enable the Foundation to decide how to adjust the administration of its demographic survey of former participants (which will not be needed for those who participate in ETAP).

5. Efforts to minimize burden on small business

Small businesses are not affected by this information collection.

6. Consequences of Less Frequent Data Collection

Less frequent data collection will impair NSF's ability to meet its information needs and respond to Administration priorities and congressional requirements. Before embarking on this effort (or its predecessor, the REU data system), NSF commissioned a feasibility study to assess whether its current data collections would be adequate to meet the congressional requirements for the REU program. This feasibility study was conducted by STPI and concluded that "new data collection will be required, as the status quo of [REU] participants providing demographic information to NSF's Research Performance Report System, coupled with voluntary tracking of participants' career choices by the REU [principal investigators], is clearly insufficient to meet the [congressional] mandate" (Zuckerman et al. 2016). An analysis of more recent cohorts of REU participants (other than those included in the STPI study) confirmed the earlier STPI findings.

Less frequent data collection may also prevent PIs and potential applicants from benefitting from ETAP, which was designed to support and leverage applications to opportunities to participate in NSF programs. These opportunities are made available through grants. Therefore, the frequency of the collection needs to align with the availability of opportunities as determined by PI implementation plans. For example, PIs determine when they open and close applications (in alignment with the plans in their grants).

7. Special Circumstances

Not applicable.

8. Federal register announcement and consultation outside the agency

a. Federal Register announcement

The 60-day notice to solicit public comments was published in the Federal Register on 4/1/2021. NSF received one comment and responded by providing this individual with a copy of this package.

b. Consultation outside the agency

A study conducted by the STPI had determined the need for NSF to create new data collection, as the status quo collection at NSF was deemed clearly insufficient to meet a congressional mandate (Zuckerman et al. 2016). To respond to the America COMPETES Act, NSF commissioned a system for the REU program. An expert advisory board was convened to obtain feedback on the system design and data collection. The current project is the evolution of this early test that originated with the REU program and is being revised to be scalable to other NSF programs.

9. Payment or gifts to respondents

A few users (PIs and applicants, such as undergraduate and graduate students) are expected to volunteer to provide feedback. PIs will not receive a payment or gift for testing the system or providing feedback if they participate in the pilot as they will be current NSF grantees in the programs participating in the pilot. Applicants (mostly undergraduate and graduate students) will be offered a gift of \$40 for their participation. This amount is roughly equivalent to four hours of time at minimum wage, and will include time to test the system, debrief on experiences using the system, and logistics (coordination and review of background information).¹ No other incentives, payments, or gifts will be offered or given to participants.

10. Assurance of confidentiality

The collection of this information constitutes a regular application to training opportunities funded by NSF. Applicants' information will be maintained in accordance with the requirements of the Privacy Act of 1974. No personal information will be released to the public. The system includes notices in two instances:

The first one appears when users first access the system to obtain a user ID and is based on the notice displayed in the NSF GRFP collection (OMB control number 3145-0023). The notice reads as follows:

¹ Source of minimum wage information: <https://www.dol.gov/whd/minwage/america.htm>

Rules of Behavior

This computer system is the property of the National Science Foundation (NSF) of the Federal Government. Any system activity may be monitored and any information stored within the system may be retrieved and used by authorized personnel for law enforcement, management, routine system operations, or other purposes. By using this computer system, you are consenting to such monitoring and information retrieval and use.

Unauthorized use of the system, including disclosure of information covered by the Privacy Act or other sensitive information, or attempts to defeat or circumvent security features, is prohibited and could result in disciplinary action, civil and/or criminal penalties. Users should be aware that they have no expectation of privacy when using the NSF-provided computer system (including any removable media used in conjunction with the system), accessing the Internet, or using electronic mail systems.

All information maintained within or retrievable through the NSF computer system, including electronic mail files, may be reviewed and retrieved by the Department of Homeland Security; NSF officials who have a legitimate reason to do so when authorized by the Director or Deputy Director; or by the Inspector General.

I acknowledge the rules of behavior

Notice

An agency may not conduct or sponsor, and a person is not required to respond to, an information collection unless it displays a valid Office of Management and Budget (OMB) control number. The OMB control number for this collection is 3145-0248. Public reporting burden for this collection of information is estimated to average 3.25 hours for registering and 7 hours for registering and submitting an application, including the time for reviewing instructions. Burden estimates for principal investigators is 4.7 hours to register and record admissions decisions and program attendance. Send comments regarding the burden estimate and any other aspect of this collection of information, including suggestions for reducing this burden, to:

Suzanne Plimpton
Reports Clearance Officer
Budget and Finance Administration
National Science Foundation
Arlington, VA 22230

Please note that information provided through the ETAP system will be used for admissions decisions, audits, and research and evaluation purposes. All applicants' information will be maintained in accordance with the requirements of the Privacy Act of 1974. No personal information will be released to the public.

The second instance of a notice appears when users certify and submit information. This notice reads as follows:

By clicking on the **SUBMIT** button below, I am certifying that the information provided is true and complete to the best of my knowledge. I understand that I am consenting to the confidential use of the information I provided for admissions decisions, audits, and research and evaluation purposes.

11. Justification for sensitive questions

Information regarding applicants' characteristics (such as gender and educational achievement) will be collected, as this information is needed by PIs (and is normally included in applications used by PIs) and is needed by NSF to monitor the program and respond to Administration guidance and congressional requirements. Date of birth will also be collected to ensure proper identification of applicants and obtain educational outcomes data from the NSC without increasing burden on participants.

12. Estimate of respondent burden

Table A1. Estimates of respondent burden by respondent type

Category of Respondent	Respondents (number)	Participation Time (hours)	Total Burden (hours)	Annual Burden (hours)
Common Registration and Application				
a. PIs (or their designated user) who will use the PI module to run applications	1,200	4.7	5,640	1,880
b. Applicants	64,910	5.7	369,987	123,329
c. Reference writers	121,899	0.5	60,950	20,317
Program Experiences and Outcomes				
d. Participant exit survey	9,600	0.33	3,168	1,056
e. Participant employment survey (former cohort of participants)	471	0.25	118	39
f. PI support in participant survey	315	0.25	79	27
g. Participant educational outcomes will be obtained from the NSC, creating no burden on participants	471 (excluded from the total count)	0	0	
h. Participant research productivity outcomes (publications and patents) will be obtained from Web of Science, Scopus, and USPTO,	471 (excluded from the total count)	0	0	

creating no burden on participants				
System User Feedback				
i. PI user satisfaction survey	384	0.16	61	20
j. User feedback interviews (Applicants & PIs)	50	2.5	125	42
Totals	198,554		440,128 hours	146,710 hours

Supporting information for Table A1

a. Principal Investigators (PIs) participating in the ETAP pilot

The pilot recruitment target is 120 PIs testing ETAP in 2022, 360 in 2023, and 720 in 2024; for a total of 1,200 system testers throughout the life of the pilot. (These could be different PIs or the same PIs participating in multiple years of the pilot.) PIs (of the award funding the given opportunity) will decide whether they will run a competitive or noncompetitive application for their award. We estimate that 2/3 of the PIs (N=792) will choose competitive applications and 1/3 noncompetitive (N=396), with each PI supporting, on average, 10 participants.

Number of respondents: We estimate that a total of 1,200 PIs or their designees—one for each opportunity offered (say, an REU Site award in a given year)—will provide information through ETAP to manage applications.

Participation time (number of burden hours per respondent): Estimated PI burden is 4.7 hours for each application cycle and includes time to:

- a) Provide information about the opportunity and submit additional application requirements (0.5 hours per opportunity)
- b) Record admission and acceptance decisions for those running a competitive application (2 items per each of the estimated 177 applicants at a rate of 1 minute per item = 5.9 hours per opportunity)
- c) Record admission and acceptance decisions for those running a noncompetitive application (2 items per each of 10 participants invited to apply at a rate of 1 minute per item = 0.33 hours per opportunity)
- d) Record participation information (1 item for 10 participants at the rate of 1 minute per item = 0.16 hours per opportunity)

Weighing the estimates in b) and c) by the number of PIs choosing a competitive and noncompetitive application cycle results in the overall burden estimate of 4.7 hours.

b. Applicants participating in the ETAP pilot

Number of respondents: We estimate that 30,678 unique applicants will use ETAP throughout the three years of the pilot to submit applications for training opportunities.

Based on data from the REU data system pilot, students apply to 2.3 opportunities, on average, and PIs running competitive applications will receive, on average, about 177 applications. We therefore estimate that 64,910 unique system users will apply to a competitive opportunity (177 applications per opportunity multiplied by 792 PIs running competitive applications and divided by 2.3 to identify unique applicants), and 3,690 will apply to a noncompetitive opportunity (10 applicants for each of the 369 noncompetitive opportunities in the pilot). This results in an estimated total of 64,910 applicants.

Participation time (number of burden hours per respondent): We estimate that applicants will take about 5.7 hours using ETAP to apply to training opportunities of their choice (3.25 hours to complete the common registration form, and, if applying to a competitive opportunity, an additional 3.8 hours preparing additional application requirements, weighted by the share of applicants applying to competitive and noncompetitive opportunities). These burden estimates are based on actual usage data from the precursor system (REU data system), included in Attachment C.

c. Reference Writers

Number of respondents: Students applying to a competitive training opportunity are asked to submit contact information for reference writers. ETAP will automatically send an email to reference letter writers inviting them to submit a letter of recommendation through ETAP. We estimate 27,548 applicants applying to competitive opportunities and, therefore, 53,436 reference writer requests (assuming two reference letters are requested, on average).

Participation time (number of burden hours per respondent): We estimate that it will take 0.5 hours for reference writers to complete the reference form in ETAP, which includes a short paragraph about the applicant and requests that applicants be rated on 11 characteristics (by selecting checkboxes).

d. Participant exit survey

Number of respondents: We estimate that a total of 9,600 participants will respond to the exit survey (10 participants per opportunity multiplied by 1,200 opportunities in the pilot and multiplied by 0.8 expected response rate). A review of similar surveys—focused on participants in programs offering research training to undergraduate and/or graduate students—and our own experience testing this exit survey with a small sample of REU Sites, suggests that response rates can reach nearly 80 percent if the survey is administered immediately after (or preferable on the last day of) the training (Lopatto 2004, 2007; Mathematica 2020).

Participation time (number of burden hours per respondent): The survey is expected to take about 20 minutes to complete (0.33 of an hour). This is based on the findings from the initial pilot with REU Sites, which showed that ninety percent of participants completed the survey in one day and took, on average, 19.8 minutes to complete it (Mathematica 2020).

e. Participant employment survey

Number of respondents: The employment survey will be tested with individuals who (1) participated in the REU data system pilot and (2) meet timing requirements of the American COMPETES Act (collect educational and employment outcomes of participants “for at least three years” after receipt of undergraduate degree). A total of 471 participants were part of the 2019 REU pilot cohort. These individuals, who participated in REU as undergraduates in the summer 2019, will be expected to have transitioned to graduate school or the labor market in the coming years. For example, by 2023, those who participated as seniors in 2019 would have been three years out of undergraduate graduation, allowing NSF to test its ability to track outcomes as required by legislation.

Participation time (burden hours per respondent): The survey will take no more than 15 minutes to complete. This time should be sufficient to answer the few questions needed to comply with the congressional requirement for the REU program, plus a few questions needed to meet the information needs of NSF leadership and program officers. We will seek to keep the survey short to avoid unnecessary burden and improve response rates and data quality. Note that the literature does not indicate a strong correlation between survey length and response rates. However, research does suggest that “stated” survey length of time is associated with lower response rates, and survey length influences data quality of questions positioned later in the survey (Galesic and Bosnjak 2009).

f. PI support in participant survey

Number of respondents: A sample of PIs (N=315) will be engaged in a study to rigorously study promising low-cost strategies to increase survey response rates and minimize response rate bias in online surveys. We will leverage one of the survey administrations to randomly assign participants to different strategies to increase survey response rates. We will test how engaging PIs to send the follow-up notification to nonrespondents (treatment 1) compares to contacting nonrespondents through social media (treatment 2) and sending automatic follow-up emails through the ETAP system (control condition). (See appendix D for more information on the study design). Findings will help inform the design of ETAP and future efforts at NSF and other agencies interested in surveying a similar population of young professionals or otherwise internet-savvy college populations.

Participation time (burden hours per respondent): We assume that PIs will take up to 15 minutes to read instructions and send an email notification (draft to be provided to them) to about 2 participants who had been randomly selected from within the pool of survey nonrespondents.

g-h. Participant educational and research productivity outcomes

Number of respondents: We will obtain educational outcomes data from the NSC and research productivity outcomes (publications and patents) from Web of Science, Scopus,

and USPTO for the 471 former REU participants who were part of the 2019 REU data system pilot.

Participation time (burden hours per respondent): Not applicable, as the data are available through existing sources.

i. PI system user satisfaction survey

Number of respondents: We estimate that 384 PIs (or their designated users) will respond to a user satisfaction survey planned to guide ongoing improvements throughout the first years of testing (480 PIs in the first two years multiplied by a 0.8 response rate).

Participation time (burden hours per respondent): The survey is intentionally designed to be a short pulse survey that should take less than 10 minutes to complete (0.16 of an hour).

j. User feedback interviews

Number of respondents: We estimate gathering user feedback with a total of 50 users (25 PIs and 25 students) through interviews to test new functionality and obtain feedback.

Participation time (burden hours per respondent): We estimate 2.5 hours to participate in these interviews, including 2 hours to test the system and debrief on experiences using the system, and an additional 1/2 hour for scheduling and other logistics.

13. Cost Burden to Respondents

There are no direct costs to respondents.

14. Cost Burden to the Federal Government

The total estimated cost of this data collection is \$4,352,034, which includes design and system development (including data analytics and data download capabilities) and 3 rounds of data collection and analyses. The resulting annualized cost of the pilot test is approximately \$1,450,678.

15. Reason for Change in Burden

The number of respondents has increased to test ETAP at scale, and burden estimates have been updated based on empirical burden evidence gathered in the original pilot with the REU data system.

16. Schedule for information collection and publication

Task	Date
ETAP– year 1 data collection	November 2022
ETAP– year 2 data collection	November 2023
ETAP– year 3 data collection	November 2024
National Student Clearinghouse and other administrative outcome data	May 2023
Employment Follow-up Survey	Spring-Summer 2023

17. Display of OMB expiration date

The expiration date for Office of Management and Budget (OMB) approval will be displayed as shown in section 10 of this document.

18. Exception to the certification statement

There are no exceptions to the certification statement.