

National Science Foundation

National Science Foundation Annual

Evaluation Plan

FY 2024

Torres del Paine National Park, Chile. Credit: Carlye Calvin, University Corporation for Atmospheric Research

Introduction

The Foundations for Evidence-Based Policymaking Act of 2018, <u>Public Law No. 115-435</u> (Evidence Act), gave impetus to ongoing federal efforts to use evidence in decision making. This legislation created an opportunity to focus attention on promoting government effectiveness and efficiency by building and using evidence in the most impactful way. This document presents the FY 2024 Annual Evaluation Plan (AEP) that NSF developed in response to this opportunity and following guidance provided by the Office of Management and Budget (<u>OMB M-21-27</u>, <u>OMB M-19-23</u>, <u>OMB M-20-12</u>, and <u>OMB Circular No. A-11</u>).

This AEP describes the evaluations prioritized by NSF for FY 2024. This includes evaluations that NSF is planning to begin or continue in FY 2024. Section 1 presents the criteria used for selecting them. Section 2 provides a summary of progress on evaluations described in the FY 2022 and/or FY 2023 AEPs. Section 3 provides the research questions guiding each evaluation. Section 4 provides overviews of the background/ rationale, timeline, technical approach, data sources, expected challenges and mitigating strategies, and use and dissemination plans for each evaluation question. These evaluations—and all other evidence-building activities—shall be conducted in accordance with NSF's Evaluation Policy.



Los Pinos mountain front, looking south. Credit: Doug Moore, courtesy of Sevilleta LTER photo archives



Section 1 Significant Evaluations

Lichen-covered boulders. Credit: Frank Bungartz, Ph.D., Arizona State University Lichen Herbarium

The following are five criteria used to select evaluation questions:

- \bigcirc
- 1. Fill a knowledge gap: the information sought is not available from existing sources, such as scholarly literature and evaluations supported by other agencies implementing similar efforts



. Have leadership support: to prioritize the staff time and commit the resources that the work demands



Have potential to support upcoming decisions: are likely to yield actionable and useful evidence in a timely fashion



Have potential for broad impacts: will likely result in findings that are useful for a broad set of stakeholders, programs, or organizations



5. Are prioritized by NSF leadership: respond to evolving requirements, Congressional mandates, and national and long-term strategic priorities

These criteria were assessed as follows:

- Individually, criteria 1-3 are necessary but not sufficient conditions
- Questions meeting criteria 1-4 are likely to be prioritized, absent resource constraints
- Criterion 5 is a sufficient condition to identify a question as significant

These criteria, and their use, may be revised as implementation of the Evidence Act and related legislation matures and as NSF responds to changing priorities and external events, such as those observed in recent years.



Section 2 Evaluation Progress

Cutler Marsh in Cache County, Utah. Credit: Matt Jensen, Utah State University

Evaluation

Annual Evaluation Plan(s)

Convergence Accelerator (CA)

In what ways does the Convergence Accelerator Innovation Training contribute to the emergence of new capacities among participating researchers to meet pressing societal needs?

Progress and Plans

- FY 2022, FY 2023
- Structured interviews of program stakeholders elicited insights about several program strengths (agility, flexibility, diverse experience of program staff) and areas for potential improvement (workload management and targeted communications). NSF is considering how to integrate these lessons learned into the program design.
- A landscape analysis examined reports, survey data, and other materials that have been, or could be, generated by the CA program and would be useful for program evaluation and monitoring. The report also highlighted additional data and metrics that could facilitate evaluation and monitoring of the program.
- NSF is examining available administrative data and its ability to yield further information on characteristics of applicants and projects that are significantly more common (1) in projects selected for CA Phase 1 funding than in those not selected and (2) in Phase 1 projects that successfully progress to Phase 2 vs those that do not.
- NSF is preparing for a future impact evaluation for the CA program. As part of that work, NSF will assess the merit and utility of a literature review on possible outcome measures for convergent research to build a foundation for determining appropriate and feasible outcome measures.

Evaluation

EPSCoR

How do Established Program to Stimulate Competitive Research (EPSCoR) program funding strategies (infrastructure, co-funding, and outreach) contribute to increasing academic research competitiveness across jurisdictions?

Annual Evaluation Plan(s)



Progress and Plans

• NSF will use the results of this study to inform data needs and the technical approach to the FY 2024 evaluation, which will examine the extent to which program funding strategies contribute to increasing academic research competitiveness across jurisdictions.



Evaluation

Annual Evaluation Plan(s)

Partnerships

What are the benefits of receiving an award from a program supported by a partnership? How do these differ from benefits associated with awards from programs not supported by a partnership? What outputs and outcomes are associated with partnership programs? To what extent can these be attributed to the partnership programs? What improvements could make partnership programs more effective or easier to implement?



Progress and Plans

- NSF has developed an algorithmic approach to identify and characterize partnerships through text analysis of unstructured and structured NSF administrative data and an internal dashboard allowing users to explore the resulting information in different ways.
- These tools now allow users to identify and understand different types of partnerships that are represented across NSF and NSF-funded activities through an efficient process that replaces time-consuming manual data calls.
- NSF awarded a contract in FY 2022 to conduct a study to identify and measure the outputs and outcomes of NSF partnerships in which NSF has a direct relationship with the private sector partner and to identify ways to enhance the outputs and outcomes. Work is anticipated to be completed in late FY 2024.

Evaluation

Missing Millions

How can NSF help increase the participation of underrepresented groups in the STEM workforce?

Progress and Plans

• NSF has awarded a contract to conduct foundational work to describe underrepresented groups in NSF's portfolio of investments.



NSF Annual Evaluation Plan - FY 2024 | March 2023 5

Annual Evaluation Plan(s)



Section 3 FY 2024 Evaluation Questions At A Glance

Dwarf fireweed (Chamaenerion latifolium). Credit: Wendi Pillars, courtesy of ARCUS

EPSCoR

How are different NSF Established Program to Stimulate Competitive Research (EPSCoR) program strategies (e.g., direct funding for infrastructure, co-funding of proposals submitted to other directorates) correlated with increasing academic research competitiveness in target jurisdictions?

NSF's Hybrid Workplace

In what ways, if any, has the NSF merit review process changed as a result of adopting a hybrid workforce model?

GRANTED

To what extent are differences in institutional support services associated with differences in proposal submissions to NSF and with differences in funding rates for those proposals?

COVID Pandemic

To what extent did the COVID pandemic influence the participation of different groups, including those that have historically been underrepresented or underserved in STEM, in the NSF portfolio of programs and activities, as compared to pre-pandemic?



Section 4 FY 2024 Evaluation Plans

This section includes a brief study plan for each prioritized evaluation question. The plans show the alignment of these questions with NSF's current Strategic Plan. These plans also provide overviews of the background/rationale, timeline, technical approach, data sources, expected challenges and mitigating strategies, and use and dissemination plans.



How are different NSF Established Program to Stimulate Competitive Research (EPSCoR) program strategies (e.g., direct funding for infrastructure, co-funding of proposals submitted to other directorates) correlated with increasing academic research competitiveness in target jurisdictions?

Strategic Goal

Discover: Create knowledge about our universe, our world, and ourselves (SG2)

Strategic Objectives

Advance the frontiers of research (SO2.1) Enhance research capability (SO2.2)

Learning Agenda Alignment

FY 2022-6 (EPSCoR)

Background and Rationale

EPSCoR seeks to foster sustainable improvements in research and development (R&D) capacity in the 28 jurisdictions that individually received 0.75% or less of total NSF funding over the most recent five-year period. EPSCoR uses multiple investment strategies to support this goal, including (1) supporting academic institution research infrastructure, (2) co-funding meritorious proposals reviewed by other NSF programs that also satisfy EPSCoR criteria, and (3) promoting interaction between NSF and the EPSCoR community through workshops and other outreach activities. NSF EPSCoR seeks to expand its capacity to generate and use evidence to monitor program progress in increasing academic research competitiveness through its three funding strategies. This effort builds on exploratory studies released in FY 2021 and more formal assessments begun in FY 2023 and remains a significant priority. NSF's continued interest in understanding how to enhance and sustain R&D research capacity in EPSCoR jurisdictions is especially important in light of the passage of the CHIPS and Science Act of 2022 (<u>Public Law No. 177-167</u>), under which Congress has directed NSF to further increase support to EPSCoR jurisdictions (<u>Section 10325</u>).

Timeline FY 2023–FY 2025





Technical Approach

NSF will conduct a descriptive analysis of the EPSCoR program which may include studies such as (1) regression and other analyses of longitudinal data of EPSCoR jurisdictions to characterize and understand associations between program participation and observed outcomes while controlling for confounding factors that may be associated with outcomes and (2) case studies of jurisdictions that have surpassed EPSCoR eligibility limitations or have made significant progress toward improving research competitiveness to understand the approaches that enabled such progress. NSF also anticipates planning for analyses of effects from increased EPSCoR funding (if provided under the CHIPS and Science Act) on program outcomes over time.

Data Sources

This study will rely on a monitoring data system developed for the EPSCoR program and will draw data from NSF administrative data systems, existing national data collections, and new collections (as needed).

Challenges and Mitigating Strategies

A prior study (released in Summer 2021) indicated that it would be challenging to detect progress toward success for EPSCoR jurisdictions when the sole outcome measure was the program's eligibility criteria. This challenge will be mitigated by relying on a rich set of output and outcome measures that can be used both to monitor institutional and jurisdictional progress and for program improvement. Another challenge will be developing a definition of "academic research competitiveness" that includes multiple factors that describe different aspects of this complex concept and can be implemented for the purposes of this study. The definition may include such elements as a jurisdiction's capacity and/or ability to obtain NSF funding, to compete with other jurisdictions, and its relative prominence among its peers; fraction of awards resulting from major NSF-wide competitions such as Major Research Instrumentation (MRI) and Mid-Scale Research Infrastructure (MSRI); number of students graduating with advanced degrees in STEM; and others.

Use and Dissemination

Findings from this study will be shared with NSF EPSCoR, grantee universities, and jurisdiction science and technology steering committees to inform decisions that may influence the academic research competitiveness of institutions and jurisdictions.





In what ways, if any, has the NSF merit review process changed as a result of adopting a hybrid workforce model?

Strategic Goal

Excel: Excel at NSF operations and management (SG4)

Strategic Objectives

Strengthen at speed and scale (SO4.1) Invest in people (SO4.2)

Learning Agenda Alignment

FY 2022-2 (COVID Pandemic), FY 2022-9 (Merit Review)

Background and Rationale

Since March 2020, NSF has operated under a remote workforce model, with most staff working remotely on a nearly full-time basis. NSF is moving forward with a long-term approach to supporting a hybrid workforce, which the agency anticipates will include a mix of onsite and remote staff. NSF seeks to understand what outcomes related to agency merit review, external outreach and engagement, and other program-related operations are associated with the hybrid workforce model and what adjustments NSF staff have made to accommodate these changes.

NSF will conduct studies to assess such questions as: What strategies have programs adopted to conduct merit review with a mix of in-person and remote staff? To what extent is the proportion of remote staff or other factors associated with differences in proposal dwell time, panel costs, and other operational indicators? To what extent are changes in NSF's merit review process reflected in changes in merit review metrics such as average time to decision for proposals and the funding rate for NSF and its component organizational units?

Timeline FY 2024–FY 2025





Technical Approach

NSF will conduct interviews and/or focus groups with NSF staff and members of the research community to understand benefits and challenges with conducting merit review and external engagement with a hybrid workforce and the operational changes implemented in response. NSF will carry out descriptive analyses of the characteristics of NSF directorates and offices and develop regression models to assess what outcomes are correlated with a higher proportion of remote staff (for example, is a higher proportion of remote staff associated with higher dwell time or increased staff engagement in outreach activities?).

Data Sources

This study will rely on NSF administrative data (on employees, organizational units, review approaches, and dwell time) and information from interviews and/or focus groups (on challenges and opportunities related to remote work, merit review, and external engagement).

Challenges and Mitigating Strategies NSF anticipates challenges in identifying a definitive set of characteristics to conduct descriptive analyses of NSF organizational units to assess what outcomes are correlated with a higher proportion of remote staff. NSF will analyze the quality and comprehensiveness of administrative data and devise approaches to fill in data gaps as needed; during interviews, EAC will seek to determine if those approaches are reasonable and, if not, identify relevant factors to adjust analyses accordingly.

Use and Dissemination

Results from this assessment will support NSF efforts to implement an effective hybrid workforce model that maintains NSF's high quality merit review processes, efficiently stewards taxpayer dollars, and strengthens NSF engagement in states and jurisdictions that have historically received a smaller share of NSF funding.





To what extent are differences in institutional support services associated with differences in proposal submissions to NSF and with differences in funding rates for those proposals?

Strategic Goal

Empower: Empower STEM talent to fully participate in science and engineering (SG1) **Discover:** Create new knowledge about our universe, the world, and ourselves (SG2)

Strategic Objectives

Ensure accessibility and inclusivity (SO1.1) Unleash STEM talent for America (SO1.2) Enhance research capability (SO2.2)

Learning Agenda Alignment

FY 2022-1 (Missing Millions), FY 2022-9 (Merit Review)

Background and Rationale

Principal investigators (PIs) from emerging and developing research institutions may lack support to identify potential funding opportunities, prepare the documentation required for Federal grant submissions, and manage Federal awards. This may put individuals from such institutions at a disadvantage relative to their peers at institutions with more institutional support, hinder their participation in the STEM workforce, and limit the benefits to the Nation of their scientific talents. NSF's Growing Research Access for Nationally Transformative Equity and Diversity (GRANTED), a new initiative in FY 2023, focuses on addressing these systemic barriers by improving research support and service capacity at emerging, developing, and underserved research institutions (hereafter, "target institutions"). NSF will conduct a needs assessment and baseline evaluation of this important new NSF initiative.

Questions of interest may include: What are the characteristics of target institutions and what research support and services do they provide relative to institutions with higher levels of research activity? What indicators can be used to measure institution-level capacity, maturity, and service delivery? To what extent are institutional support services associated with increased proposals submissions to NSF and funding rates? What support do institutions need to grow research capacity? How can GRANTED best provide the needed support?

Timeline FY 2024–FY 2025





Technical Approach

The technical approach includes a descriptive analysis of the characteristics of target and comparator institutions using data from the federal statistical system and from NSF administrative data. Qualitative data will be collected from focus groups, interviews, and listening sessions, which will be used to engage institutions to better understand what research support and services they provide, how they think the GRANTED initiative can best support them, and potential indicators of progress. This will be supplemented with a literature review to identify other relevant outcome and impact measures of research capacity. NSF will also explore the feasibility of designing a capacity assessment that could be used to establish a baseline and measure changes over time at both the institution and population levels.

Data Sources

This study will rely on the following data sources: NSF administrative data on PIs, proposals, reviewers, and award decisions, the National Center for Science and Engineering Statistics (NCSES) for nationally representative survey data on the characteristics of the scientific workforce, and information on institutional characteristics from the Integrated Postsecondary Education Data System and Carnegie Classification of post-secondary institutions. The study will also rely on data from interviews, focus groups, and listening sessions with PIs and institutional representatives, and information gathered from a survey of relevant literature.

Challenges and Mitigating Strategies

NSF anticipates challenges in identifying a definitive set of characteristics to conduct descriptive analyses of target and comparator institutions and to construct wellmatched comparison groups. NSF will analyze the quality and comprehensiveness of administrative and survey data and devise approaches to fill in data gaps as needed. Proposed studies will also place burden on respondents asked to participate in interviews or focus groups and NSF anticipates challenges with obtaining adequate response rates to enable robust inferences. NSF will seek to address these challenges by collaborating with stakeholders to develop approaches that rely on existing data, leverage moments when respondents have strong incentives to provide information, and clearly communicate benefits of participation.

Use and Dissemination

Findings will be used by NSF to inform and improve program design. A well-designed and evidence-based capacity assessment and accompanying indicators can be used to measure progress of the initiative and assist institutions and NSF prioritize where to invest GRANTED resources.





To what extent did the COVID pandemic influence the participation of different groups, including those that have historically been underrepresented or underserved in STEM, in the NSF portfolio of programs and activities as compared to pre-pandemic?

Strategic Goal

Empower: Empower STEM talent to fully participate in science and engineering (SG1)

Strategic Objectives

Ensure accessibility and inclusivity (SO1.1)

Learning Agenda Alignment

FY 2022-1 (Missing Millions), FY 2022-2 (COVID Pandemic)

Background and Rationale

The COVID-19 pandemic disrupted NSF operations. In mid-March 2020, the agency transitioned to remote work and cancelled in-person activities, including panels through which thousands of proposals (more than 40,000 yearly) are peer reviewed to receive funding recommendations. NSF grantees also experienced disruptions. Some institutions reported closing laboratories or limiting field work, which affected research conducted by faculty, researchers, post-docs, and students. NSF-supported facilities were affected as well; for example, needed resources could not be deployed to some facilities due to travel restrictions. Concerns about COVID-driven disruptions on the scientific enterprise—and on the careers of those most at risk (such as early career and female scientists)—were voiced at NSF and beyond (Cui, Ding, and Zhu 2021; NASEM 2021; Myers et al. 2020, Morgan et al. 2021). These included warnings of grant applications delayed, papers left unwritten, and research careers stalled, particularly among groups underrepresented in science, technology, engineering, and mathematics. NSF used administrative data to monitor key indicators (such as proposals received by gender) and leveraged its deep community connections to hear from external stakeholders regarding problems encountered and strategies used to address them. What emerged was a complex picture that requires careful assessment. Disruptions seemed to have led to both negative and positive outcomes. For instance, the switch to virtual work disrupted in-person panels but also opened the door for increasing reviewer diversity through remote panels (by removing the barrier that travel may represent for some, such as scientists with caregiver responsibilities or with disabilities that make traveling difficult). Building a deeper understanding of this complexity is an important step in developing or revising interventions to (1) address any inequities that may have been exacerbated or introduced during the pandemic, (2) reinforce positive outcomes observed, and (3) prepare for future disruptions.

Timeline FY 2024





Technical Approach

A descriptive analysis of the characteristics of different groups in NSF's portfolio over time will be conducted. This will include the characteristics of principal investigators (PIs) and co-PIs submitting proposals and of reviewers participating in panels or conducting ad hoc reviews—overall, by directorates and offices, and by whether proposals were awarded or declined. This exploratory work will facilitate analyses of data and the specification of regression models as part of an interrupted time-series (ITS) design to determine changes that might be attributed to COVID—by modeling (and comparing) the expected pre-COVID and observed since-COVID trends, controlling for relevant factors. The qualitative component will rely on information gathered through semi-structured interviews with NSF program officers (POs), PIs, and reviewers. Once collected, these gualitative data will assist in the interpretation of guantitative findings, and model specification (to ensure important relationships are not overlooked) and understanding of relevant factors (positive and negative) that influenced participation in NSF's portfolio since the onset of the pandemic. If helpful for programming decisions, interview findings may be used to design a survey to be administered to a representative sample of PIs/ reviewers to estimate the influence of different factors on participation in NSF's portfolio of programs.

Data Sources

This study will rely on the following data sources: NSF administrative data (on PIs, proposals, reviewers, and award decisions), NCSES (for nationally representative survey data on the characteristics of the scientific workforce), the Integrated Postsecondary Education Data System and Carnegie Classification of post-secondary institutions (for information on the characteristics of institutions of PIs and reviewers), and interview data (from POs, PIs, and reviewers).

Challenges and Mitigating Strategies

This study faces at least three limitations related to existing data quality, methodological assumptions, and respondents. (1) The share of reviewers providing information on their demographic characteristics has been declining over time, which limits NSF's ability to produce valid and reliable estimates and tease out whether changes observed are due to changes in the composition of individuals in our data (resulting from missing data) or to changes in participation. The degree to which this data is missing likely reduces the ability to infer important relationships and obtain accurate estimates. NSF will attempt to mitigate this challenge by using imputation techniques to replace missing data and conducting sensitivity analyses to test the extent to which key findings change as different techniques or methods are used. (2) A key assumption of the ITS design is that pre-COVID trends would have continued unchanged and that no other external factors systematically affected the groups of interest during the post-COVID period. During interviews, we will seek to determine if these assumptions are reasonable and, if not, identify relevant factors to adjust analyses accordingly. (3) Devising a sampling strategy that enables us to identify a group of POs, PIs, and reviewers to interview (to obtain the insights we are looking for) and that agree to participate in this study will be challenging. We will work closely with NSF POs and develop a sample with appropriate replacement cases.

Use and Dissemination

Findings will be shared with NSF stakeholders to inform programming and policy decisions to address inequities and promote the inclusion of underrepresented groups in STEM. As permitted, they will also be disseminated to other Federal Government Agencies that have similar programs.





National Science Foundation <u>www.nsf.gov</u>



Spring sunset at Palmer Station, Antarctica. Credit: Mindy Piuk