About

The National Science Foundation (NSF)

NSF was created “to promote the progress of science; to advance the national health, prosperity, and welfare; to secure the national defense...” (1950, as amended).

NSF seeks to achieve these goals through an integrated strategy that advances the frontiers of knowledge; cultivates a world-class, broadly inclusive science and engineering workforce; expands the scientific literacy of all citizens; builds the nation’s research capability through investments in advanced instrumentation and facilities; and supports excellence in science and engineering research and education.

NSF is committed to evaluating the efficacy and efficiency of its strategy, leveraging evaluation to help the agency achieve its mission. Evaluations and other evidence-building activities conducted or supported by NSF are expected to adhere to NSF’s Evaluation Policy.

The Evaluation and Assessment Capability (EAC) Section

EAC bolsters NSF efforts to make informed decisions and promote a culture of evidence. Located in the Office of Integrative Activities of the Office of the Director, EAC provides centralized technical support, tools, and resources to conduct evidence-building activities and to build capacity for evidence generation and use across the agency.

Questions?

Please contact Clemencia Cosentino, Chief Evaluation Officer, at eac@nsf.gov.
Introduction

The Foundations for Evidence-Based Policymaking Act of 2018, Public Law No. 115-435 (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 2023 Annual Evaluation Plan (AEP) that NSF developed in response to this opportunity and following guidance provided by the Office of Management and Budget (OMB M-21-27, OMB M-19-23, OMB M-20-12, and OMB Circular No. A-11).

This AEP describes the evaluations prioritized by NSF for FY 2023. This includes evaluations that NSF is planning to begin or continue in FY 2023. Section 1 presents the criteria used for selecting them. Section 2 provides the research questions guiding each evaluation. Section 3 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 adherence to NSF’s Evaluation Policy.
Acknowledgments

NSF gratefully acknowledges the contributions of a wide range of stakeholders who were consulted or otherwise participated in the preparation of this Annual Evaluation Plan.

NSF Leadership and Staff
Leadership and staff from all NSF Directorates and Offices joined brainstorming sessions, helped prioritize learning questions, and drafted/reviewed the plans to answer those questions.

Federal Government Agencies
NSF consulted with other government agencies with similar investment portfolios to assess the merits of the questions, formulate technical approaches to answer the questions, and determine the questions’ potential for generating evidence that is useful for other agencies.

Other Stakeholders
NSF consulted with evaluators and researchers across multiple sectors—in academia, private and philanthropic organizations, and state and local government—and solicited input from the public through a request for information published in the Federal Register.
The following are five criteria used to select evaluation questions:

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.

2. **Have leadership support** — to prioritize the staff time and commit the resources that the work demands.

3. **Have potential to support upcoming decisions** — are likely to yield actionable and useful evidence in a timely fashion.

4. **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.

During NSF’s initial phase of Evidence Act implementation, 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 (COVID-19, government shutdowns, and delays in appropriations).
Section 2
Evaluation Questions At A Glance - FY 2023

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

COVID pandemic
In what ways did the COVID pandemic influence the participation of different groups in the NSF portfolio of programs and activities, such as merit review?

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

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

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?

*NSF plans to initiate these evaluations in FY 2022.
Questions are presented in alphabetical order.
Section 3
Evaluation Plans - FY 2023

This section includes a brief study plan for each prioritized evaluation question. They 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.
In what ways does the Convergence Accelerator Innovation Training contribute to the emergence of new capacities among participating researchers to meet pressing societal needs?

**Strategic Goal**  
Impact: Benefit society by translating knowledge into solutions

**Strategic Objectives**  
Deliver benefits from research  
Lead globally

**Guiding Question**  
How can NSF mobilize knowledge most effectively to impact society?

**Background and Rationale**  
The NSF Convergence Accelerator is a unique organizational structure within NSF that was initiated in FY 2019. The Convergence Accelerator seeks to (1) accelerate the transition of use-inspired convergence research into practice and (2) build team capacity to pursue exploratory, high-risk projects in topics that vary yearly. One of the signature approaches of the Convergence Accelerator that distinguishes it from other NSF efforts is the training the program provides to grantees to prepare them to transition their research ideas into investment-ready deliverables. This training is important for the success of the program in achieving its goals. This study seeks to determine in what ways and to what extent the curriculum developed for the program and the training provided using this curriculum helped teams acquire capabilities (attitudes and skills) that promote the Convergence Accelerator program's goals of building team capacity to transition research ideas into market-ready investments.

**Timeline**  
FY 2022–FY 2023

**Technical Approach**  
This study focuses on the FY 2022 cohort of Convergence Accelerator grantees and has two components to study training outcomes associated with program participation. The first component is a quantitative analysis of changes in grantees' understanding and, if possible, application of design thinking, team management, partnership development, and strategic communication concepts and practices, as these are the focus of Convergence Accelerator training. The analysis will be based on data collected through pre- and post-training surveys completed by participants. The second component will be based on a qualitative analysis of how artifacts evolved over time and may demonstrate how teams' research ideas are refined, packaged, and delivered after exposure to the Convergence Accelerator curriculum with grantee participation in trainings. This component of the study will be based on a comparison of the proposals submitted by grantees in Phase I versus Phase II and the oral pitches delivered as part of the Phase II competition. To conduct this comparison, we will develop and apply a rubric that aligns elements of grantees' work with program learning objectives.
### Data Sources
This study will rely on the Convergence Accelerator training material (agendas, presentations, workbooks, and other materials); grantee proposals, annual reports, and final deliverables/reports; pre- and post-training surveys of participants; and pitch videos. Convergence Accelerator instructors and coaches will be interviewed as sources for information about instrument development and testing.

### Challenges and Mitigating Strategies
Two main challenges stand out for this study. The first is the potential for low survey response rates, based on early experiences. To address this challenge, NSF plans to motivate participants by increasing their understanding of the importance of responding to surveys. Convergence Accelerator staff will also seek to revise the solicitation and award letters to make participation in evaluation activities a program requirement. The second challenge is construct validity and reliability of the rubric developed to analyze proposals and pitches. To mitigate this challenge, NSF will interview coaches and instructors for additional calibration of the rubric and train the analysts for using the rubric to ensure high inter-rater reliability.

### Use and Dissemination
Findings from this study will be shared with key NSF stakeholders and used to refine Convergence Accelerator’s grantees’ training.
In what ways did the COVID pandemic influence the participation of different groups in the NSF portfolio of programs and activities?

**Strategic Goal**
Engage: Empower STEM talent to fully participate in science and engineering

**Strategic Objective**
Ensure accessibility and inclusivity

**Guiding Question**
How can NSF help grow STEM talent and opportunities for all Americans?

**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 over the impacts of these 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 underrepresented minorities 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 2022–FY 2023
This evaluation will include quantitative and qualitative components. The quantitative component will begin with a descriptive analysis of the characteristics of different groups in NSF’s portfolio over time. This will include the characteristics of principal investigators (PIs) and teams 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 through a difference-in-differences approach (to measure differences in measures, such as proposals submitted by gender before and after the pandemic) 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 qualitative data will assist in the interpretation of quantitative 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.

This study will rely on the following data sources: NSF administrative data (on PIs, reviewers, proposals, panel reviews, and award decisions), the National Center for Science and Engineering Statistics (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 principal investigators and 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. NSF will attempt to mitigate this challenge by conducting sensitivity analyses to test the robustness of findings and use imputation techniques where possible. (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.
How do EPSCoR program funding strategies (infrastructure, co-funding, and outreach) contribute to increasing academic research competitiveness across jurisdictions?

### Strategic Goal
Discover: Create knowledge about our universe, our world, and ourselves

### Strategic Objectives
- Advance the frontiers of research
- Enhance research capability

### Guiding Question
How can NSF fuel transformative discoveries most effectively?

**Background and Rationale**
As its name indicates, the Established Program to Stimulate Competitive Research (EPSCoR) seeks to foster sustainable improvements in research and development (R&D) capacity in the 28 jurisdictions (states and territories) that individually received 0.75 percent or less of total NSF funding over the most recent five-year period. The EPSCoR program employs three investment strategies: (1) it supports physical, human, and cyber infrastructure in academic institutions through its Research Infrastructure Improvement funding tracks; (2) it co-funds meritorious proposals reviewed by other NSF programs that also satisfy EPSCoR programmatic criteria; and (3) it promotes interaction within the EPSCoR community and NSF through workshops and other outreach activities that help build mutual awareness and develop areas of potential strength. The program’s theory of change asserts that EPSCoR jurisdictions have opportunities to use EPSCoR funds and other available resources to improve their science, technology, engineering, and mathematics (STEM) ecosystems by strengthening academic research competitiveness—that is, the research competitiveness of the academic institutions in their jurisdictions. 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.

**Timeline**
FY 2023–FY 2025

**Technical Approach**
This outcomes evaluation will build on prior work, such as an exploratory study completed in FY 2020, to develop a design that helps NSF determine whether and how EPSCoR, through its different funding tracks, may be associated with observed project outcomes. The technical approach will be developed once background work is
Technical Approach cont'd
completed and may include analyses overall and by funding track, such as (1) descriptive analyses of jurisdictional characteristics, outputs, and outcomes to determine variation in characteristics and progress in implementation and outcomes over time; (2) a regression analysis of longitudinal data on EPSCoR jurisdictions (most likely done using a lower unit of analysis, such as participating institutions) to establish associations between observed outcomes and program participation, controlling for other factors that are known or hypothesized to be associated with outcomes; and (3) case studies of former EPSCoR program jurisdictions (or those nearing graduation or improving their research competitiveness) to understand the strategies that enabled them to increase their research competitiveness.

Data Sources
This study will rely on a monitoring data system that will be 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.

Use and Dissemination
Findings from this study will be shared with EPSCoR NSF program officers, grantee universities, and jurisdiction science and technology steering committees to inform decisions that may influence the ARC of institutions and jurisdictions.
How can NSF help increase the participation of underrepresented groups in the STEM workforce?

**Strategic Goal**
Engage: Empower STEM talent to fully participate in science and engineering

**Strategic Objectives**
Ensure accessibility and inclusivity
Unleash STEM talent for America

**Guiding Question**
How can NSF help grow STEM talent and opportunities for all Americans most equitably?

**Background and Rationale**
The National Science Board’s (NSB) report, *Vision 2030*, notes that “women and underrepresented minorities remain inadequately represented in S&E relative to their proportions in the U.S. population.” NSF awards more than $1 billion to broadening participation programs each year. These include programs focused on broadening participation, and programs that support research that contributes to these efforts by engaging students, post-docs, and early career faculty. Programs also vary in the strategies used to broaden participation—including scholarships, fellowships, mentorships, research experiences, and other interventions targeting individuals, teams, networks, and institutions. NSF has evaluated some of its efforts (examples include the quasi-experimental evaluations of the Louis Stokes Alliances for Minority Participation and the Graduate Research Fellowship Program) and evaluation activities will continue throughout the years of NSF’s new Strategic Plan as specific research questions are developed. These questions will guide further studies that contribute useful evidence that helps NSF bolster the efficacy of its initiatives to broaden participation and reduce inequities in how it delivers programs to its communities.
Continued...

**Background and Rationale cont’d**

NSF will pursue a series of studies designed to answer specific research questions, which might include the following: What intersectional groups are extremely underrepresented in STEM, and why? How could NSF leverage tools at its disposal—policies, strategies, programs, and so on—to increase the participation of these (most extremely underrepresented intersectional) groups in the STEM enterprise? What are the characteristics and, among individuals, educational and workforce outcomes of beneficiaries of NSF workforce development programs? What are the impacts of NSF policies and programs on the diversity of the STEM workforce and the participation of the most underrepresented groups? What changes to current NSF policies and programs might further catalyze improvements in the participation of extremely underrepresented groups in the STEM enterprise? What does success look like? Answers to these questions will help NSF identify best practices and align programs and policies toward closing gaps in participation in the STEM enterprise.

**Timeline**

FY 2022–FY 2026

**Technical Approach**

Technical approaches will be developed once the results of ongoing studies are available (such as the ongoing evaluations of the ADVANCE program and the Emerging Frontiers in Innovation Research Experience and Mentoring program) and new questions are finalized. NSF might pursue foundational studies (1) to further diagnose the problem of underrepresentation in STEM (and develop targeted interventions) and (2) to understand the characteristics of beneficiaries from NSF’s portfolio of investments and (3) determine the success of current NSF strategies and programs in achieving their goals equitably. More specifically, next steps may include the following: (1) an analysis that helps NSF identify extremely underrepresented groups (as characterized by intersectional characteristics, such as disabled women of color) and diagnose barriers to their participation; (2) a systematic review of broadening participation approaches used by NSF or emerging from the scholarly/policy literature to inform decisions regarding the portfolio of strategies that NSF will pursue; (3) a meta-analysis of existing evaluations related to NSF investments in broadening participation to identify the most impactful strategies leading to equitable outcomes and gaps in knowledge; and (4) additional evaluations with well-matched comparison group designs to measure the causal impacts of NSF programs and contribute useful knowledge to guide agency efforts to dismantle barriers to equitable participation in the STEM enterprise.
## Data Sources

Studies will rely on the following data sources: NSF administrative records (including annual and final reports and existing monitoring data systems to identify individuals), the National Center for Science and Engineering Statistics (for nationally representative survey data), the Integrated Postsecondary Education Data System and Carnegie Classification of post-secondary institutions (for information on the characteristics of institutions), the National Student Clearinghouse (for educational outcomes data), and individuals who participate in data collections (such as students, postdoctoral research fellows, university administrators, and principal investigators (PIs) surveyed or interviewed).

## Challenges and Mitigating Strategies

NSF anticipates challenges in identifying participants and nonparticipants and obtaining data on their characteristics to conduct descriptive analyses and construct well-matched comparison groups. NSF will rely on its data systems and national data, analyze the quality of existing data, and devise approaches to fill in data gaps, such as collecting demographic and prior achievement information through collections conducted as part of the new studies.

Proposed studies will also place burden on respondents asked to participate in surveys, interviews, or focus groups. NSF will seek to collaborate 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. A related challenge will be obtaining adequate response rates from participants and nonparticipants to enable robust and causal inferences. NSF will draw on its extensive experience recruiting respondents to devise appropriate strategies for each respondent group.

## Use and Dissemination

Findings will help NSF describe, reduce, and address barriers to full participation by updating programs and policies, identifying best practices to consider adopting, and aligning efforts to broaden participation of groups underrepresented in STEM. As appropriate, findings will also be shared with the NSB, Committee on Equal Employment Opportunity in Science and Engineering, communities implementing NSF-funded programs (such as PIs), beneficiaries of NSF programs, and the public.
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?

**Strategic Goal**
Impact: Benefit society by translating knowledge into solutions

**Strategic Objectives**
Deliver benefits from research
Lead globally

**Guiding Question**
How can NSF mobilize knowledge most effectively to impact society?

**Background and Rationale**
Building partnerships is a high priority for NSF, as evidenced by two consecutive agency Priority Goals (APGs for FY 2020 and FY 2021) focused on developing a partnerships strategy. The importance of partnerships is echoed in the recent National Science Board’s Vision 2030 report and reflected in the new Directorate for Technology, Innovation, and Partnerships (TIP) proposed in the NSF FY 2022 budget request. Partnerships can accelerate discovery in several ways. They can expand the kinds of questions that can be addressed, enable access to expertise and infrastructure, and expand communities of researchers. NSF engages in two types of partnerships—direct and indirect. Direct partnerships are established by NSF with other federal agencies, industry, private foundations, non-governmental organizations, and foreign science agencies. Indirect or “NSF-stimulated” partnerships are required or encouraged by NSF and established by principal investigators (PIs) on NSF grants seeking collaborators with complementary expertise or resources. These types of partnerships are common in many NSF programs, such as the Established Program to Stimulate Competitive Research, and can vary greatly in their characteristics. Having acquired deep experience in building, managing, sustaining, and ending partnerships, NSF is prioritizing evaluation activities that complement other ongoing learning efforts (such as conducting a landscape study) to reap the greatest benefits from partnerships. This study will be the second of several conducted to learn about the efficacy of NSF’s partnership strategy and identify ways to improve it.
This study will rely on the design developed in FY 2021 to begin evaluating NSF partnerships by studying direct partnerships with industry through the Directorate for Computer and Information Science and Engineering (CISE). NSF selected this type of partnership for the first evaluation for several reasons. Partnerships with industry are a priority for NSF and those in CISE (1) account for a substantial share of existing partnerships (for example, six of the seven new industry partnerships in FY 2019 were in CISE), (2) have sufficient cohorts of grantees to support retrospective or prospective evaluations, and (3) may have comparable non-partnership programs that could be used in support of a more rigorous (quasi-experimental) design to evaluate measurable outputs and outcomes. This study will also rely on qualitative analyses—such as analyses of interviews with partners and grantees—to uncover the benefits of partnerships and the barriers and facilitating factors to successful implementation (from the perspective of participants). These analyses will identify opportunities for improvements and dissemination of promising practices. NSF will use findings from the quantitative analyses to select samples of partners and grantees for surveys and/or interviews to ensure that NSF is able to tease out factors that are likely associated with successful partnerships.

Data sources will be determined after the design is completed and are likely to include NSF administrative data and documents (such as grantee annual and final reports), data on productivity (publications, patents, funding raised, startups launched, and so on), and surveys and interviews with different stakeholders (such as partners and grantees).

Two potential challenges stand out. The first is related to the complexity of creating a high-quality data file with information across programs, years, and data sources. The design phase of this project will enable NSF to devise a data strategy. The second challenge is methodological, as many factors stand in the way of effective evaluation of investments in basic science, such as long timelines to observe outcomes. In the design phase, NSF will identify opportunities to employ designs that enable causal inferences and identify cohorts for which outcomes can reasonably be expected by the time of this study.

Findings will be shared with NSF leadership and program officers. They will be used for program improvements and to inform the design of evaluations of other types of partnerships.