National Science Foundation
Merit Review
For a Changing Landscape
2025

National Science Board

Table of Contents

Memorandum from the Chair of the National Science Board

In the National Science Board (Board) report *Vision 2030*, which showcases our vision for U.S. science and engineering in 2030, we emphasized that delivering benefits from research to society is essential if the United States is to remain the world leader in basic research and innovation. In that report we called for a reexamination of the Board's Merit Review policy to ensure that the U.S. National Science Foundation's (NSF; Foundation) award portfolio furthers the national interest.

The Board established the Commission on Merit Review to ensure that NSF's award decisions remain the best in class – **based on merit** – and that the resulting portfolio of awards accomplishes NSF's mission to promote the progress of science, advance the national health, prosperity, and welfare, and secure the national defense.

This report, the product of extensive study, deliberation, and consultation, arrives at a moment of transformation for the Foundation, as it responds to evolving national priorities and a changed science and engineering landscape, where the achievements and contributions of other countries and sectors of the U.S. economy are more important than at any time since NSF's founding 75 years ago. The recommendations herein provide a map for building a next generation NSF, one that is increasingly nimble, efficient, and impactful.

First, and most importantly, this report emphasizes that taxpayer-funded research and STEM education must both advance science and engineering *and deliver benefits to Americans*. Second, it elevates the need to strategically build and track a portfolio of awards that aligns with NSF's mission and the priorities charted by Congress and the President.

This report presents recommendations for an updated Merit Review policy comprised of Principles, Criteria, and Considerations. NSF is now charged to implement the Board's updated policy, to be rolled out as soon as practicable.

Victor R. McCrary
Chair, National Science Board

Memorandum from the Chair of the National Science Board Commission on Merit Review

In December 2022, the National Science Board (NSB) established the NSB Commission on Merit Review to "assess the efficacy of the current Merit Review policy and associated criteria and processes at supporting the U.S. National Science Foundation's (NSF; Foundation) mission to create new knowledge, fully empower diverse talent to participate in STEM, and benefit society by translating knowledge into solutions."

Our two-year examination – which included expert deliberations, literature and legislative reviews, data collection, and analyses – revealed that NSF's merit-based approach to awarding funds, with its two Merit Review criteria, remains the best in class. As represented on the cover of this report, we conceptualize those criteria, Intellectual Merit and Broader Impacts, as a Möbius strip – two "sides" that are one continuous and intrinsic "surface" of merit. Yet some researchers and reviewers continue to perceive that benefits to society are less important in determining proposal quality than intellectual merit. Furthermore, there is an opportunity to better track and report the specific outcomes of NSF-funded research.

Broader Impacts, which we propose to rename Societal Benefits, is not an optional or bonus criterion. It is intrinsic to identifying projects worthy of federal funding and has been repeatedly codified by Congress. This public return on investment – whether by increasing American economic competitiveness, supporting national security, developing a globally competitive STEM workforce, or expanding STEM opportunities for all Americans – is a distinguishing feature of publicly-funded research.

Public funding also uniquely requires public transparency. As the nation's principal funding agency for all fields of basic research and STEM education, it is imperative that **outcomes**, reflecting the advancement of knowledge for the benefit of society, are effectively communicated to stakeholders. A nimble response to the changing landscape and national priorities requires a renewed focus on creating a portfolio of awards that reflects agency and national goals and tracking the portfolio's outcomes.

Our recommendations and guidance to the Foundation offer opportunities to strengthen the public trust with greater transparency, accountability, and clarity, and to build a next generation NSF whose portfolio of awards delivers on all facets of its mission.

Wanda E. Ward
Chair, NSB Commission on Merit Review

Executive Summary

In 2025, the U.S. National Science Foundation (NSF) and the National Science Board (NSB) celebrate 75 years of catalyzing world-changing discoveries and supporting generations of science and engineering (S&E) talent. Merit Review is central to NSF's past and future success.

Continued U.S. S&E leadership requires that NSF's Merit Review policy and processes remain agile and adaptable to a rapidly changing S&E landscape, emerging national priorities, and the enhanced expectations of Congress and the American people. In this context, the NSB Commission on Merit Review (Commission) was charged to re-examine the current NSB policy on NSF's Merit Review to make recommendations for revisions to ensure that "the Merit Review criteria, process, and reporting are delivering both new knowledge and societal benefits."[1]

The Merit Review policy and associated implementation processes govern how to evaluate proposals in S&E. The Commission found three fundamental strengths:

- A competitive process that relies on expert review is the best method to award funds.
- The NSB's Merit Review criteria, currently called **Intellectual Merit** and **Broader Impacts**, are appropriate and sufficient to identify proposals that advance NSF's mission.
- NSF's program directors are integral to the Merit Review process and should be empowered in decision-making and portfolio construction to the maximum extent possible.

The Commission also found opportunities to improve Merit Review. To maintain public trust, clarify the process for all participants, and demonstrate NSF's value, the Commission encourages the NSF and NSB to:

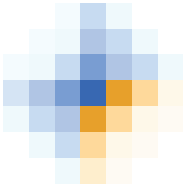
- Continue to fund only the highest quality proposals across all fields of S&E: there must be no compromise on merit
- In accordance with statute, deliver societal benefits from NSF-funded science and engineering research and education, including enhanced national security and economic competitiveness
- Intentionally and transparently build a portfolio of awards that covers NSF's intellectual and societal mission across all fields of science, technology, engineering, and mathematics research and education, in a manner that gives NSF the ability to adapt rapidly to opportunities in the science and technology ecosystem and be responsive to policymakers, national priorities, the research community, and the public
- Encourage investigators to embrace in their proposals the full range of societal benefits, including those outlined in statute as well as agency and national strategic goals
- Seek proposals and expert reviewers from a wide range of institution types, geographic areas and communities, economic backgrounds, and employment sectors, as permitted by law
- Use evidence-based practices to continuously improve the Merit Review process.

1 See Charge to the NSB-NSF Commission on Merit Review, December 2022

Based on these findings, the Commission developed the following policy recommendations to the Board:

1. Pursue renaming the "Broader Impacts" criterion to "Societal Benefits"
2. Maintain the two current review criteria, presently named Intellectual Merit and Broader Impacts, refine the definitions of each criterion, and clarify the use of additional review considerations
3. Emphasize the imperative to both promote the progress of science and engineering and deliver societal benefits from NSF-funded research and education, in accordance with statute
4. Strengthen the emphasis on expanding participation, including (but not limited to) institution type, geography, demographics, field of expertise, and sector of employment, as permitted by law
5. Revise NSB's Merit Review policy to clarify the alignment of award portfolios with agency strategy and outcomes
6. Update Merit Review policies and processes to strengthen transparency and accountability and promote continuous improvement.

With the recommendations and guidance in this report, the Commission seeks to ensure that proposers, reviewers, and NSF staff give **full and documented consideration of both criteria** throughout the Merit Review process. Transparent evaluation of each criterion will ensure NSF fulfills its vital mission to promote the progress of science in service to the nation.



Roadmap

CHANGED LANDSCAPE

NEW LAWS & DIRECTIVES

COMMISSION CHARGE

BOARD MEETINGS

EXTERNAL REPORTS

The work of the Commission involved the analysis of multiple sources (pg. 27-30)

Commission Charge
External consulations and listening sessions
Committee of Visitor (COV) reports and other analyses
Analytical questions for public RFI & NSF staff / leadership collection

DATA CONSULTED AND ANALYZED

FINDINGS

The findings encapsulate themes and ideas distilled from Commission work inputs and environment (pg. 14-18)

The Commission's work generated 6 recommendations (pg. 19) that produced updates to the Merit Review policy (pg. 20-23). In addition, the Commission generated a set of 24 ideas for NSF and NSB on the implementation of those recommendations, marked as guidance (pg. 24-26)

RECOMMENDATIONS

High level directions reflective of findings that the Commission desires to strongly communicate

REVISED MERIT REVIEW POLICY

Updated policy, incorporating recommendations
Proposed pending NSB consideration and a formal public process

GUIDANCE

Ideas for implementation of recommendations into Merit Review process

The Merit Review Commission was established because of a number of factors that also shaped the entire workstream (pg. 36-38)

Report

Merit Review for a Changing Landscape 2025

Why Re-examine Merit Review? And Why Now?

The U.S. National Science Foundation (NSF; Foundation) serves the American people with a clear statutory mission: enable U.S. discovery and innovation across all science and engineering (S&E) fields, secure the national defense, improve the nation's health and economic competitiveness, and educate and train the science, technology, engineering, and mathematics (STEM) workforce of the future.

The NSF Statutory Mission

To promote the progress of science; to advance the national health, prosperity, and welfare; to secure the national defense; and for other purposes.

NSF meets its mission through Merit Review. Merit Review is the process used to select proposals for funding by having experts assess their quality and potential impact. It involves evaluating both the intellectual merit (potential to advance knowledge) and broader impacts (potential to benefit society) of the proposed research. Through this rigorous and competitive process, NSF identifies projects to fund that align with its mission to promote scientific progress and contribute to national goals. Through Merit Review, the Foundation also supports the nation's most creative and innovative scientists, engineers, and students. These scientists and engineers advance fields ranging from cyberinfrastructure to ocean sciences, advanced manufacturing to artificial intelligence, and semiconductors to biotechnology (see Sidebar: NSF Success Stories).

The National Science Board (NSB; Board) is statutorily responsible for establishing the policies of NSF.[2] To maintain the Foundation's high standard for evaluating and recommending proposals for funding, NSB periodically conducts comprehensive examinations of Merit Review policy and processes. These reviews allow the Board to assess how effectively the Merit Review policy, resources, and NSF culture advance knowledge and deliver societal benefits.

Furthermore, the Board is charged to advise Congress and the President on S&E research and education policy matters.[3] NSF is key to U.S. success as a global leader in S&E. As such, NSB must also assess Merit Review in the context of the dynamic global S&E ecosystem. Responding to a rapidly changing S&E landscape demands agility. Thus, Merit Review should not be static – there must be continuous process improvements, while maintaining its core strengths.

Rigorous stewardship of federal funds requires transparency and accountability. While NSF has steadily improved in both, the NSB recognizes opportunities to update Merit Review to address remaining gaps. In addition, the S&E community (see Appendix G[4]), NSF Committees of Visitors (COVs), and Congress, most recently with the CHIPS & Science Act and the America COMPETES Reauthorization Act of 2010 (see Appendix C for relevant excerpts), have all articulated challenges in understanding the Broader Impacts criterion's importance and application. The Board has also highlighted this issue in *Vision 2030*[5] and at numerous Board meetings.[6] A reassessment of Merit

2 National Science Foundation Act of 1950, Pub. L. No. 81-507 (1950)
3 Ibid.
4 These include reports from the Committee on Equal Opportunities in Science and Engineering (CEOSE), Advancing Research in Society (ARIS), and the National Academies of Sciences, Engineering, and Medicine (NASEM).
5 National Science Board, *Vision 2030* (Alexandria, 2020). Excerpt from section entitled Deliver Benefits from Research, states that NSB should "Evaluate how NSF's broader impacts merit review criterion could better meet societal needs."
6 For example: NSB's Committee on Oversight, February 22, 2022 meeting (conversations around the Merit Review Digest, Broader Impacts (BI) categorization and how to assess impacts of BI); Committee on Oversight (CO) July 29, 2022 meeting (discussed CO goals, which included: (1) Merit Review Digest continued review to ensure that NSF's merit review process is fair, competitive, and transparent, reflects the Board's Vision, and enables meaningful assessment of NSF's progress towards achieving agency goals. (2) Broader Impacts. The Committee will continue to explore

Review presents opportunities to revise evaluation criteria to address these issues.

Since the last assessment of Merit Review in 2010-2012,[7] changes in the global and national S&E landscape,[8] in law, and within NSF itself,[9] also present new opportunities. For example, evidence-based policymaking has become a greater focus of the federal government including through the establishment of the U.S. Commission on Evidence-Based Policymaking in 2017 and the Foundations for Evidence-Based Policymaking Act in 2018.

In response to its statutory obligations – and to address major changes since its last Merit Review assessment – the Board established on December 1, 2022, the NSB Commission on Merit Review (Commission) *"to assess the efficacy of the current Merit Review policy and associated criteria and processes at supporting NSF's mission to create new knowledge, fully empower diverse talent to participate in STEM, and benefit society by translating knowledge into solutions"* (see Appendix D for the full Charge to the Commission).

Three broad questions emerged as Commission members considered revising the Merit Review process and which motivated its work.

Does NSF's Merit Review process provide sufficient transparency to meet the expectations of NSF stakeholders and to facilitate strategic, evidence-based decisions within the Foundation?

Beyond helping the Foundation demonstrate to Congress how it is fulfilling its mission (see text box), transparency and accountability at multiple levels – project, program, agency, and interagency – help NSF build trust with all stakeholders, including the S&E community and the public. Transparency and accountability also improve federal program efficiency and effectiveness. They help answer questions such as: Does NSF collect, analyze, and accurately report the data most useful for evaluating outcomes from NSF-funded grants and awards? How do the Foundation's strategic goals and award portfolio support the nation's S&E research and STEM education priorities?

the possibilities for capturing and communicating NSF-funded societal impacts of research. (3) Re-examination of NSF's Merit Review Process. The reexamination will take into consideration the Board's vision for Developing STEM Talent and Delivering Benefits from Research, and a commitment to ensure the criteria and process continue to meet the needs of our nation and achieve agency goals).

7 National Science Board, "National Science Foundation's Merit Review Criteria: Review and Revisions," (Arlington, VA: National Science Board, 2012)

8 National Science Board (NSB), "A Changed Science and Engineering Landscape," (Alexandria, VA: National Science Board, 2024)

9 For example, the establishment of the Directorate for Technology, Innovation and Partnerships Translation, Innovation, and Partnerships (TIP) at NSF through the CHIPS & Science Act.

On evidence-building, transparency, and accountability

Transparency is the clear, consistent, thorough, and accessible communication of information regarding the characteristics, evaluation, and funding of proposals, outcomes of awards, as well as the assessments and outcomes of award portfolios at the program, directorate, and agency levels. Transparency provides accountability – defined here as showing taxpayers that awards made by NSF contribute to the advancement of knowledge and benefit society.

Over the past decades and across multiple Administrations (examples below), the Legislative and Executive branches have required increased public access to information demonstrating the responsible use of taxpayer dollars and the federal government's responsiveness to national priorities.

"The Administration is committed to building and using evidence to improve policy, program, budget, operational, and management decision-making. Our vision for effective and efficient government includes ensuring accountability for results, having the necessary analytical tools to measure outcomes and impacts, identifying and investing in effective practices, and transforming data into evidence that informs action. With stronger evidence, we can learn from and improve programs to better serve the American people." (Office of Management and Budget: Evidence and Evaluation 2017-2025)

AICA Sec 102(b)(1): "The Director of the Foundation shall issue and periodicaly update, as appropriate, policy guidance for both Foundation staff and other Foundation merit review process participants on the importance of transparency and accountability to the outcomes made through the merit review process."(2017)

Does the Merit Review process help NSF strategically build a portfolio that covers a wide range of intellectual and societal benefits, including those areas mandated by statute?

The Merit Review process is key to building a world-leading award portfolio accomplishing the breadth of the Foundation's mission. Portfolios are NSF's mechanism to align its funding and strategic priorities with results for the nation.

Active, transparent, and strategic portfolio construction and management at the program, directorate, and agency levels provide the Foundation's leadership with a command of the full range of NSF-funded research and education and its societal impacts. Maintaining active and transparent portfolios allows the Foundation to respond more quickly to requests for information from stakeholders, including Congress and the White House, as well as to detect emerging opportunities and respond with agility as national priorities change.

Portfolios

A portfolio is a collection of awards that share one or more characteristics. Portfolio is used in several contexts related to NSF Merit Review:

Award Portfolio: A collection of awards resulting from the Merit Review process and categorized by a shared characteristic such as funding opportunity, organizational unit or level, topical focus, or approach to specific research and education objectives

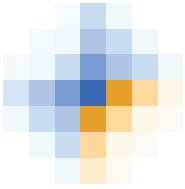
Program Portfolio: A collection of programs or funding opportunities categorized by a shared characteristic such as the managing organizational unit, topical focus, or proposer eligibility

Portfolio Balance: A mix of activities across both Merit Review criteria within a portfolio that has the potential to produce intellectual and societal benefits in alignment with disciplinary and/or program objectives, NSF strategic goals, and national priorities.

Does the Merit Review process – including decisions about what constitutes Intellectual Merit and Broader Impacts – reflect the full breadth of S&E expertise and relevant expert perspectives from across communities and sectors?

Ultimately, Merit Review aims to advance S&E research and education across all fields to serve the American people. As such, we must understand the breadth of participants in and beneficiaries of NSF-funded activities. On the beneficiary side, it includes answering: Who conducts research? Who assesses the outcomes? Who determines the scientific and societal impacts of NSF-funded research and education? On the process side, this includes answering: Who identifies which problems are worth pursuing and solving? Who is involved in solicitation development and reviewer selection?

Answering these questions well and implementing the resulting recommendations and policy revisions will ensure that NSF responds to evolving national priorities and new opportunities to advance the progress of science and engineering. In a rapidly changing S&E landscape, it is vital to U.S. global leadership that NSF stays at the forefront of discovery, learning, and innovation while providing unprecedented transparency and accountability to the American taxpayer.



NSF Success Stories

For 75 years, NSF funding has driven American innovation, touched every American, and changed the world. Here are examples of how NSF-funded research has transformed lives, improved U.S. economic prosperity, and strengthened national defense:

- **Building the internet:** NSF (along with the Department of Defense) helped develop early networks like NSFNET, the foundation for the modern-day internet. NSF programs supported two graduate students who developed a novel algorithm for ranking webpages – an innovation central to the company they founded, Google. These breakthroughs made today's digital world possible – pumping trillions of dollars annually into the U.S. economy, generating millions of jobs, and impacting global communication, work, and learning.

- **Creating advanced medical technology:** NSF-funded basic research in physics, engineering, computer science, and biological sciences laid the groundwork for developing Magnetic Resonance Imaging (MRI). In the 1970s, scientists – partially funded by NSF – developed the MRI imaging technique. This critical medical technology saves lives through early detection of conditions like strokes and heart abnormalities, and through surgical planning and treatment monitoring of diseases. NSF-funded research continues to expand this technology's impact and build next generation MRI machines.

- **3D printing:** 3D-printing technologies originated in the 1970s and 1980s, when NSF-supported researchers began exploring new ways to make products. Today, 3D printing is used for additive manufacturing to build parts for our military and American businesses, for patient care in implants, prosthetics, and drug design, and even for food.

- **Accessible communication:** NSF funding helped create the first dictionary of American Sign Language, a crucial resource for learners and the Deaf community. It also supported research contributing to language-learning apps like Duolingo, making language education more accessible globally.

- **From discoveries to applications:** NSF support for the Laser Interferometer Gravitational Wave Observatory (LIGO) led to the groundbreaking detection of gravitational waves in 2015, opening a new way to study the universe. LIGO's development of ultra-sensitive instruments and data analysis also spurred innovations in medical imaging (critical for disease diagnostics and treatments) and big data analysis.

Visit the *NSF Impacts website* to learn more about how NSF leads in discovery and innovation and delivers benefits from research.

Methods

The Commission's 18-month fact-finding process included literature reviews (including internal and external reports), discussions with NSF leadership, internal NSF surveys and interviews, a public Request for Information (RFI), listening sessions, and focus groups. The Commission consulted numerous experts and constituencies, including many external NSF stakeholders, to understand their perceptions of the Merit Review policy and its implementation. Appendix F provides a comprehensive timeline of Commission activities, including links to video recordings of public meetings. The Summary of Data Collection and Analysis presents a synopsis of the Commission's information-gathering and analysis, with full details in Appendices G-J.

Anchored by NSF's mission, the Commission assessed

- The Board's Merit Review (MR) policy
- How the Merit Review policy is understood and implemented by proposers, reviewers, program directors, and others
- How and to what extent the current Merit Review policy is successful in fulfilling NSF's statutory mission.

A number of themes emerged during this assessment. Based on the Commission's analysis of these inputs, its perspective on the changing geopolitical landscape for S&E, and the dynamic nature of science and engineering itself, this section of the report distills the key strengths of Merit Review and areas for improvement into eight Findings. The Findings inform the Policy Recommendations to the NSB (presented in the next section), which in turn shape the Proposed Merit Review Policy Language and the Guidance on the Application of the Revised Merit Review Policies found in the subsequent sections of this report.

Findings

Based on its extensive study, deliberation, and consultations, the Commission presents three significant Findings about the strengths of the current Merit Review policy and process, knowledge and skills necessary to compete.

FINDING 1: A competitive Merit Review process that relies on expert review continues to be the best in class for awarding funds for research and education in fields of science and engineering. It serves as an exemplar for other entities and nations.

FINDING 2: The two current Board-approved review criteria – presently named Intellectual Merit (IM) and Broader Impacts (BI) – remain appropriate to identify S&E research and STEM education proposals that advance NSF's mission.

FINDING 3: NSF program directors (PD) are a vital connection between the Foundation and the science and engineering community and are integral to the review process. Merit Review is strengthened when PDs are empowered in decision making and portfolio construction to the maximum extent possible.

Building on this solid foundation, the following five Findings identify opportunities to improve Merit Review and enhance NSF's agility to respond to a changing S&E landscape.

FINDING 4: NSF can improve the evaluation of proposals in ways that inform strategy and improve accountability. Two examples:

- Transparent and separate evaluation of each Merit Review criterion (IM and BI) could improve holistic assessments of proposals and improve evidence-based assessments of how NSF is meeting its statutory requirements.
- Expanding the pool of expert reviewers by drawing from a wider array of sectors, communities, geographies, and institution types would help to reflect the full range of perspectives on the intellectual and societal benefits of S&E research and education.

The Commission emphasizes that NSF must not compromise on merit; it must continue to fund only the highest quality proposals. At the same time, it is important to underscore that a broad range of potentially meritorious S&E research and education activities exist beyond those identified through a traditional academic lens.

Today, industry funds almost as much basic research as the federal government,[10] more nations are joining the S&E discovery and innovation race, and taxpayers expect scientific and technological solutions to urgent and emerging problems. This new S&E landscape requires partnership among government, academia, industry, and philanthropy to harvest the fruits of discoveries in fundamental S&E, speed the translation of research and education to societal benefits, nurture STEM talent, support cross-disciplinary research, foster STEM-based economic development nationwide, and

10 National Center for Science and Engineering Statistics, "InfoBrief: Analysis of Federal Funding for Research and Development in 2022: Basic Research," (Alexandria, VA: National Science Foundation, 2024)

develop critical and emerging technologies. To stay at the forefront of discovery *and* deliver benefits to the nation, it is essential to seek out a wider range of perspectives and expertise to provide input on what science questions to ask, what problems are urgent to solve, and what constitutes "excellent science" and "a societal benefit" (Figure 1). Expanding the pool of participants could help to enhance trust in the Merit Review process and broaden multi-sectoral support for the Foundation itself.

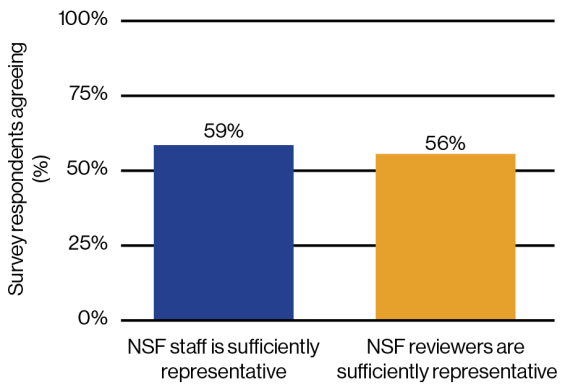


Figure 1: Just over half of NSF staff agreed that staff (blue, n=222) and reviewers (gold, n=224) represent a sufficiently representative range of individual and institutional characteristics to achieve the goals of the Merit Review process. (Appendix J, p. 130, exhibit 13.2).

FINDING 5: Documented construction and management of award portfolios above the level of individual programs would help inform agency strategy and improve Foundation accountability in the delivery of intellectual and societal benefits.

NSF does not have a singularly applied definition of "portfolio"; at the program level, staff are instructed to consider multiple factors when recommending proposals for funding.[11, 12] Input from NSF leadership and staff via surveys (see Appendix J) and discussions between NSF leaders and the Commission make clear the thoughtfulness and effort that goes into portfolio construction at the program level.

The Commission found opportunities for a more strategic and structured approach to manage award portfolios above the program level. Active portfolio construction and management above individual programs (e.g., at Directorate and agency-wide levels) would improve NSF's ability to anticipate emerging opportunities and more readily respond as national priorities shift. Higher-level portfolio construction and management would also enable NSF to demonstrate alignment of NSF funding with the agency's strategic plan and national needs. Consistent, documented, and accessible processes and outputs for portfolio construction and management should be available for internal NSF leaders and external stakeholders (to the extent appropriate) for evaluation and assessment. This would allow for more strategic management of NSF's award portfolio, including improved multisectoral partnerships and coordination across federal agencies.

FINDING 6: Strategic management of NSF's award portfolio requires establishment and tracking of consistent metrics, including improved and additional data to evaluate, track, and report on accomplished Intellectual Merit and Broader Impacts.

11 Factors include internal priorities (e.g. the NSF Strategic Plan, NSB's *Vision 2030*), external directives (e.g., Executive Orders, legislation, and Supreme Court rulings), sub-fields of inquiry, and the goal of supporting a variety of project sizes, approaches, and institutions.

12 Most COV reports address a series of questions that examine the resulting portfolio of awards, including items about the balance of sub-disciplines, size and duration of awards, the potential for innovative and transformative research, the balance of interdisciplinary research, geographical distribution, types of institutions, balance of new and early career researchers, whether awards integrate research and education, information about the equality of opportunity for participation by all, and relevance to national priorities, agency mission and other external needs. See NSF Performance: Committees of Visitors.

Data provide useful, cross-cutting information about funding allocation, particularly along dimensions beyond academic discipline or demographics. Such data are invaluable internally, as input to analysis and evaluation of award portfolios. Furthermore, data are a key aspect of agency accountability, demonstrating to Congress and the American people that NSF has an award portfolio that meets the full range of statutory intellectual and societal benefits.

NSF collects and reports significant data, including in the Merit Review Digest and public dashboards (see Example: NSF data dashboards). However, the Commission found that NSF needs improved agency-wide metrics for both Intellectual Merit and Broader Impacts to evaluate award portfolio outcomes at the program level and above. For example, 61% of NSF staff survey respondents reported that the data NSF collects are somewhat or very effective in assessing whether funded projects contribute to advancing scientific knowledge. Concerningly, only 39% of staff respondents reported the data collected are somewhat or very effective at helping to assess whether funded projects support NSF's mission to benefit society (Appendix J, p. 41).[13]

Example: NSF data dashboards

The NSF By the Numbers dashboard collates and publicly presents data historically released through static tables and summary reports. It allows users to generate tailored insights into proposal and award trends by NSF Directorate, institution characteristics, and geography. Similarly, the Translation, Innovation, and Partnerships (TIP) directorate has an interactive public dashboard showing geographic and institutional distribution of its awards.

Grant recipients also play a key role. It is important and would be valuable for the grant recipient to report how the grant fulfilled its original Intellectual Merit and Broader Impacts goals or provided other valuable outcomes. The intent is not to increase administrative burden. Instead, the Commission's goal is to ensure required reports and data collection are maximally useful in assessing the agency's award portfolio impact. To this end, it would be beneficial for NSF to evaluate its data collection and reporting requirements and remove any that do not significantly aid in assessing NSF-funded research and education outcomes.

FINDING 7: Revising and clarifying Merit Review policy and guidance would help the S&E community, including proposers and reviewers, develop a better understanding of the purpose and intent of the Broader Impacts criterion and the range of societally relevant goals it encompasses.

Since before the passage of the America COMPETES Reauthorization Act of 2010 and the publication of the U.S. National Science Foundation's Merit Review Criteria: Review and Revisions, the Board has promoted stakeholder understanding and use of NSB's Broader Impacts criterion to drive delivery of societal benefits from NSF's scientific research and education investments. Societal benefits are statutorily required. Yet challenges persist in interpreting and applying the Broader

13 See Exhibit 15.2 of the Mathematica report in Appendix J, p. 41. Percentage of NSF staff who think the data NSF collects about awarded projects are somewhat or very effective at helping NSF assess whether funded projects support NSF's mission to advance scientific knowledge (n=223) and benefit society (n=222).

Impacts criterion, often resulting in the perception that it is secondary in importance to Intellectual Merit. For example, respondents[14] to the Commission's Request for Information (RFI)"indicated that full consideration [of the two criteria] is not achieved due to factors such as...the undervaluing of Broader Impacts." Within NSF, 91% of staff survey respondents indicated that there are times when the criteria are weighted unevenly. Furthermore, 61% reported they place more weight on Intellectual Merit than on Broader Impacts, and staff indicated that it is less clear how to assess a proposal's merit for the Broader Impacts criterion (Figure 2; Appendix J, p. 10).

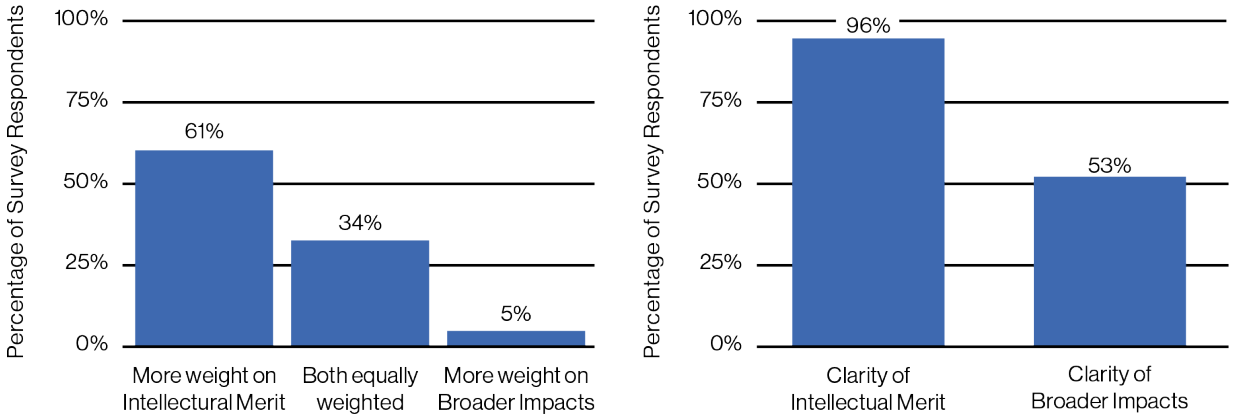


Figure 2: (Left) 61% of NSF staff survey respondents (n=204) reported they place more weight on Intellectual Merit than on Broader Impacts, and only approximately one-third of NSF staff report that they weight the two criteria equally. (Appendix J, page 92, exhibit 2.6). (Right) NSF Staff (n=232) also reported that nearly all staff (96%) feel it is less clear how to assess a proposal's merit for the Intellectual Merit criterion, but little more than half think it is clear how to assess the Broader Impacts criterion (Appendix J, page 84, exhibit 1.5).

In recent years, NSF has successfully conveyed to the S&E community the importance of attracting and developing domestic STEM talent. However, NSF should further emphasize the need to deliver other statutory societal benefits from NSF research and education, including economic competitiveness, national security, health, partnerships with industry, and public engagement and scientific literacy. Renaming the Merit Review criterion from Broader Impacts to Societal Benefits would likely assist in that effort.

FINDING 8: Expanding participation is valuable to both the conduct and the impact of science and engineering research and education. Improvements to Merit Review policy and processes could help increase participation in NSF-funded research and education from underrepresented regions and states and across sectors and institution types.

Talent is the treasure upon which the U.S. S&E enterprise rests.[15] The nation faces a STEM talent crisis threatening its security, economic prosperity, and global prominence in science and engineering, particularly in critical and emerging technology areas.[16,17] In a world where S&E is the

14 A Request for Information (RFI), published in the Federal Register and in an NSF Dear Colleague Letter, elicited input from a broad set of people and groups who participate in or are affected by the Merit Review process, including principal investigators (PIs), reviewers, academic institutions, organizations that submit proposals to NSF, and the general public. (Appendix J, p. ix)

15 NSB, "Talent is the Treasure," (Alexandria, VA: National Science Board, 2024).

16 Ibid.

17 Written Testimony of Daniel Reed, NSB Chair, "Oversight and Examination of the National Science Foundation's Priorities for 2025 and Beyond." (Research and Technology Subcommittee, Committee on Science, Space, & Technology, U.S. House of Representatives, May 16, 2024).

new currency of power,[18] our nation cannot afford to sideline domestic STEM talent. Data in *Science and Engineering Indicators* and other statistically robust reports[19] show the U.S. is not on track to have an S&E workforce that fully represents all Americans. As the White House stated in its 2018 *Strategy for STEM Education*, "all Americans deserve the chance to master STEM skills and methods, both for their own success and for America's competitiveness."[20] The Commission emphasizes that building an S&E workforce that reflects the U.S. population requires increasing the participation of individuals including those from rural areas,[21,22] lower socioeconomic backgrounds, and underrepresented regions and states,[23] as well as first-generation college students. Furthermore, talented individuals develop S&E knowledge and skills at various degree levels via a range of institution types, from research universities to community colleges and trade schools, and pursue careers across all S&E employment sectors. Jobs requiring S&E knowledge and skills (at all degree levels) pay more on average and are more resilient during economic downturns compared to non-STEM jobs.[24] In an S&E-driven world, it is a national imperative to equip Americans from all backgrounds and all regions with the STEM knowledge and skills necessary to compete.

The benefits of expanding participation in S&E go beyond individual success and national S&E competitiveness. It also brings intellectual breadth to the science itself.[25] Researchers and collaborations that consider and incorporate a wide variety of perspectives have a higher potential to create innovative methodologies,[26] drive curiosity,[27] and translate knowledge[28] – all of which are critical aspects of Intellectual Merit. In addition, bringing a broad range of voices into STEM enhances public trust in science and improves alignment of S&E research and education with societal needs to better the lives of all Americans.

In sum, Merit Review must evolve to realize an NSF award portfolio that uniquely contributes to science and engineering within a changing multisectoral and global S&E landscape, addresses current and emerging national priorities, and meets taxpayer expectations for progress and accountability. To demonstrate the current and potential future value of NSF investments to the nation, the Foundation must collect, analyze, and openly report data on awards and outcomes. A next generation NSF will require a next generation Merit Review that delivers intellectual and societal benefits from NSF-funded research and education, reinforcing appreciation of and support for NSF by Congress, the White House, and the American people. To this end, in the following sections, the Commission presents its recommendations and guidance to the Foundation for revising Merit Review.

18 NSB, "Connected Horizons."
19 NCSES, "Women, Minorities, and Persons with Disabilities in Science and Engineering," (Alexandria, VA: National Science Foundation, 2021)
20 Committee on STEM Education of the National Science & Technology Council, "Charting a Course for Success: America's Strategy for STEM Education" (December 2018).
21 Committee on Equal Opportunities in Science and Engineering "Making Visible the Invisible: STEM Talent of Rural America, 2024 CEOSE Report to Congress,"(Alexandria, VA; National Science Foundation, 2024).
22 National Academies of Sciences, Engineering, and Medicine. "K-12 STEM Education and Workforce Development in Rural Areas," (Washington, DC: The National Academies Press, 2024).
23 Committee on the Future of NSF EPSCoR, "Envisioning the Future of NSF EPSCoR," (Alexandria, VA; National Science Foundation, August 2022).
24 NSB, Science and Engineering Indicators 2024: The State of U.S. Science and Engineering. (Alexandria, VA; NSF, 2024).
25 Report of the Advisory Committee for GPRA Performance Assessment FY 2009 (NSF 09-068).
26 Page, S. E. Diversity Powers Innovation. Center for American Progress. (January 26, 2007).
27 Dixon-Fyle, S., Huber, C., Del Mar Martínez Márquez, M., Prince, S., & Thomas, A. Diversity matters even more: The case for holistic impact. McKinsey & Company (2023).
28 Creating and maintaining high performing collaborative research teams: the importance of diversity and interpersonal skills." Frontiers in Ecology and the Environment 12.1 (2014): 31-38.

Policy Recommendations to the National Science Board

The Commission developed and refined the following Merit Review policy recommendations for NSF and NSB based on its findings and data examinations. The Commission recommends that NSB and NSF implement these recommendations by translating them into NSF's relevant policy documents. The report's next section presents the Commission's revised Merit Review Policy.

In addition to the policy recommendations, the Commission identified opportunities to improve how NSF and NSB use the Merit Review policy. The Commission developed suggested guidance for NSF and NSB to consider when implementing the policy recommendations, including in agency strategic planning and budgeting. The section following the revised Merit Review policy language presents this guidance.

The Commission recommends that:

1. NSB pursue renaming the "Broader Impacts" criterion to "Societal Benefits"

2. NSB maintain the two current review criteria, presently named Intellectual Merit and Broader Impacts, refine the definitions of each criterion, and clarify the use of additional review considerations

3. NSB improve Merit Review policies to emphasize both promoting the progress of science and engineering and delivering societal benefits from NSF-funded research and education, in accordance with the NSF Act and other applicable statutes

4. Throughout Merit Review policy and processes, NSB and NSF strengthen the emphasis on expanding participation, including (but not limited to) institution type, geography, demographics, field of expertise, and sector of employment, as permitted by law

5. NSB revise its Merit Review policy to clarify the alignment of award portfolios with agency strategy and outcomes to improve transparency and accountability in the delivery of intellectual impacts and societal benefits

6. NSB and NSF update Merit Review policies and processes to strengthen transparency and accountability and promote continuous improvement

Proposed Merit Review Policy

The National Science Board establishes the agency's Merit Review policy, in consultation with NSF. The six policy recommendations developed and approved by the Commission, presented in the preceding section, informed these proposed revisions to the Merit Review Policy language.

Key Points about the New Policy Language

The Policy continues to consist of three sections: Merit Review Principles, Criteria, and Considerations (formerly "Elements").

The updated Principles emphasize transparent review processes that fully consider both Merit Review criteria. They take a broad approach to identifying scientific progress, societal benefits, and contributors to agency mission, while emphasizing appropriate project and portfolio evaluations to improve funding outcomes.

The Commission retains the existing Intellectual Merit and Broader Impacts Criteria but refines their definitions. The revised policy articulates the many dimensions of Intellectual Merit – including, for instance, reproducibility, research on expanding participation, and curiosity-driven research – while emphasizing the full scope of statutory Broader Impacts goals. The Commission's work suggests renaming Broader Impacts may contribute to a better understanding of what is meant by that criterion. The revised Policy clarifies that the separate, transparent, high-quality evaluation of each criterion is necessary input to a holistic proposal assessment and clarifies the role of solicitation-specific guidance.

The updated Considerations language removes "novelty" from both criteria to eliminate misconceptions that proposals for reproducibility studies and established methods are unwelcome. Based on NSF staff feedback that proposers and reviewers poorly understand Merit Review Considerations, as well as other components of Merit Review policy (Appendix J, Section II.4), this section has been renamed and reorganized to provide targeted direction to reviewers and NSF program staff, as appropriate.

This proposed policy reflects Commission recommendations and is not yet in effect as of the release of this publication. Policy updates require NSB consultation with NSF to ensure consistency with agency procedures, plus consideration of public comments. Once completed, the revised Merit Review policy will be incorporated into the Proposal & Award Policies & Procedures Guide (PAPPG) and published. NSF and NSB should work collaboratively to ensure expeditious implementation of the revised policy.

Merit Review Principles

In 1950, Congress charged the U.S. National Science Foundation "To promote the progress of science; advance the national health, prosperity and welfare; and secure the national defense; and for other purposes." Consequently, all activities NSF funds should both advance scientific progress and deliver societal benefits to the nation. NSF uses the Merit Review process to evaluate proposals to assess their potential to advance this dual mission. NSF consistently strives to conduct a fair, competitive, transparent Merit Review process for the selection of projects. For this process to succeed, all stakeholders benefit from a clear set of guiding principles, including policymakers, investigators, institutions, reviewers, and NSF program staff.

The following six Principles reflect NSF's statutory mission and obligation to taxpayers. They should be adhered to by PIs and organizations when preparing proposals and managing projects, by reviewers when reading and evaluating proposals, and by NSF program staff when determining whether to recommend proposals for funding and while overseeing awards.

- All awards and the full portfolio of NSF investments should reflect the highest standards of research quality across the full spectrum of Intellectual Merit and Broader Impacts.
- The national importance of NSF's mission demands that successful proposals have the potential to both advance the frontiers of science and engineering and provide benefits to society.
- Full, documented, and transparent consideration of both criteria is essential throughout the proposal preparation, review, and decision-making processes.
- Diverse viewpoints, including institutional, geographic, demographic, disciplinary, and sector of employment, are critical to evaluating the potential impact of new or expanded knowledge and societal benefits.
- Investigators, reviewers, and NSF program staff must enable enhanced accountability for project funding through meaningful assessment and evaluation of projects based on appropriate metrics established in the proposal.
- Consistent and documented evaluation of portfolios at the Foundation-wide and Directorate levels, based on appropriate metrics for both criteria, is essential to communicating and assuring the value of federally-funded research.

Merit Review Criteria

All proposals should be written by investigators, evaluated by reviewers, and decided on by program staff using these two criteria:

- **Intellectual Merit:** the potential to create new or expand existing knowledge
- **Broader Impacts:** the potential to benefit society.[29]

Collectively, NSF-funded projects must contribute to a range of intellectual outcomes. Individual project outcomes should include the creation of science and engineering knowledge, ranging from fundamental to applied research. Projects can be motivated by one or a combination of curiosity, potential end use, translation to application, transformative potential, robust examination of a risky hypothesis with the prospect of high rewards, or verification of previously reported results. Other critical aspects of funded research derive from the need to increase the nation's STEM talent reservoir, for example, through increasing understanding and exploring approaches to expanding participation in STEM, including (but not limited to) institution type, geography, demographics, field of expertise, and sector of employment, as permitted by law.

Simultaneously, NSF-funded projects must have the potential to benefit society – either through the near or long-term impacts of the proposed research itself, or from activities associated with the proposed research program. A wide range of societal benefits is identified in legislation including the

29 Policy Recommendation 6 directs NSB to discuss the benefits and costs of changing the name of this criterion to "Societal Benefits" or other descriptive name and raise a suggested change to Congress. If Congress changes the name of the Broader Impacts criterion to Societal Benefits or to another name that more accurately conveys the outcomes intended by the criterion, the agency will do so as well.

America COMPETES Reauthorization Act of 2010, the American Innovation and Competitiveness Act of 2017, and the CHIPS and Science Act of 2022. Individual project outcomes should include, but are not limited to: increasing U.S. economic competitiveness; advancing public health and welfare; supporting the national defense; enhancing partnerships between academia and industry; developing a globally competitive American STEM workforce; improving STEM education and instruction at all levels; use of science and technology to inform public policy; improving public scientific literacy; expanding participation in STEM; and enhanced infrastructure for research and education. Congress continues to emphasize the importance of considering potential societal benefits when NSF is selecting projects to fund.

These examples of intellectually and societally relevant outcomes are neither comprehensive nor prescriptive. Investigators may include other appropriate outcomes not covered by these examples.

Proposers must give **full consideration** to **both** criteria in their proposals. Reviewers and NSF staff must do the same during the review, decision-making, and award management processes. Each criterion is necessary, but neither is sufficient. Therefore, proposers, reviewers, and NSF staff should equally and transparently address both criteria. Reviewers will evaluate each criterion individually as input to a holistic proposal assessment. These evaluations will be considered by NSF program officers as input to funding decisions and by NSF leadership and Board members as input to the development of balanced award portfolios that address a range of Intellectual Merit and Broader Impacts objectives, in accordance with the Foundation's statutory mission.

Merit Review Considerations

NSF proposals should address the following: what is to be done, why it should be done, how will it be done, how will success be identified, what benefits could occur if the project is successful, and what opportunities may be lost if the project does not go forward. These considerations apply to both criteria. NSF may, when necessary, employ supplementary guidance to emphasize aspects of the criteria in certain programs and activities. However, supplementary guidance does not constitute additional criteria.

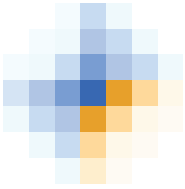
For both criteria, reviewers and NSF program staff should consider the following:

1. The potential for the proposed activities to
    a. Create new knowledge or expand existing knowledge, including through high risk / high reward, potentially transformative, curiosity-driven, expanding participation, and replication studies (Intellectual Merit); and
    b. Benefit society in ways beyond knowledge impacts, including current statutory goals (42 USC § 1862p-14) (Broader Impacts)
2. The extent to which the plans for the proposed activities are well-reasoned, well-organized, cohesive, adaptable, and feasible
3. The inclusion of clear mechanisms and measures in the proposal to assess and report outcomes.

4. Evidence of the potential of the individual, team, or organization to conduct the proposed activities, including:
    a. Relevant expertise;
    b. Demonstrated achievement of anticipated and/or unanticipated outcomes from prior support, if applicable; and
    c. Availability of adequate resources (either at the home organization or through collaborations).

For both criteria, in addition to the above, NSF program staff should consider the following (individually for each proposal and in the aggregate for award portfolios):

1. The potential for the proposed activities to contribute to a balanced award portfolio that includes a range of Intellectual Merit and Broader Impacts, in accordance with NSF's statutory mission and program objectives
2. How the proposal may contribute to assembling a balanced award portfolio with a mix of awardees, including but not limited to consideration of career stage, institution type, and geographic location.



Guidance on the Application of the Revised Merit Review Policies

During its deliberations, the Commission identified strengths and opportunities to improve NSF's use of the Merit Review principles, criteria, and considerations. Upon adoption by NSB and NSF, the guidance below is intended to enhance responsiveness to stakeholders, address national priorities, and improve Merit Review's efficacy in delivering benefits to the nation.

Because applying the suggested guidance is primarily the responsibility of NSF leadership and staff, the Commission encourages NSF to consider such guidance in its ongoing strategic planning and budgeting. The Commission also encourages NSB to consider proposed guidance in its governance roles, and where appropriate, work with NSF to develop and implement reasonable, feasible timelines. NSF and NSB should work collaboratively to ensure effective progress toward implementing this proposed guidance.

The proposed guidance for applying the policy recommendations will help the agency in three ways:

- Inform agency strategy
- Demonstrate agency accountability to immediate stakeholders in the Merit Review process (*e.g.* proposers, reviewers, NSF staff) and external stakeholders (*e.g.* Congress, the White House, the S&E community, and taxpayers and the broader public)
- Facilitate continuous improvement and clarification of the purpose, process, and implementation of Merit Review.

Informing agency strategy

The Commission advocates for improvements to the Merit Review process–including how proposers, reviewers, awardees, and staff engage with Merit Review–that will enable the agency to better develop strategies to fulfill its mission in a complex and rapidly changing global S&E landscape.

NSF should consider the following actions:

- Develop mechanisms to empower, encourage, and recognize program directors (PD) for increasing funding allocated toward high-risk / high-reward research and potentially transformative research, and work with NSB on defining these categories and setting targets across award portfolios
- Develop a mechanism to document transparent and evidence-based evaluation of each review criterion including assessing how and to what degree each proposal addresses the requirements of each criterion, and report as required by law. Such a mechanism could be explored in a study/pilot program implementing criterion-specific review ratings
- Update and enhance the agency's information infrastructure to enable insight into awards and outcomes through systematic, standardized categorization of NSF-funded activities
- Publish periodic reports summarizing the results of external reviews, including Committee of Visitors (COV) reviews across NSF to provide a holistic picture of the state of NSF's Merit Review process, as recommended by the internal 2013 report on NSF's Use of the Committee of Visitors Mechanism. These reports should highlight trends in COV recommendations that may elicit agency-level action
- Seek regular advice from external experts and community driven mechanisms to provide formal, documented advice on Directorate- and agency-level portfolio goals and strategy.

NSB should consider the following actions:

- Pursue strategies to expand participation in S&E research and education to engage individuals from a range of institutional, geographic, sectoral, and demographic backgrounds in all aspects of Merit Review, as permitted by law.

Accountability to stakeholders

The Commission believes NSF must provide better evidence that it is accomplishing its mission to the S&E community, Congress, the White House, and, crucially, the public—the source of the agency's funding and the principal beneficiaries of NSF-funded research—through enhanced documentation and publicly-accessible data.

NSF should consider the following actions:

- Implement and document the impact of efforts to broaden the reviewer pool, including (but not limited to) institutional, geographic, demographic, field of expertise (including criterion-specific expertise), and sector of employment, as permitted by law
- Update project final reporting mechanisms to require NSF-funded PIs to report: criterion-specific goals, the degree of success achieving those goals, criterion-specific impacts, and insights into potential future impacts. Such updates may include introducing new mechanisms to capture long-term outcomes beyond the award lifecycle
- Identify and implement effective policies and mechanisms NSF-wide to ensure safe and harassment-free research and education environments
- Publish data on criterion-specific themes as proposed in funded awards—including those defined in statute—in NSF by the Numbers and the Merit Review Digest
- Communicate publicly the factors considered when balancing award portfolios.

NSB should consider the following actions:

- Reimagine reports on Merit Review to include proposal and award data, as well as information on reviewer pool characteristics, including (but not limited to) institutional, geographic, demographic, field of expertise (including criterion-specific expertise), and sector of employment, as permitted by law
- Continue to work with NSF on developing IM and BI outcome metrics and appropriately publishing of outcomes, in alignment with agency strategy and NSB's Vision and goals
- Regularly assess NSF progress toward implementing proposed Merit Review policy revisions.

Improving and clarifying the Merit Review process

The Commission encourages NSF to make the Merit Review process and decision-making as transparent as possible to stakeholders. NSF and NSB should also continue to pursue continuous improvement and innovation of the Merit Review process through piloting and expeditious implementation of new mechanisms and protocols and regular study, evaluation, and analysis of Merit Review outcomes.

NSF should consider the following actions:

- Clarify that "additional review criteria" are not equivalent to IM or BI, and when used, expressly emphasize program-specific aspects of IM or BI pertaining to a solicitation
- Develop an internal repository of additional guidance utilized in solicitations to help identify common themes and terminology across NSF funding opportunities
- Continue to pilot and implement mechanisms to further strengthen BI, such as:
  - Raise the quality of BI criterion written reviews to a level similar to IM criterion written reviews
  - Update panel discussion practices and protocols to ensure full consideration is given to both Board-approved criteria and document the impacts of these efforts
- Enhance BI and IM criterion-specific expertise (including community and multi-sectoral partners) among reviewers to assess the full spectrum of proposed activities
- Periodically examine and report on the interpretation and application of the BI and IM criteria in proposals, evaluations, decisions, and outcomes
- Building on existing efforts, revise training to augment program director preparedness for actively ensuring fairness in review panels and other components of the Merit Review process
- Continue to identify and pilot mechanisms that could expand access and increase opportunities for proposers to participate in the Merit Review process, propagating successful approaches throughout the agency, and regularly reporting to NSB on outcomes of these efforts
- Regularly assess and update, as necessary, practices to more efficiently process and evaluate proposals and more rapidly inform proposers of funding decisions.

Example: CMMI reviewer training

NSF's Civil, Mechanical, and Manufacturing Innovation (CMMI) Division in the Engineering Directorate (in partnership with the Kardia group) has developed an immersive, cohort-based professional development program for future CMMI research review panelists. This program builds group behavior, conflict resolution, and other skills to generate more robust and fair review processes. Over 70% of participants report having used the skills from this program in NSF panel reviews.

Example: Modernizing Research.gov

Recent upgrades to Research.gov have significantly combined access to NSF Merit Review functions under a single account and self-managed user profile. This change improves user profile quality and consistency. This will help NSF to better understand PI and Reviewer characteristics—from demographics to institutions to expertise—and reduce burden by eliminating duplicate information collection.

NSB should consider the following actions:

- Communicate the breadth of research avenues that contribute to advancing knowledge. Such a list may include, but is not limited to basic research, curiosity-driven research, translational research, research on expanding participation, STEM education research, potentially transformative research, and replication studies
- Advocate for flexibility and reduced burden in reporting requirements at the interagency level and collaborate with NSF in developing innovative approaches to awardee reporting.

Summary of Data Collection and Analysis

To inform its recommendations on NSF's Merit Review policy and process, the Commission's fact-finding process included open and closed meetings and meetings with internal and external experts. The Commission conducted surveys and interviews within NSF and issued a public Request for Information (RFI). The Commission conducted an extensive literature review including internal and external reports. It also held a variety of listening sessions and focus groups.

The Commission took care to consult with a broad range of constituencies to understand their perceptions of the Merit Review policy and associated criteria and their implementation. Specifically, the Commission gathered input from discussions with NSF's Executive Leadership Team, via surveys of NSF staff and interviews of NSF leadership, public comments from key external stakeholders, a longitudinal review of past Committee of Visitors (COV) reports from 2012 to 2023 and conducting focus group interviews.

Data Collection

The Commission contracted with Mathematica Inc. to design and implement a plan to gather input from key constituencies on the Merit Review criteria, process, and reporting mechanisms at NSF. Mathematica's data collections included surveys, interviews, focus groups, and a Request for Information (RFI). The survey of NSF staff included program directors, division directors, deputy division directors, deputy assistant directors, and deputy office heads. The study team emailed 472 NSF staff an invitation to participate in the web survey. During the four weeks of web survey administration, 241 NSF staff (51%) participated.[30]

Interviews with senior leadership included the NSF Director, NSB chair, and NSF assistant directors and office heads. Twelve leaders were invited to participate in a 60-minute interview conducted on Zoom. Nine (75%) responded and completed an interview; those who were unable to or did not wish to participate were asked to recommend another leader from their directorate for interview.[31] Mathematica also solicited input from individuals and groups external to NSF, including visiting personnel assigned to program director roles via the Intergovernmental Personnel Act, via a Request for Information posted in the Federal Register and disseminated in an NSF Dear Colleague Letter (DCL). The RFI consisted of six open-ended questions developed by the Commission. Responses were counted if they commented on at least one item on the RFI. RFI responses were collected from August 26 to September 20th, 2024, and the Commission received 130 responses, some of which were group responses. The volume of responses received was consistent with other NSF RFIs, especially considering the time available for response under the Commission's work plan.[32]

30 Staff from the following directorates or offices were invited to participate in the survey: Biological Sciences (BIO); Computer and Information Science and Engineering (CISE); Engineering (ENG); Geosciences (GEO); Mathematical and Physical Sciences (MPS); Social, Behavioral and Economic Sciences (SBE); Science, Technology, Engineering, and Mathematics Education (EDU); Technology, Innovation and Partnerships (TIP); Office of Integrative Activities (OD/OIA); and Office of International Science and Engineering (OISE).

31 The assistant director or office head from each of the following directorates and offices were invited: BIO, CISE, ENG, EDU, GEO, MPS, OIA, OISE, SBE, and TIP.

32 RFI published as a Notice in the Federal Register (Request for Information (RFI) on National Science Board-National Science Foundation Merit Review Commission Review of NSF's Merit Review Policy and Processes).

In addition, Mathematica held focus groups with vice presidents of research (VPR) or similar role at institutions of higher education and individuals who have served on an NSF Advisory Committee (AC) and Committee of Visitors. Focus group participants were sourced from 25 potential participants for the AC focus group and 15 potential participants for the VPR focus group. Two participants attended the AC focus group and four attended the VPR focus group.[33]

Data Analysis

The Mathematica study team conducted descriptive quantitative analyses of the responses to multiple-choice items on the NSF staff survey. After closing the web survey, the study team cleaned the survey data file and prepared the data for analysis. Data cleaning included range verification and skip logic verification to confirm that each variable only contained allowable values. As needed, the study team constructed variables to address small cell sizes. The study team calculated an overall survey response rate. Each record was categorized as a complete case, partial response, or nonresponse. The study team computed frequency tabulations for each multiple-choice item on the NSF staff survey. Select survey responses are reported by position type, directorate, gender, race, ethnicity, and disability status. We conducted significance testing by using a chi-square statistic on disaggregated results.

Descriptive qualitative analyses were performed on interview notes, focus group transcripts, and verbatim responses to the 15 free-response items from the staff survey and six free-response items from the RFI. Interview, staff survey, and RFI data were analyzed using framework analysis.[34] The framework analysis approach entailed first coding the data using a set of broad descriptive codes aligned to the guiding questions then assessing the coded data for data-driven themes. For each theme, the theme and the number of supporting cases is described along with quotes or examples that illustrate the supporting cases.

Select Findings from the Mathematica Report

The final Mathematica report surfaced several themes across responses from various audiences. Respondents were mixed on whether the Merit Review criteria are clear, with a greater proportion of respondents indicating that there was less clarity about the Broader Impacts criterion than the Intellectual Merit criterion, as well as noting ambiguity in how it is applied during review. Respondents indicated that Broader Impacts generally suffers from uneven weighting, unclear standards, training, and expertise when compared to Intellectual Merit and stated a need for more clarity and training. The report also explored perceptions of fairness and bias, noting that most respondents identified several potential areas that could introduce unfair bias. The report also noted that respondents said that leadership and staff have a high to moderate level of expertise in both Merit Review criteria, and work to identify reviewers with the proper level of expertise to assess the criteria, while fewer respondents said that reviewers and PIs have a high level of understanding of the criteria.

When considering portfolio development, reviewers' assessments of the criteria were factored into portfolio management decisions, but it was reported that some reviewers lacked necessary

33 This information collection was conducted with approval from the Office of Management and Budget under NSF's Generic Clearance for the Collection of Qualitative Feedback on Agency Service Delivery (approval no. 3145-0215).

34 Srivastava, A., and S.B. Thomson. 2009. "Framework Analysis: A Qualitative Methodology for Applied Research Note Policy Research." Journal of Administration and Governance, 4(2): 72–79.

knowledge and training. Respondents also indicated that final and interim reports were more effective at assessing the outputs and outcomes of funded projects in advancing scientific knowledge (Intellectual Merit) than in assessing benefits to society (Broader Impacts). Mathematica's findings were organized under a set of guiding questions that were developed throughout the data collection process. The full report may be found in Appendix J.

Literature Review

Numerous sources in the scholarly and policy literature informed the Commission's discussions about relevant work done pertaining to Merit Review and informed recommendations and suggestions. These sources were included in Commission materials (written documents and presentations), and included reports from NSB, NSF, the National Academies of Science, Engineering, and Medicine (NASEM), the Council of Canadian Academies,[35] the National Academy of Public Administration,[36] and the Committee on Equal Opportunities in Science and Engineering (CEOSE). Additionally, the Commission reviewed internal NSF policy documents, and an analysis of relevant legislative sources informed the Commission's deliberations. These sources are referenced throughout this report and can be found listed in the bibliography in Appendix G.

To help inform the Commission's efforts, NSBO staff compiled and read through select publicly-available COV reports produced between 2012 to 2023. NSBO staff reviewed 28 Committee of Visitors Reports produced between 2012-2023 (representing 25% of all COV Reports since the previous Merit Review Taskforce) for any issues raised by COVs related to the use of the Merit Review criteria. The findings of this report were discussed during a closed meeting of the Commission and the full report is included in Appendix K.

Analysis of Committee of Visitors (COV) Reports

All programs at NSF prior to 2025 were reviewed by COVs composed of external experts with NSF reviewer experience as part of NSF's larger performance assessment. NSF relies on these external experts to review the quality and integrity of the Merit Review process, operations, and technical and managerial matters pertaining to proposal decisions. COV reports are available upon request.[37]

To help inform the Commission's efforts, NSBO staff compiled and read through select publicly available COV reports produced between 2012 to 2023. NSBO staff reviewed 28 Committee of Visitors Reports produced between 2012-2023 (representing 25% of all COV Reports since the previous Merit Review Taskforce) for any issues raised by COVs related to the use of the Merit Review criteria. The findings of this report were discussed during a closed meeting of the Commission and the full report is included in Appendix I.

35 Council of Canadian Academies, The Expert Panel on International Practices for Funding Natural Sciences and Engineering Research, "Powering Discovery," (Ottawa, ON: Council of Canadian Academies, 2021)

36 National Academy of Public Administration (NAPA), "A Study of the National Science Foundation's Criteria for Project Selection," (Washington D.C.: National Academies of Public Administration, 2001)

37 COV records are available for routine inspection and copying, in accordance with 5 USC 1009(b). Please contact cov@nsf.gov for further information.

Stakeholder Input: Experts, Panels, Listening Sessions, and Focus Groups

During meetings of the Commission, members held conversations with a variety of other groups and external experts to inform the Commission's discussion and deliberations. Internal to NSF, the Commission benefited from multiple discussions with the NSF's executive leadership team and talks with internal Foundation experts. To gain an understanding of federal STEM agency review processes across the government, the Commission convened a panel discussion in August 2023 focusing on grant review policies and practices at other STEM federal agencies (National Institutes of Health, Department of Energy, NASA, DARPA, ARPA-H).[38] Commission members had the opportunity to question the panelists following their presentations, covering topics from resubmission policies and review approaches to management of portfolios and translational research.

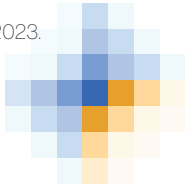
Commission members also held listening sessions to hear feedback from the broader community of S&E stakeholders and leaders to better understand the external perception of Merit Review and its implementation. These conversations helped frame the Commission's understanding of the stakeholder landscape when considering the full range of other data collected, though the conversations themselves were not used as evidence to make recommendations. Executive summaries of these discussions can be found in Appendix H.

In March 2023, Commission members held a listening session at the 2023 Advancing Research Impacts in Society (ARIS) Summit in Baltimore, MD. Over 120 Broader Impacts experts were present during the 90-minute listening session. This conversation focused on specific suggestions related to the Broader Impacts criterion and helped inform the Commission's discussions.

In June of 2023 and 2024, the Commission Chair and Vice Chair met with members of the Committee on Equal Opportunities in Science and Engineering (CEOSE) in a public discussion on Merit Review.[39] These discussions focused on recent CEOSE reports, the status of recommendations made to NSF regarding specific areas of interest, and engaging on specific suggestions related to the work of the Commission.

In January of 2024, Commission members held a focus group with 28 provosts, vice presidents of research, graduate school deans, sponsored research officers, and public engagement officials from institutions within the Big Ten Academic Alliance. This conversation focused on how these institutions support principal investigators in their efforts related to Broader Impacts and Intellectual Merit, and how those criteria are used in evaluating faculty for tenure and promotion.

38 The panel "Federal agency perspectives on grant review processes" occurred during an open meeting of the Commission on August 14, 2023.
39 Written accounts of these discussions are in the June 2023 and June 2024 minutes, found at https://www.nsf.gov/od/oia/ceose



Looking Ahead

The Merit Review process is key to building a first-class STEM research and education portfolio that encompasses the breadth of the NSF mission. It has been essential to the Foundation's success over the last 75 years. Setting NSF on a course for continued leadership in national and global science and engineering requires continuous improvement of NSF's processes, portfolios, and – most importantly – delivering the benefits and impacts of science and engineering research and education to the American people. Leadership in this effort requires adapting Merit Review policies and processes to meet the changing global landscape for S&E and responsive to emerging national priorities. The Commission has explored key areas in strategy, transparency, and accountability, and its findings lay a strong foundation for the next 75 years.

The Commission views this report as launching a collaborative process between NSF and NSB to update and enhance Merit Review. For example, this report recommends the clear alignment of NSF's award portfolios with agency strategy and project outcomes. Metrics to demonstrate this alignment will be tailored to the settings in which they will be applied. The Commission expects NSF and NSB to collaborate on developing metrics NSF can implement. For example, when appropriate metrics are in place, the resulting information could be used to better demonstrate to Congress how Merit Review outcomes align with national priorities and geographic distribution of funding.

The Commission also looks forward to NSF's developing data-driven, modernized training for NSF staff and reviewers. In addition, we encourage NSF and NSB to work together to improve proposer, reviewer, and staff understanding of both review criteria, particularly, what is encompassed within the full spectrum of Intellectual Merit and statutory Broader Impacts. Further, we encourage NSF and NSB to collaborate in identifying opportunities to reduce administrative burden. For example, for any new reporting or assessment requirement that is instituted, requirements to provide less valuable information could be removed so that the total reporting burden remains the same or decreases. NSB will also work with NSF on setting investment priorities for potentially transformative and high-risk-high reward research across appropriate award portfolios.

Publication of this report on NSF's and NSB's 75th anniversary offers an opportunity for reflection and recommitment to supporting science and engineering research and education that advances the progress of science and the prosperity and security of the United States. While the Foundation's mission remains the same, an ever-evolving S&E ecosystem demands a next generation NSF with a more nimble and transparent funding process that benefits from the full participation of and delivers benefits back to all American people.

Members of the National Science Board

Victor R. McCrary
National Science Board Chair
Vice President for Research and Professor of Chemistry
The University of the District of Columbia

Aaron Dominguez
National Science Board Vice Chair
Provost
Catholic University of America

Sudarsanam Suresh Babu
Clark Distinguished Chair
Professor of Materials Science & Engineering
University of Maryland, College Park

Roger N. Beachy
Professor Emeritus of Biology
Washington University, St. Louis

Joan Ferrini-Mundy
President
University of Maine

Yolanda Gil
Fellow and Research Professor
Information Sciences Institute, University of Southern California

Juan E. Gilbert
Andrew Banks Family Preeminence Endowed Professor
University of Florida Distinguished Professor
Chair of the Computer & Information Science & Engineering Department
University of Florida

Dorota Grejner-Brzezinska
Vice Chancellor for Research
University of Wisconsin-Madison

Melvyn Huff
Assistant Teaching Professor
University of Massachusetts-Dartmouth

Jeffrey Isaacson
Chief Executive Officer Emeritus
Universities Space Research Association

Matthew Malkan
Distinguished Professor of Physics and Astronomy
University of California, Los Angeles

Marvi Matos Rodriguez
Director of Engineering in Mobility & Surveillance for the P-8 Air Vehicle Integrated Product Team
Boeing Company

Willie E. May
Vice President for Research and Economic Development
Morgan State University

Sarah O'Donnell
Distinguished Chief Engineer
National Security Engineering Center

Ryan Panchadsaram
Advisor
Kleiner Perkins

Julia M. Phillips
Executive Emeritus
Sandia National Laboratories

Scott Stanley
Vice President of Technology
Techno Planet

Keivan G. Stassun
Stevenson Chair in Physics & Astronomy
Vanderbilt University

Merlin Theodore
Group Leader for the Advanced Fibers Manufacturing Group in the Manufacturing Science Division
Oak Ridge National Laboratory

Wanda E. Ward
Executive Associate Chancellor for the Office of the Chancellor – Public Engagement
University of Illinois Urbana-Champaign

Bevlee Watford
(Emeritus) Professor of Engineering Education
Associate Dean for Equity and Engagement
Virginia Polytechnic Institute and State University

Heather Wilson
President
University of Texas, El Paso

Members of the NSB Commission on Merit Review

Wanda E. Ward
Chair, NSB Commission on Merit Review
Executive Associate Chancellor for the Office of the Chancellor – Public Engagement
University of Illinois Urbana-Champaign

Julia M. Phillips
Vice Chair, NSB Commission on Merit Review
Executive Emeritus
Sandia National Laboratories

Roger N. Beachy
Professor Emeritus of Biology
Washington University, St. Louis

Steven Leath*
Former President
Auburn University
Iowa State University

Former Commission member

Matthew Malkan
Distinguished Professor of Physics and Astronomy
University of California, Los Angeles

Scott Stanley
Vice President of Technology
Techno Planet

Keivan G. Stassun
Stevenson Chair in Physics & Astronomy
Vanderbilt University

Stephen H. Willard*
Former Chair, NSB Commission on Merit Review
Chief Executive Officer
ICaPath Inc.

Acknowledgements

The NSB Merit Review Commission wishes to thank the many individuals who gave their time and thoughtful input to the Commission in the process of collecting information, prioritizing areas for analysis, and developing the report, including the policy recommendations and implementation guidance.

The Commission extends its gratitude to the numerous individuals who participated in listening sessions, interviews, and focus groups, and responded to our survey and our Request for Information.

The Commission wishes to acknowledge NSB and NSF leadership, including NSB Chair Victor McCrary, former NSB Chairs Darío Gil and Dan Reed, and former NSF Director Sethuraman Panchanathan.

The Commission thanks the initial Chair of the Commission Stephen Willard for his guidance during the first year of the Commission's work. He and former Board member Anneila Sargent were instrumental to launching the Commission in their capacity as Chairs of the Board's Committee on Oversight.

The Commission acknowledges the NSF program directors, staff, and leadership of Divisions, Directorates, and Offices who graciously shared their thoughts and expertise on the future of Merit Review, as well as a host of external reviewers who generously gave their balanced nonpartisan views and expertise. We are also grateful for the insights, analyses, and energy of John Adamec, Executive Secretary to the Commission.

The Commission also expresses its sincerest thanks to the National Science Board Office staff for their tireless support throughout the Commission's reexamination of Merit Review. Albert Einstein Distinguished Educator Fellows Danielle Taylor and Vida Treviño helped organize and staff our listening sessions. Einstein Fellow Michael Stewart played a leading role in our Committee of Visitors report analysis. American Association for the Advancement of Science S&T Policy Fellow Andrew Czeidinksi helped organize and analyze findings from listening sessions and contributed to the drafting and production of the report. Chris Blair, Brandon Powell, Michelle McCrackin, and Clint Lohse assisted with various parts of the Commission's work. DeMonica Fooks, Kathy Jacquart, Brad Gutierrez, Andrea Rambow, and Elise Lipkowitz provided administrative guidance and thoughtful advice. Nadine Lymn and Elizabeth Jeffers provided communications insights. Faith Hixson served as our Contracting Officer's Representative. Alexandra Surcel provided strategic support and data analysis throughout the Commission's work and contributed to the drafting and production of the report. Reba Bandyopadhyay played a leading role in the drafting, management, and production of the Commission report. Ann Bushmiller provided strategic, legal, and policy guidance throughout the Commission's examination. Portia Flowers served as senior liaison to the Commission and guided the steps of the Commission's examination development from start to finish with skill and dedication. Finally, we thank Executive Officer John Veysey for his wise counsel, unwavering clarity of purpose, and the many roles that he has played in ensuring the Commission's success.

Appendices

Appendix A: Brief History of Merit Review at NSF

Throughout its seven-decade history, technical and nontechnical elements have always been vital to the U.S. National Science Foundation's official proposal review criteria. The National Science Board has consistently aimed to improve and clarify review processes, criteria, and implementation.

In December 1951, one year after the establishment of the NSF, the initial criterion for proposal evaluation was "the scientific merit of proposed research, including the competence of the investigator."[40] However, additional criteria for external reviewers included: duplication of effort, reasonableness of budget, and quality of available personnel and facilities at the host institution. In 1967, the NSB approved a clarification of five review criteria for academic research: promise of scientific results; the potential scientific impact; the degree of novelty, originality, or uniqueness; the educational value of the proposed research; and the relevance of the proposed work to potential applications.[41] In 1974, in response to applied research becoming a growing share of NSF's research portfolio, NSB established 11 criteria across four categories: technical competence of the researcher and their institutional base; the internal structure of the science, with an emphasis on the possibility of impacting other disciplines; the utility or relevance of the research; and the long-term scientific potential (including impact on students, institutional structure, and practice).[42]

In 1976, Congress stated that NSB should have primary responsibility for establishing policies governing peer review.[43] NSB reduced the number of review criteria from 11 to 4 in 1981. [44] Out of concern that the increasing directed appropriations from Congress into federal funding of science and engineering facilities throughout the early 1980s could undermine the peer review process, NSB called upon NSF to re-examine its Merit Review process.[45] In 1986, NSF's Advisory Committee on Merit Review presented a set of recommendations reaffirming its review process, but called for improvements to the process and quality of reviews.[46] It also recommended a change from the term "peer review" to "merit review" to more accurately describe the agency's selection process.

Following mandates in the 1993 Government Performance and Results Act (GPRA),[47] NSF and NSB sought to link NSF goals and strategies to outcomes. In addition, findings from reviewer and program director surveys showed a lack of attention to and understanding of nontechnical criteria among reviewers. This led NSB to seek additional clarification and revision of review criteria.[48] Upon recommendation by the 1996-97 NSB-NSF Staff Task Force on Merit Review, the Board approved the peer review" to "merit review" to more accurately describe the agency's selection process.

Following mandates in the 1993 Government Performance and Results Act (GPRA), NSF and NSB sought to link NSF goals and strategies to outcomes. In addition, findings from reviewer and program director surveys showed a lack of attention to and understanding of nontechnical criteria

40 National Science Foundation (NSF), "Annual report, 1952," (Washington, DC: Government Printing Office 1952).
41 NSF, "Annual report, 1967," (Washington, DC: Government Printing Office 1967).
42 NSF, "Annual report, 1974," (Washington, DC: Government Printing Office 1974).
43 U.S. Congress, House, Subcommittee on Science, Research, and Technology report, Section E (1976), located in NSB report to House Subcommittee on Science, Research, and Technology Regarding Peer Review Procedures at NSF, Appendix A (Washington, DC: National Science Board 1977).
44 NSF, "Grant proposal guide (NSF-95-27)," (Washington, DC: National Science Foundation, 1995).
45 National Science Board, (NSB), "Report of the NSB Committee on Excellence in Science and Engineering," (NSB-85-50) (Washington, DC: National Science Board, 1985).
46 NSF, Advisory Committee on Merit Review (NSF 86-93) (Washington, DC: National Science Foundation, 1986).
47 Government Performance and Results Act, Pub. L. 103-62 (1993)
48 NSB, National Science Board and National Science Foundation Staff Task Force on Merit Review, Discussion report (NSB/MR-96-15), (Arlington, VA: National Science Board, 1996)

among reviewers. This led NSB to seek additional clarification and revision of review criteria. Upon recommendation by the 1996-97 NSB-NSF Staff Task Force on Merit Review, the Board approved the simplification of the four criteria with two new criteria – Intellectual Merit and Broader Impacts.[49] Several years later, in September 2004, Congress requested NSB to conduct a review of the NSF Merit Review process. [50] The Board conducted the review and issued its report in September 2005, concluding that the NSF Merit Review process is fair and effective, and "remains an international 'gold standard' for review of science and engineering research proposals."[51] In the report, the Board provided several recommendations for NSF to improve the transparency and effectiveness of the NSF Merit Review process, while preserving the ability of program directors to identify the most innovative proposals and effectively diversify and balance NSF's research and education portfolio. In response to the Board's recommendations, NSF implemented an agency-wide effort to address quality of reviews, transparency of the award/decline decision, and support of transformative research.[52]

The America COMPETES Reauthorization Act of 2010,[53] which became law in early 2011, directed NSF to apply the Broader Impacts criterion to achieve a specific array of societal goals. NSB established an NSB-NSF Task Force to examine both Merit Review criteria and their effectiveness in achieving NSF's goals in support of science and engineering research and education. A final report from the NSB-NSF Task Force, released in early 2012, concluded that while the Intellectual Merit and Broader Impacts criteria remained appropriate for evaluating NSF proposals, clarification and guidance in implementation and evaluation were recommended.[54]

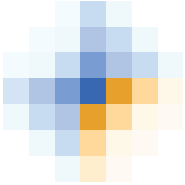
More recently, with the 2020 release of the Board's *Vision 2030* report,[55] NSB committed itself to evaluating "how NSF's broader impacts Merit Review criterion could better meet societal needs" and work in partnership with NSF leadership to "undertake an organizational review and offer recommendations on changes to directorate structure, funding models, or programmatic offerings, including for convergent research and questions inspired by societal problems." In 2021, the Board passed resolutions requesting the implementation of policies to improve reviewer preparedness[56] and facilitate Broader Impacts expertise on COV panels.[57]

Signed into law in August 2022, the CHIPS and Science Act directed NSF to conduct an external assessment of the application of the Broader Impacts criterion across NSF and make recommendations for improving the criterion's effectiveness in meeting legislative goals.[58] In December 2022, NSB established a special Commission on Merit Review to assess "the efficacy of

49 NSB, National Science Board and National Science Foundation Staff Task Force on Merit Review Final Recommendations, (NSB/MR-97-05) (Arlington, VA: National Science Board, 1997
50 H. Rept. 108-674 – Departments of Veterans Affairs and Housing and Urban Development, and Independent Agencies Appropriations Bill, 2005.
51 NSB, "Report of the National Science Board on the National Science Foundation's Merit Review System (NSB-05-119)," (Arlington, VA: National Science Board, 2005)
52 NSB, "Report to the National Science Board on the National Science Foundation's Merit Review Process Fiscal Year 2005," (Arlington, VA: National Science Board, 2006)
53 America Competes Reauthorization Act of 2010, Pub. L. No. 111–358 (2011)
54 NSB, "National Science Foundation's Merit Review Criteria: Review and Revisions (NSB/MR-11-22)," (Arlington, VA: National Science Board, 2011)
55 Vision 2030, https://www.nsf.gov/nsb/publications/2020/nsb202015.pdf
56 NSB, Resolution, "Training to Improve Peer Reviewing in Merit Review Process, (NSB-2021-10)," (Alexandria, VA: National Science Board, 2021)
57 NSB, Resolution, "Broader Impacts Experts to Serve on Committees of Visitors, (NSB-2021-11)," (Alexandria, VA: National Science Board, 2021)
58 CHIPS and Science Act, Pub. L. No. 117-167 (2022)

the current Merit Review policy and associated criteria and processes at supporting NSF's mission to create new knowledge, fully empower diverse talent to participate in STEM, and benefit society by translating knowledge into solutions."[59] The Commission completed its work and was officially discharged on March 10, 2025.

59 NSB, Major Actions and Approvals at the December 1-2, 2022 Meeting, NSB-2022-45, (Alexandria, VA: National Science Board, 2022)



Appendix B: Current and Proposed Merit Review Policy

CURRENT Proposal & Award Policies & Procedures Guide (NSF 24-1), Chapter III.A	PROPOSED Merit Review Policy
The U.S. National Science Foundation strives to invest in a robust and diverse portfolio of projects that creates new knowledge and enables breakthroughs in understanding across all areas of science and engineering research and education. To identify which projects to support, NSF relies on a merit review process that incorporates consideration of both the technical aspects of a proposed project and its potential to contribute more broadly to advancing NSF's mission "to promote the progress of science; to advance the national health, prosperity, and welfare; to secure the national defense; and for other purposes." NSF makes every effort to conduct a fair, competitive, transparent merit review process for the selection of projects.	In 1950, Congress charged the U.S. National Science Foundation "To promote the progress of science; advance the national health, prosperity and welfare; and secure the national defense; and for other purposes." Consequently all activities NSF funds should both advance scientific progress and deliver societal benefits to the nation. NSF uses the Merit Review process to evaluate proposals to assess their potential to advance this dual mission. NSF consistently strives to conduct a fair, competitive, transparent Merit Review process for the selection of projects. For this process to succeed, all stakeholders benefit from a clear set of guiding principles, including policymakers, investigators, institutions, reviewers, and NSF program staff.
Merit Review Principles	Merit Review Principles
These principles are to be given due diligence by PIs and organizations when preparing proposals and managing projects, by reviewers when reading and evaluating proposals, and by NSF program staff when determining whether or not to recommend proposals for funding and while overseeing awards. Given that NSF is the primary Federal agency charged with nurturing and supporting excellence in basic research and education, the following three principles apply:	The following six Principles reflect NSF's statutory mission and obligation to taxpayers. They should be adhered to by PIs and organizations when preparing proposals and managing projects, by reviewers when reading and evaluating proposals, and by NSF program staff when determining whether to recommend proposals for funding and while overseeing awards.

Merit Review Principles (cont.)	Merit Review Principles (cont.)
• All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge. • NSF projects, in the aggregate, should contribute more broadly to achieving societal goals. These broader impacts may be accomplished through the research itself, through activities that are directly related to specific research projects, or through activities that are supported by, but are complementary to, the project. The project activities may be based on previously established and/or innovative methods and approaches, but in either case must be well justified. • Meaningful assessment and evaluation of NSF funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of broader impacts and the resources provided to implement projects. If the size of the activity is limited, evaluation of that activity in isolation is not likely to be meaningful. Thus, assessing the effectiveness of these activities may best be done at a higher, more aggregated, level than the individual project.	• All awards and the full portfolio of NSF investments should reflect the highest standards of research quality across the full spectrum of Intellectual Merit and Broader Impacts. • The national importance of NSF's mission demands that successful proposals have the potential to both advance the frontiers of science and engineering and provide benefits to society. • Full, documented, and transparent consideration of both criteria is essential throughout the proposal preparation, review, and decision-making processes. • Diverse viewpoints, including institutional, geographic, demographic, disciplinary, and sector of employment, are critical to evaluating the potential impact of new or expanded knowledge and societal benefits. • Investigators, reviewers, and NSF program staff must enable enhanced accountability for project funding through meaningful assessment and evaluation of projects based on appropriate metrics established in the proposal. • Consistent and documented evaluation of portfolios at the Foundation-wide and Directorate levels, based on appropriate metrics for both criteria, is essential to communicating and assuring the value of federally funded research.

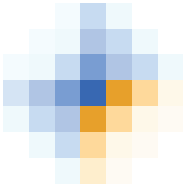
Merit Review Principles (cont.)	Merit Review Principles (cont.)
With respect to the third principle, even if assessment of Broader Impacts outcomes for particular projects is done at an aggregated level, PIs are expected to be accountable for carrying out the activities described in the funded project. Thus, individual projects should include clearly stated goals, specific descriptions of the activities that the PI intends to do, and a plan in place to document the outputs of those activities. These three merit review principles provide the basis for the merit review criteria, as well as a context within which the users of the criteria can better understand their intent.	
Merit Review Criteria	**Merit Review Criteria**
All NSF proposals are evaluated through use of two National Science Board approved merit review criteria. In some instances, however, NSF will employ additional criteria as required to highlight the specific objectives of certain programs and activities. The two merit review criteria are listed below. Both criteria are to be given full consideration during the review and decision-making processes; each criterion is necessary but neither, by itself, is sufficient. Therefore, proposers must fully address both criteria. (Chapter II.D.2.d(i) contains additional information for use by proposers in development of the Project Description section of the proposal.) Reviewers are strongly encouraged to review the criteria, including Chapter II.D.2.d(i), prior to the review of a proposal.	All proposals should be written by investigators, evaluated by reviewers, and decided on by program staff using these two criteria:

Merit Review Criteria (cont.)	Merit Review Criteria (cont.)
When evaluating NSF proposals, reviewers will be asked to consider what the proposers want to do, why they want to do it, how they plan to do it, how they will know if they succeed, and what benefits could accrue if the project is successful. These issues apply both to the technical aspects of the proposal and the way in which the project may make broader contributions. To that end, reviewers will be asked to evaluate all proposals against two criteria:	
• **Intellectual Merit:** The Intellectual Merit criterion encompasses the potential to advance knowledge; and • **Broader Impacts:** The Broader Impacts criterion encompasses the potential to benefit society and contribute to the achievement of specific, desired societal outcomes.	• **Intellectual Merit:** the potential to create new or expand existing knowledge; and • **Broader Impacts:** the potential to benefit society.
	Collectively, NSF-funded projects must contribute to a range of intellectual outcomes. Individual project outcomes should include the creation of science and engineering knowledge, ranging from fundamental to applied research. Projects can be motivated by one or a combination of curiosity, potential end use, translation to application, transformative potential, robust examination of a risky hypothesis with the prospect of high rewards, or verification of previously reported results. Other critical aspects of funded research derive from the need to increase the nation's STEM talent reservoir, for example, through increasing understanding and exploring approaches to expanding participation in STEM, including (but not limited to) institution type, geography, demographics, field of expertise, and sector of employment, as permitted by law.

Merit Review Criteria (cont.)	Merit Review Criteria (cont.)
	Simultaneously, NSF-funded projects must have the potential to benefit society – either through the near or long-term impacts of the proposed research itself, or from activities associated with the proposed research program. A wide range of societal benefits is identified in legislation including the America COMPETES Reauthorization Act of 2010, the American Innovation and Competitiveness Act of 2017, and the CHIPS and Science Act of 2022. Individual project outcomes should include, but are not limited to, increasing U.S. economic competitiveness; advancing public health and welfare; supporting the national defense; enhancing partnerships between academia and industry; developing a globally competitive American STEM workforce; improving STEM education and instruction at all levels; use of science and technology to inform public policy; improving public scientific literacy; expanding participation in STEM; and enhanced infrastructure for research and education. Congress continues to emphasize the importance of considering potential societal benefits when NSF is selecting projects to fund. These examples of intellectually and societally relevant outcomes are neither comprehensive nor prescriptive. Investigators may include other appropriate outcomes not covered by these examples. Proposers must give **full consideration** to both criteria in their proposals. Reviewers and NSF staff must do the same during the review, decision-making, and award management processes. Each criterion is necessary but neither, by itself, is sufficient. Therefore, proposers, reviewers, and NSF staff should equally and transparently address both criteria. Reviewers will evaluate each criterion individually as input to a holistic proposal assessment. These evaluations will be considered by NSF program officers as input

Merit Review Criteria (cont.)	Merit Review Criteria (cont.)
	to funding decisions and to the development of balanced award portfolios that address a range of Intellectual Merit and Broader Impacts objectives, in accordance with the Foundation's statutory mission.
The following **elements** should be considered in the review for both criteria: 1. What is the potential for the proposed activity to: a. Advance knowledge and understanding within its own field or across different fields (Intellectual Merit); and b. Benefit society or advance desired societal outcomes (Broader Impacts)? 2. To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts? 3. Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success? 4. How well qualified is the individual, team, or organization to conduct the proposed activities? 5. Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?	NSF proposals are expected to address the following: what is to be done, why it should be done, how will it be done, how will success be identified, what benefits could occur if the project is successful, or what opportunities may be lost if the project does not go forward. These considerations apply to both criteria. NSF may, when necessary, employ supplementary guidance to emphasize aspects of the criteria in certain programs and activities. However, supplementary guidance does not constitute additional criteria. With respect to both criteria, reviewers and NSF program staff should consider the following: 1. The potential for the proposed activities to: a. Create new knowledge or expand existing knowledge, including through high risk / high reward, potentially transformative, curiosity-driven, broadening participation, and reproducibility research (Intellectual Merit); and b. Benefit society in ways that may go beyond knowledge impacts, including current statutory goals (Section 102(c) of the American Innovation and Competitiveness Act of 2017) (Broader Impacts)

Merit Review Criteria (cont.)	Merit Review Criteria (cont.)
	2. The extent to which the plans for the proposed activities are well-reasoned, well-organized, cohesive, adaptable, and feasible 3. The inclusion of clear mechanisms and measures in the proposal to assess and report outcomes 4. The potential of the individual, team, or organization to conduct the proposed activities, including: a. Relevant expertise; b. Demonstrated achievement of planned outcomes from prior support, if applicable; and c. Availability of adequate resources (either at the home organization or through collaborations). For both criteria, in addition to the above, NSF program staff should consider the following (individually for each proposal and in the aggregate for award portfolios): 1. The potential for the proposed activities to contribute to a balanced award portfolio that includes a range of Intellectual Merit and Broader Impacts, in accordance with NSF's statutory mission and program objectives. 2. How the proposal may contribute to assembling a balanced award portfolio with a mix of awardees, including but not limited to consideration of career stage, institution type, and geographic location.



Appendix C: Excerpts from Relevant Legislation

1 – Section 2 of the

Government Performance and Results Act of 1993

Public Law 103-62
103d Congress

An Act

To provide for the establishment of strategic planning and performance measurement in the Federal Government, and for other purposes.

Aug. 3, 1993
[S. 20]

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

Government Performance and Results Act of 1993.
31 USC 1101 note.

SECTION 1. SHORT TITLE.

This Act may be cited as the "Government Performance and Results Act of 1993".

31 USC 1115 note.

SEC. 2. FINDINGS AND PURPOSES.

(a) FINDINGS.—The Congress finds that—

(1) waste and inefficiency in Federal programs undermine the confidence of the American people in the Government and reduces the Federal Government's ability to address adequately vital public needs;

(2) Federal managers are seriously disadvantaged in their efforts to improve program efficiency and effectiveness, because of insufficient articulation of program goals and inadequate information on program performance; and

(3) congressional policymaking, spending decisions and program oversight are seriously handicapped by insufficient attention to program performance and results.

(b) PURPOSES.—The purposes of this Act are to—

(1) improve the confidence of the American people in the capability of the Federal Government, by systematically holding Federal agencies accountable for achieving program results;

(2) initiate program performance reform with a series of pilot projects in setting program goals, measuring program performance against those goals, and reporting publicly on their progress;

(3) improve Federal program effectiveness and public accountability by promoting a new focus on results, service quality, and customer satisfaction;

(4) help Federal managers improve service delivery, by requiring that they plan for meeting program objectives and by providing them with information about program results and service quality;

(5) improve congressional decisionmaking by providing more objective information on achieving statutory objectives, and on the relative effectiveness and efficiency of Federal programs and spending; and

(6) improve the management of the Federal Government.

2 – Section 526 of the

America COMPETES Reauthorization Act of 2010

(1) collaborate with industry in the development of standards supporting trusted cloud computing infrastructures, metrics, interoperability, and assurance; and

(2) support standards development with the intent of supporting common goals.

SEC. 525. TRIBAL COLLEGES AND UNIVERSITIES PROGRAM.

42 USC 1862p–13. Grants.

(a) IN GENERAL.—The Director shall continue to support a program to award grants on a competitive, merit-reviewed basis to tribal colleges and universities (as defined in section 316 of the Higher Education Act of 1965 (20 U.S.C. 1059c), including institutions described in section 317 of such Act (20 U.S.C. 1059d), to enhance the quality of undergraduate STEM education at such institutions and to increase the retention and graduation rates of Native American students pursuing associate's or baccalaureate degrees in STEM.

(b) PROGRAM COMPONENTS.—Grants awarded under this section shall support—

(1) activities to improve courses and curriculum in STEM;

(2) faculty development;

(3) stipends for undergraduate students participating in research; and

(4) other activities consistent with subsection (a), as determined by the Director.

(c) INSTRUMENTATION.—Funding provided under this section may be used for laboratory equipment and materials.

SEC. 526. BROADER IMPACTS REVIEW CRITERION.

42 USC 1862p–14.

(a) GOALS.—The Foundation shall apply a Broader Impacts Review Criterion to achieve the following goals:

(1) Increased economic competitiveness of the United States.

(2) Development of a globally competitive STEM workforce.

(3) Increased participation of women and underrepresented minorities in STEM.

(4) Increased partnerships between academia and industry.

(5) Improved pre-K–12 STEM education and teacher development.

(6) Improved undergraduate STEM education.

(7) Increased public scientific literacy.

(8) Increased national security.

(b) POLICY.—Not later than 6 months after the date of enactment of this Act, the Director shall develop and implement a policy for the Broader Impacts Review Criterion that—

(1) provides for educating professional staff at the Foundation, merit review panels, and applicants for Foundation research grants on the policy developed under this subsection;

(2) clarifies that the activities of grant recipients undertaken to satisfy the Broader Impacts Review Criterion shall—

(A) to the extent practicable employ proven strategies and models and draw on existing programs and activities; and

(B) when novel approaches are justified, build on the most current research results;

(3) allows for some portion of funds allocated to broader impacts under a research grant to be used for assessment and evaluation of the broader impacts activity;

(4) encourages institutions of higher education and other nonprofit education or research organizations to develop and provide, either as individual institutions or in partnerships thereof, appropriate training and programs to assist Foundation-funded principal investigators at their institutions in achieving the goals of the Broader Impacts Review Criterion as described in subsection (a); and

(5) requires principal investigators applying for Foundation research grants to provide evidence of institutional support for the portion of the investigator's proposal designed to satisfy the Broader Impacts Review Criterion, including evidence of relevant training, programs, and other institutional resources available to the investigator from either their home institution or organization or another institution or organization with relevant expertise.

42 USC 1862p–15. Grants.

SEC. 527. TWENTY-FIRST CENTURY GRADUATE EDUCATION.

(a) IN GENERAL.—The Director shall award grants, on a competitive, merit-reviewed basis, to institutions of higher education to implement or expand research-based reforms in master's and doctoral level STEM education that emphasize preparation for diverse careers utilizing STEM degrees, including at diverse types of institutions of higher education, in industry, and at government agencies and research laboratories.

(b) USES OF FUNDS.—Activities supported by grants under this section may include—

(1) creation of multidisciplinary or interdisciplinary courses or programs for the purpose of improved student instruction and research in STEM;

(2) expansion of graduate STEM research opportunities to include interdisciplinary research opportunities and research opportunities in industry, at Federal laboratories, and at international research institutions or research sites;

(3) development and implementation of future faculty training programs focused on improved instruction, mentoring, assessment of student learning, and support of undergraduate STEM students;

(4) support and training for graduate students to participate in instructional activities beyond the traditional teaching assistantship, and especially as part of ongoing educational reform efforts, including at pre-K–12 schools, and primarily undergraduate institutions;

(5) creation, improvement, or expansion of innovative graduate programs such as science master's degree programs;

(6) development and implementation of seminars, workshops, and other professional development activities that increase the ability of graduate students to engage in innovation, technology transfer, and entrepreneurship;

(7) development and implementation of seminars, workshops, and other professional development activities that increase the ability of graduate students to effectively communicate their research findings to technical audiences outside of their own discipline and to nontechnical audiences;

(8) expansion of successful STEM reform efforts beyond a single academic unit to other STEM academic units within an institution or to comparable academic units at other institutions; and

3 – Sections 101 and 102 of the

American Innovation and Competitiveness Act

42 USC 1862s note.

SEC. 2. DEFINITIONS.

In this Act, unless expressly provided otherwise:

(1) APPROPRIATE COMMITTEES OF CONGRESS.—The term "appropriate committees of Congress" means the Committee on Commerce, Science, and Transportation of the Senate and the Committee on Science, Space, and Technology of the House of Representatives.

(2) FEDERAL SCIENCE AGENCY.—The term "Federal science agency" has the meaning given the term in section 103 of the America COMPETES Reauthorization Act of 2010 (42 U.S.C. 6623).

(3) FOUNDATION.—The term "Foundation" means the National Science Foundation.

(4) INSTITUTION OF HIGHER EDUCATION.—The term "institution of higher education" has the meaning given the term in section 101(a) of the Higher Education Act of 1965 (20 U.S.C. 1001(a)).

(5) NIST.—The term "NIST" means the National Institute of Standards and Technology.

(6) STEM.—The term "STEM" has the meaning given the term in section 2 of the American COMPETES Reauthorization Act of 2010 (42 U.S.C. 6621 note).

(7) STEM EDUCATION.—The term "STEM education" has the meaning given the term in section 2 of the STEM Education Act of 2015 (42 U.S.C. 6621 note).

TITLE I—MAXIMIZING BASIC RESEARCH

42 USC 1862s.

SEC. 101. REAFFIRMATION OF MERIT-BASED PEER REVIEW.

(a) SENSE OF CONGRESS.—It is the sense of Congress that—

(1) sustained, predictable Federal funding of basic research is essential to United States leadership in science and technology;

(2) the Foundation's intellectual merit and broader impacts criteria are appropriate for evaluating grant proposals, as concluded by the 2011 National Science Board Task Force on Merit Review;

(3) evaluating proposals on the basis of the Foundation's intellectual merit and broader impacts criteria should be used to assure that the Foundation's activities are in the national interest as these reviews can affirm that—

(A) the proposals funded by the Foundation are of high quality and advance scientific knowledge; and

(B) the Foundation's grants address societal needs through basic research findings or through related activities; and

(4) as evidenced by the Foundation's contributions to scientific advancement, economic growth, human health, and national security, its peer review and merit review processes have identified and funded scientifically and societally relevant basic research and should be preserved.

(b) MERIT REVIEW CRITERIA.—The Foundation shall maintain the intellectual merit and broader impacts criteria, among other specific criteria as appropriate, as the basis for evaluating grant proposals in the merit review process.

(c) UPDATES.—If after the date of enactment of this Act a change is made to the merit-review process, the Director shall submit a report to the appropriate committees of Congress not later than 30 days after the date of the change. Deadline.

SEC. 102. TRANSPARENCY AND ACCOUNTABILITY. 42 USC 1862s–1.

(a) FINDINGS.—

(1) building the understanding of and confidence in investments in basic research is essential to public support for sustained, predictable Federal funding;

(2) the Foundation has improved transparency and accountability of the outcomes made through the merit review process, but additional transparency into individual grants is valuable in communicating and assuring the public value of federally funded research; and

(3) the Foundation should commit to transparency and accountability and to clear, consistent public communication regarding the national interest for each Foundation-awarded grant and cooperative agreement.

(b) GUIDANCE.—

(1) IN GENERAL.—The Director of the Foundation shall issue and periodically update, as appropriate, policy guidance for both Foundation staff and other Foundation merit review process participants on the importance of transparency and accountability to the outcomes made through the merit review process.

(2) REQUIREMENTS.—The guidance under paragraph (1) shall require that each public notice of a Foundation-funded research project justify the expenditure of Federal funds by—

(A) describing how the project—

(i) reflects the statutory mission of the Foundation, as established in the National Science Foundation Act of 1950 (42 U.S.C. 1861 et seq.); and

(ii) addresses the Foundation's intellectual merit and broader impacts criteria; and

(B) clearly identifying the research goals of the project in a manner that can be easily understood by both technical and non-technical audiences.

(c) BROADER IMPACTS REVIEW CRITERION UPDATE.—Section 526(a) of the America COMPETES Reauthorization Act of 2010 (42 U.S.C. 1862p–14(a)) is amended to read as follows:

Applicability.

"(a) GOALS.—The Foundation shall apply a broader impacts review criterion to identify and demonstrate project support of the following goals:

"(1) Increasing the economic competitiveness of the United States.

"(2) Advancing of the health and welfare of the American public.

"(3) Supporting the national defense of the United States.

"(4) Enhancing partnerships between academia and industry in the United States.

"(5) Developing an American STEM workforce that is globally competitive through improved pre-kindergarten through grade 12 STEM education and teacher development, and improved undergraduate STEM education and instruction.

"(6) Improving public scientific literacy and engagement with science and technology in the United States.

"(7) Expanding participation of women and individuals from underrepresented groups in STEM.".

SEC. 103. EPSCOR REAFFIRMATION AND UPDATE.

(a) FINDINGS.—Section 517(a) of the America COMPETES Reauthorization Act of 2010 (42 U.S.C. 1862p–9(a)) is amended—

(1) in paragraph (1)—

(A) by striking "The National" and inserting "the National"; and

(B) by striking "education," and inserting "education";

(2) in paragraph (2), by striking "with 27 States" and all that follows through the semicolon at the end and inserting "with 28 States and jurisdictions, taken together, receiving only about 12 percent of all National Science Foundation research funding;";

(3) by striking paragraph (3) and inserting the following:

"(3) each of the States described in paragraph (2) receives only a fraction of 1 percent of the Foundation's research dollars each year;"; and

(4) by adding at the end the following:

"(4) first established at the National Science Foundation in 1979, the Experimental Program to Stimulate Competitive Research (referred to in this section as 'EPSCoR') assists States and jurisdictions historically underserved by Federal research and development funding in strengthening their research and innovation capabilities;

"(5) the EPSCoR structure requires each participating State to develop a science and technology plan suited to State and local research, education, and economic interests and objectives;

"(6) EPSCoR has been credited with advancing the research competitiveness of participating States, improving awareness of science, promoting policies that link scientific investment and economic growth, and encouraging partnerships between government, industry, and academia;

4 – Sections 10341, 10343, 10399, 10502, 10503 of the CHIPS and Science Act

Subtitle E—Fundamental Research

SEC. 10341. BROADER IMPACTS. 42 USC 19051.

(a) ASSESSMENT.—Not later than 120 days after the date of enactment of this Act, the Director shall enter into an agreement with a qualified independent organization to assess how the Broader Impacts review criterion is applied across the Foundation and make recommendations for improving the effectiveness for meeting the goals established in section 526 of the America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Reauthorization Act of 2010 (42 U.S.C. 1862p–14). Deadline. Contracts. Recommendations.

(b) ACTIVITIES.—The Director shall make awards on a competitive basis, to institutions of higher education or non-profit organizations (or consortia of such institutions or organizations) to support activities to increase the efficiency, effectiveness, and availability of resources for implementing the Broader Impacts review criterion, including—

(1) training and workshops for program officers, merit review panelists, award office administrators, faculty, and students to improve understanding of the goals and the full range of potential broader impacts available to researchers to satisfy this criterion;

(2) repositories and clearinghouses for sharing best practices and facilitating collaboration; and

(3) tools for evaluating and documenting societal impacts of research.

SEC. 10342. SENSE OF CONGRESS.

It is the sense of Congress that the Director should continue to identify opportunities to reduce the administrative burden on researchers.

SEC. 10343. RESEARCH ETHICS. 42 USC 19052.

(a) SENSE OF CONGRESS.—It is the sense of Congress that—

(1) a number of emerging areas of research have potential ethical, social, safety, and security implications that might be apparent as early as the basic research stage;

(2) the incorporation of ethical, social, safety, and security considerations into the research design and review process for Federal awards, may help mitigate potential harms before they happen;

(3) the Foundation's agreement with the National Academies to conduct a study and make recommendations with respect to governance of research in emerging technologies is a positive step toward accomplishing this goal; and

(4) the Foundation should continue to work with stakeholders to promote best practices for governance of research in emerging technologies at every stage of research.

(b) INCORPORATION OF ETHICS CONSIDERATIONS.—Drawing on stakeholder input, not later than 24 months after the date of enactment of this Act, the Director shall revise proposal instructions to require that ethical and societal considerations are to be included as part of a proposal for funding prior to making the award, where such considerations are applicable. Such considerations shall be evaluated by the Director in the review of proposals, taking into account any relevant input from the peer-reviewers for the proposal, and shall factor into award decisions, as deemed necessary by Deadline. Revised proposal. Requirement.

Evaluation. Review.

the Director. When incorporating such considerations, proposers may include, as appropriate—

(1)(A) any readily foreseeable or quantifiable risks to society, including how the research could enable products, technologies, or other outcomes that could intentionally or unintentionally cause significant societal harm; or

(B) an assertion that no readily foreseeable potential ethical, social, safety, or security implications are apparent;

(2) how technical or social solutions can mitigate such risks and, as appropriate, a plan to implement such mitigation measures; and

(3) how partnerships and collaborations in the research can help mitigate potential harm and amplify potential societal benefits.

(c) GUIDANCE.—The Director shall solicit stakeholder input to develop clear guidance on what constitutes a readily foreseeable or quantifiable risk as described in subsection (b)(1), and to the extent practicable harmonize this policy with existing ethical policies or related requirements for human subjects.

(d) RESEARCH.—The Director shall make awards, on a competitive basis, to institutions of higher education or non-profit organizations (or consortia of such institutions or organizations) to support—

Assessment.

(1) research to assess the potential ethical and societal implications of Foundation- supported research and products or technologies enabled by such research, including the benefits and risks identified pursuant to subsection (b)(1); and

Verification.

(2) the development and verification of approaches to proactively mitigate foreseeable risks to society, including the technical and social solutions identified pursuant to subsection (b)(1).

(e) ANNUAL REPORT.—The Director shall encourage recipients to update their consideration of potential risks and benefits as appropriate as part of the annual reports required by all awardees under the award terms and conditions.

42 USC 19053.

SEC. 10344. RESEARCH REPRODUCIBILITY AND REPLICABILITY.

Public information.

(a) IN GENERAL.—Consistent with existing Federal law for privacy, intellectual property, and security, the Director shall facilitate public access to research products, including data, software, and code, developed as part of Foundation-supported projects.

(b) DATA MANAGEMENT PLANS.—

Requirement.

(1) IN GENERAL.—The Director shall require that every proposal for funding for research include a machine-readable data management plan that includes a description of how the awardee will archive and preserve public access to data, software, and code developed as part of the proposed project.

(2) REQUIREMENTS.—In carrying out the requirement in paragraph (1), the Director shall—

(A) provide necessary resources, including trainings and workshops, to educate researchers and students on how to develop and review high quality data management plans;

(B) ensure program officers and merit review panels are equipped with the resources and training necessary to review the quality of data management plans; and

42 USC 19119.

SEC. 10399. REPORTS AND ROADMAPS.

(a) ANNUAL REPORT.—The Director shall provide to the relevant authorizing and appropriations committees of Congress an annual report describing projects supported by the Directorate during the previous year.

Deadline. Time period.

(b) ROADMAP.—Not later than 1 year after the date of enactment of this Act, the Director shall provide to the relevant authorizing and appropriations committees of Congress a roadmap describing the strategic vision that the Directorate will use to guide investment decisions over the following 3 years.

Time periods. Consultation.

(c) REPORTS.—Not later than 1 year after the date of enactment of this Act and every 3 years thereafter, the Director, in consultation with the heads of relevant Federal agencies, shall prepare and submit to Congress—

Strategic vision.

(1) a strategic vision for the next 5 years for the Directorate, including a description of how the Foundation will increase funding for research and education for populations underrepresented in STEM and geographic areas; and

(2) a description of the planned activities of the Directorate to secure federally funded science and technology pursuant to section 1746 of the National Defense Authorization Act for Fiscal Year 2020 (Public Law 116–92; 42 U.S.C. 6601 note) and section 223 of William M. (Mac) Thornberry National Defense Authorization Act for Fiscal Year 2021 (Public Law 116–283) and the requirements under subtitle D of this title and subtitle E of title VI .

(d) SELECTION CRITERIA REPORT.—Not later than 24 months after the establishment of the Directorate, the Director shall prepare and submit a report to Congress regarding the use of alternative methods for the selection of award recipients and the distribution of funding to recipients, as compared to the traditional peer review process.

42 USC 19120.

SEC. 10399A. EVALUATION.

Time period. Contracts.

(a) IN GENERAL.—After the Directorate has been in operation for 6 years, the Director shall enter into an agreement with the National Academies to provide an evaluation of how well the Directorate is achieving the purposes identified in section 10382.

Assessments.

(b) INCLUSIONS.—The evaluation shall include—

(1) an assessment of the impact of Directorate activities on the Foundation's primary science mission;

(2) an assessment of the Directorate's impact on the challenges and key technology focus areas under section 10387;

(3) an assessment of efforts to ensure coordination between the Directorate and other Federal agencies, and with external entities;

(4) a description of lessons learned from operation of the Directorate; and

Recommendations.

(5) recommended funding levels for the Directorate;

Public information.

(c) AVAILABILITY.—On completion of the evaluation, the evaluation shall be made available to Congress and the public.

Time period.

(2) reporting such data on an annual basis to the Director in such form as required by the Director.

42 USC 19152.

SEC. 10502. COLLECTION AND REPORTING OF DATA ON FEDERAL RESEARCH AWARDS.

(a) COLLECTION OF DATA.—

(1) IN GENERAL.—Each Federal research agency shall collect, as practicable, with respect to all applications for merit-reviewed research and development awards made by such agency, standardized record-level annual information on demographics, primary field, award type, institution type, review rating, budget request, funding outcome, and awarded budget.

Consultation.

(2) UNIFORMITY AND STANDARDIZATION.—The Director, in consultation with the heads of each Federal research agency, shall establish, and update as necessary, a policy to ensure uniformity and standardization of the data collection required under paragraph (1).

(3) RECORD-LEVEL DATA.—

Deadline.
Time period.

(A) REQUIREMENT.—Beginning not later than two years after the issuance of the policy under paragraph (2) to Federal research agencies, and on an annual basis thereafter, each Federal research agency shall submit to the National Center for Science and Engineering Statistics record-level data collected under paragraph (1) in the form required by the Director of the National Science Foundation.

(B) PREVIOUS DATA.—As part of the first submission under subparagraph (A), each Federal research agency, to the extent practicable, shall also submit comparable record-level data, if it is available to the agency, for the five years preceding the date of such submission, or an analysis for why such data cannot be provided.

Publication.

(b) REPORTING OF DATA.—The Director of the National Science Foundation shall publish statistical summary data, as practicable, collected under this section, disaggregated and cross-tabulated by race, ethnicity, sex, socioeconomic indicators, which may include employment status, occupation, educational attainment, parental education, and income, geographic location, and years since completion of doctoral degree, including in conjunction with the National Science Foundation's report required by section 37 of the Science and Engineering Equal Opportunities Act (42 U.S.C. 1885d; Public Law 96–516).

42 USC 19153.

SEC. 10503. POLICIES FOR REVIEW OF FEDERAL RESEARCH AWARDS.

(a) ASSESSMENT OF POLICIES.—Federal research agencies shall regularly assess, and update as necessary, policies, and practices to remove or reduce cultural and institutional barriers limiting the recruitment, retention, and success of groups historically underrepresented in STEM research careers, including policies and practices relevant to the unbiased review of Federal research applications.

(b) CONSIDERATIONS AND ACTIVITIES.—In carrying out the requirements under subsection (a), Federal research agencies shall—

(1) review current levels of participation of groups historically underrepresented in STEM in peer-review panels and consider approaches for expanding their participation;

(2) analyze the data collected under section 10502, including funding rates of proposals from all groups, including those historically underrepresented in STEM; Analysis.

(3) collect and disseminate best practices to remove or reduce cultural and institutional barriers limiting the recruitment, retention, and success of groups historically underrepresented in STEM research careers; and

(4) implement evidence-based policies and practices to achieve the goals of this section.

SEC. 10504. COLLECTION OF DATA ON DEMOGRAPHICS OF FACULTY. 42 USC 19154.

(a) COLLECTION OF DATA.—

(1) IN GENERAL.—Not later than 5 years after the date of the enactment of this Act and at least every five years thereafter, the Director of the National Science Foundation shall carry out a survey to collect data from award recipients on the demographics of STEM faculty, by broad fields of STEM, at different types of institutions of higher education that receive Federal research funding. Deadline. Time period. Survey.

(2) SURVEY CONSIDERATIONS.—To the extent practicable, the Director of the National Science Foundation shall survey, by sex, race, socioeconomic indicators, which may include employment status, occupation, educational attainment, parental education, and income, geographic location, ethnicity, citizenship status, and years since completion of doctoral degree—

(A) the number and percentage of faculty;

(B) the number and percentage of faculty at each rank;

(C) the number and percentage of faculty who are in nontenure-track positions, including teaching and research;

(D) the number and percentage of faculty who are reviewed for promotion, including tenure, and the percentage of that number who are promoted, including being awarded tenure;

(E) faculty years in rank;

(F) the number and percentage of faculty to leave tenure-track positions;

(G) the number and percentage of faculty hired, by rank; and

(H) the number and percentage of faculty in leadership positions.

(b) EXISTING SURVEYS.—The Director of the National Science Foundation, may, in modifying or expanding existing Federal surveys of higher education (as necessary)—

(1) take into account the considerations under subsection (a)(2) by collaborating with statistical centers at other Federal agencies; or

(2) make an award to an institution of higher education or nonprofit organization (or consortia thereof) to take such considerations into account.

(c) REPORTING DATA.—The Director of the National Science Foundation shall publish statistical summary data collected under this section, including as part of the National Science Foundation's report required by section 37 of the Science and Engineering Equal Opportunities Act (42 U.S.C. 1885d; Public Law 96–516). Publication.

Appendix D: Charge to the NSB-NSF Commission on Merit Review

February 16, 2023

NSB hereby establishes a Commission to re-examine the current Merit Review policy, associated criteria, and process for reasons including:

- The current criteria have been in effect since 1997, with the last NSB review occurring in 2011. In the past decade, three laws have expressed relevant Congressional intent and interest.
- Committee of Visitors reports revealed discrepancies in balancing Broader Impacts (BI) with Intellectual Merit, and questions persist about the role of BI.
- While the 2011 NSB policy stressed that "assessment and evaluation of NSF funded projects should be based on appropriate metrics," the NSB Committee on Oversight has identified challenges in measurement of outcomes.
- In light of enduring funding disparities, it is critical that NSF's merit review processes be fair and unbiased, particularly with respect to researchers of color and other underrepresented researchers.
- The development of communities of expertise in critical areas such as BI and data analysis brings an opportunity to improve process and policy.

Finally, the newly passed CHIPS + Science Act, which reinforces NSF's strategic focus on both delivering benefits from research and developing diverse domestic STEM talent, increases the importance of the broader impact review criterion. Ensuring that the Merit Review criteria, process, and reporting are delivering both new knowledge and societal benefits, is essential to our nation's competitiveness and security.

A group, now named the **NSB-NSF Commission on Merit Review**, was established at the **December 1-2, 2022** National Science Board meeting, *to assess the efficacy of the current Merit Review policy and associated criteria and processes at supporting NSF's mission to create new knowledge, fully empower diverse talent to participate in STEM, and benefit society by translating knowledge into solutions.*

The Commission is soliciting input widely from the research and stakeholder communities and may solicit special studies as appropriate. This reexamination may lead to recommendations regarding the current policy, process, and reporting mechanism. The Commission expects to deliver a report with recommendations by May 2025.

Appendix E: Glossary

Accountability – The responsibility to evaluate the use of federal funds, assure the effectiveness of programs, and demonstrate to the public that awards made by NSF contribute to the advancement of knowledge and benefit society.

Award – is used in this report as a generic term encompassing all funding instruments issued by NSF following merit review, including but not limited to grants, fellowships, and cooperative agreements.

Broader Impacts (BI) - The Broader Impacts criterion encompasses the potential to benefit society – either inherent in the near- or long-term impacts of the proposed research, or from activities associated with the proposed research program.

Committee of Visitors (COV) - A committee of visitors (COV) is a committee of external experts convened by NSF to conduct a review of the quality and integrity of the Merit Review process, operations, and technical and managerial matters pertaining to proposal determinations made within one or more programs under review.

Demographics - characteristics of different groups and subgroups within a population.

Expanding Participation - is the comprehensive phrase describing efforts to increase the participation of individuals, institutions, and communities across the nation – ensuring broad access to resources and opportunities for discovery and innovation.

Intellectual Merit (IM) - The Intellectual Merit criterion encompasses the potential to advance knowledge.

Merit Review (MR) – The process which NSF relies on to ensure the proposals it receives are reviewed in a fair, competitive, transparent and in-depth manner.

Portfolio - A portfolio is a collection of components that is collated based on one or more shared characteristics, including thematic area or administrative level. Portfolio is used in several contexts related to NSF Merit Review:

> ***Award Portfolio*** refers to a collection of awards resulting from the Merit Review process and categorized by a shared characteristic such as funding opportunity, organizational unit or level, topical focus, or approach to specific research and education objectives.
>
> ***Program Portfolio*** refers to a collection of programs or funding opportunities categorized by a shared characteristic such as the managing organizational unit, topical focus, or proposer eligibility.
>
> ***Portfolio Balance*** is a mix of activities across both Merit Review criteria within a portfolio that has the potential to produce intellectual and societal benefits in alignment with disciplinary and/or program objectives, NSF strategic goals, and national priorities.

Principal Investigator (PI) - A category of senior/key personnel, designated by the proposer/recipient organization and approved by NSF, who contributes in a substantive, meaningful way to the scientific development or execution of a research and development project.

Transparency – The clear, consistent, and accessible communication (within NSF and to the public, to the extent appropriate) of information regarding the evaluation of proposals and outcomes of awards, and of the assessments and outcomes of award portfolios at the program, directorate, and agency levels. This includes making information more available, searchable, discoverable, usable, and re-usable for internal and external stakeholders, to the extent consistent with legal requirements and appropriate.

Acronyms

AICA – American Innovation and Competitiveness Act of 2017

BI – Broader Impacts

COV – Committee of Visitors

IM – Intellectual Merit

MR – Merit Review

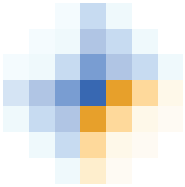
NSB – National Science Board

NSF – U.S. National Science Foundation

RFI – Request for Information

S&E – Science and Engineering

STEM – Science, Technology, Engineering, and Mathematics, including Computer Science

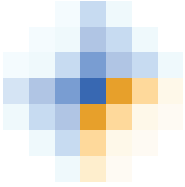


Appendix F: Commission Timeline

Date	Task
November 28, 2022	National Science Board (NSB) Committee on Oversight retreat to reach consensus on recommendation to establish Commission and discuss draft Commission Charge.
December 1, 2022	NSB establishes NSB-NSF Commission on Merit Review.
Jan-Mar 2023	NSB Chair selects Commission members
February 15, 2023	NSB approves Commission Charge (Board meeting; Appendix D)
March 23, 2023	Listening session with Broader Impacts experts at ARIS Summit
April 19, 2023	Discussion of potential topical areas of inquiry and Commission workplan (Commission meeting)
May 8, 2023	Discussion of current Merit Review Policy (Commission meeting)
May 10, 2023	Quarterly Commission update to NSB (Board meeting)
May 24, 2023	Discussion of Broader Impacts criterion and ARIS listening session findings (Commission meeting)
June 15, 2023	Discussion with expanding participation experts at CEOSE Advisory Committee meeting
June 28, 2023	Discussion of Intellectual Merit criterion (Commission meeting)
August 14, 2023	Panel discussion on STEM federal agency grant review policies and practices (Commission meeting)
August 16, 2023	Quarterly Commission update to NSB (Board meeting)
September 27, 2023	Discussion of the potential for additional Merit Review Criteria (Commission meeting)
October 25, 2023	Discussion of the evaluation of Merit Review criteria (Commission meeting)
November 28, 2023	Discussion of high-risk/high-reward research and portfolio management (Commission meeting)
November 29-30, 2023	Quarterly Commission update to NSB (Board meeting)
December 13, 2023	Discussion of data collection, workplan, and instruments, and increased participation in Merit Review (Commission meeting)
January 18, 2024	Discussion prioritizing topics for potential preliminary policy recommendations (Commission meeting)

January 24, 2024	Listening Session with University Administrators at Big Ten Summit
January 29, 2024	Discussion prioritizing topics for potential implementation and accountability guidance (Commission meeting)
Jan-Feb 2024	Contractor data collection from NSF staff and leadership via survey and interviews
February 22, 2024	Quarterly Commission update to NSB (Board meeting)
April 3, 2024	Discussion with Broader Impacts experts at ARIS Summit
April 30, 2024	Commission meeting to finalize and approve preliminary policy recommendations
May 2, 2024	Quarterly Commission update to NSB (Board meeting)
June 13, 2024	Discussion with expanding participation experts at CEOSE Advisory Committee meeting
July 23, 2024	Commission meeting to finalize and approve preliminary guidance
July 25, 2024	Quarterly Commission update to NSB (Commission meeting)
August 12, 2024	Commission meeting to discuss preliminary accountability guidance
August 21, 2024	Commission meeting to finalize and approve preliminary guidance
Aug 26-Sept 20, 2024	Request for Information/Dear Colleague Letter response window
Aug 26-Sept 27, 2024	University Vice Presidents of Research and NSF Advisory Committee focus group period
October 3, 2024	Commission meeting to finalize and approve preliminary guidance
October 21, 2024	Quarterly Commission update to NSB (Commission meeting)
November 4, 2024	Commission meeting to discuss recommendations and guidance
November 18, 2024	Commission meeting to discuss draft report
November 25, 2024	Delivery of Mathematica report (Appendix J) to Commission
December 3, 2024	Commission meeting to approve final recommendations and guidance
December 5, 2024	Quarterly Commission update to NSB (Board meeting)

December 11, 2024	Commission meeting to approve final guidance
December 20, 2024	Delivery of draft report to Commission and external reviewers
January 13, 2025	Commission meeting to discuss revised Commission report
February 11, 2025	Quarterly Commission update to NSB
February 28, 202	Delivery of revised draft report to Commission
March 10, 2025	Commission meeting to approve sending report to NSB leadership team; Commission formally discharged
May 7, 2025	Quarterly Commission update to NSB (Board meeting)
July 23, 2025	Quarterly Commission update to NSB (Board meeting)
August 18, 2025	NSB approval of final report



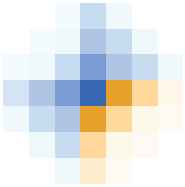
Appendix G: Bibliography

1. America COMPETES Reauthorization Act of 2010, Pub. L. 111-358 (2011)
2. American Innovation and Competitiveness Act, 2017, Pub. L. 114-329 (2017)
3. CHIPS and Science Act, Pub. L. 117-167 (2022)
4. Cheruvelil, Kendra S., et al. "Creating and maintaining high-performing collaborative research teams: the importance of diversity and interpersonal skills." Frontiers in Ecology and the Environment 12.1 (2014): 31-38.
5. Committee on the Future of NSF EPSCoR, "Envisioning the Future of NSF EPSCoR," (Alexandria, VA; National Science Foundation, 2022).
6. CEOSE, "Making visible the invisible: STEM Talent of Rural America," (Alexandria, VA: Committee on Equal Opportunities in Science and Engineering, 2020)
7. Committee on STEM Education of the National Science & Technology Council, "Charting a Course for Success: America's Strategy for STEM Education" (December 2018).
8. Council of Canadian Academies, The Expert Panel on International Practices for Funding Natural Sciences and Engineering Research, "Powering Discovery," (Ottawa, ON: Council of Canadian Academies, 2021)
9. Dixon-Fyle, S., Huber, C., Del Mar Martínez Márquez, M., Prince, S., & Thomas, A. "Diversity matters even more: The case for holistic impact," McKinsey & Company (2023).
10. Government Performance and Results Act, Pub. L. 103-62 (1993)
11. 11) H. Rept. 108-674 – Departments of Veterans Affairs and Housing and Urban Development, and Independent Agencies Appropriations Bill, 2005.
12. National Academy of Public Administration (NAPA), "A Study of the National Science Foundation's Criteria for Project Selection," (Washington D.C.: National Academies of Public Administration, 2001)
13. National Academies of Sciences, Engineering, and Medicine (NASEM), "Advancing Diversity, Equity, Inclusion, and Accessibility in the Leadership of Competed Space Missions," (Washington, DC: The National Academies Press, 2022)
14. National Center for Science and Engineering Statistics (NCSES), "Analysis of Federal Funding for Research and Development in 2022: Basic Research," (Alexandria, VA: National Science Foundation, 2024)
15. NCSES, "Women, Minorities, and Persons with Disabilities in Science and Engineering," (Alexandria, VA: National Science Foundation, 2021)
16. National Science Board (NSB), "A Changed Science and Engineering Landscape," (Alexandria, VA: National Science Board, 2024)
17. NSB, "Charge to the NSB-NSF Commission on Merit Review", (Alexandria, VA: National Science Board, 2022)
18. NSB, "Connected Horizons", (Alexandria, VA: National Science Board, 2024)
19. NSB, Major Actions and Approvals at the December 1-2, 2022 Meeting, NSB-2022-45, (Alexandria, VA: National Science Board, 2022)

20. NSB, "Merit Review Process: Fiscal Year 2021 Digest," (Alexandria, VA: National Science Board, 2023)
21. NSB, "NSF's Merit Review Criteria: Review and Revisions," (Arlington, VA: National Science Board, 2011
22. NSB, National Science Board and National Science Foundation Staff Task Force on Merit Review, Discussion report (NSB/MR-96-15), (Arlington, VA: National Science Board, 1996)
23. NSB, "National Science Board and National Science Foundation Staff Task Force on Merit Review: Final Recommendations," (Arlington, VA: National Science Board, 1997)
24. NSB, "Report of the NSB Committee on Excellence in Science and Engineering," (NSB-85-50) (Washington D.C.: National Science Board, 1985)
25. NSB, "Report to the National Science Board on the National Science Foundation's Merit Review Process Fiscal Year 2005," (Arlington, VA: National Science Board, 2006)
26. NSB, "Report of the National Science Board on the National Science Foundation's Merit Review System (NSB-05-119)," (Arlington, VA: National Science Board, 2005)
27. NSB, Resolution, "Training to Improve Peer Reviewing in Merit Review Process, (NSB-2021-10)," (Alexandria, VA: National Science Board, 2021)
28. NSB, Resolution, "Broader Impacts Experts to Serve on Committees of Visitors, (NSB-2021-11)," (Alexandria, VA: National Science Board, 2021)
29. NSB, "Talent is the Treasure," (Alexandria, VA: National Science Board, 2024)
30. NSB, "Vision 2030," (Alexandria, VA: National Science Board, 2020)
31. National Science Foundation Act of 1950, Pub. L. No. 81-507 (1950)
32. National Science Foundation (NSF), Advisory Committee on Merit Review (NSF 86-93) (Washing-ton, DC: National Science Foundation, 1986).
33. NSF, "Annual report, 1952," (Washington, DC: Government Printing Office 1952)
34. NSF, "Annual report, 1967," (Washington, DC: Government Printing Office 1967).
35. NSF, "Annual report, 1974," (Washington, DC: Government Printing Office 1974).
36. NSF, "Grant proposal guide (NSF-95-27)," (Arlington, VA: National Science Foundation, 1995).
37. NSF, "NSF by the Numbers," (Alexandria, VA: National Science Foundation, accessed 11/2024)
38. NSF, "NSF Proposal & Award Policies & Procedures Guide (PAPPG)(NSF-24-1)," (Alexandria, VA: National Science Foundation, 2024, accessed 11/2024)
39. NSF, "On the Origins of Google," (Alexandria, VA: National Science Foundation, 2024, accessed 12/2024)
40. Page, S. E. "Diversity Powers Innovation," Center for American Progress. (January 26, 2007).
41. Report of the Advisory Committee for GPRA Performance Assessment FY 2009 (NSF 09-068).
42. Srivastava, A., and S.B. Thomson. 2009. "Framework Analysis: A Qualitative Methodology for Applied Research Note Policy Research." Journal of Administration and Governance, 4(2): 72–79.

43. U.S. Congress, House, Subcommittee on Science, Research, and Technology report, Section E (1976), located in NSB report to House Subcommittee on Science, Research, and Technology Regarding Peer Review Procedures at NSF, Appendix A (Alexandria, VA: National Science Board 1977).

44. Written Testimony of Daniel Reed, NSB Chair, "Oversight and Examination of the National Science Foundation's Priorities for 2025 and Beyond," (Research and Technology Subcommittee, Committee on Science, Space, & Technology, U.S. House of Representatives, 118th Cong., 2024



Appendix H: Listening Sessions and Focus Groups

Advancing Research Impact in Society Summit Listening Session
March 23, 2023

The NSB Commission on Merit Review conducted a listening session at the 2023 Advancing Research Impacts in Society (ARIS) Summit in Baltimore, MD. Over 120 Broader Impacts (BI) experts were present during the 90-minute listening session. The session began with an introduction to the Commission and its charge.

Attendees offered their views in response to questions posed by Commission leadership regarding NSF's Merit Review process and review criteria. Attendees also provided feedback on the broader impacts review criterion, perspectives on the Merit Review process and criteria, and offered comments on institutional support and tracking of Broader Impacts related award outcomes. These conversations helped Commission members understand the perspective of Broader Impacts professionals and practitioners when deliberating throughout the Commission's work.

Committee on Equal Opportunities in Science and Engineering Listening Session
June 15, 2023

The Committee on Equal Opportunities in Science and Engineering (CEOSE) convened a public discussion with the NSB Commission on Merit Review leadership in June of 2023. Conversations centered around sharing the work of the Commission to date, as well as sharing with CEOSE the overarching questions guiding the Commission's work and receiving feedback on NSF's Merit Review criteria and review process.

CEOSE members and Commission leadership discussed their perspectives on the existing review criteria, Expanding Participation, and barriers to participation in NSF's Merit Review policy and process, and members provided feedback and recommendations to the Commission. CEOSE members also discussed the role of reviewer expertise in both IM and BI and the importance of reviewer selection in the Merit Review process. Other topics of discussion included the importance of institutional support of Merit Review, and the relative weighting of IM and BI in the review process. This discussion helped Commission members understand the Advisory Committee members' perspective on Merit Review policy and process.

STEM Federal Funding Agency Panel
August 14, 2023

The NSB Commission on Merit Review convened a public panel discussion with representatives from the National Institutes of Health (NIH), the Department of Energy, the National Aeronautics and Space Administration (NASA), the Advanced Research Projects Agency for Health (ARPA-H), and the Defense Advanced Research Projects Agency (DARPA). The discussion focused on approaches to research grant review within the federal government.

Big Ten Listening Session
January 24, 2024

The Big Ten Listening session on Merit Review was designed to hear directly from members of the Big Ten Academic Alliance both on the Merit Review process and on institutional commitment to Merit Review criteria. To capture that perspective, the Big Ten listening session audience consisted of provosts, vice presidents of research, graduate school deans, sponsored research officers, and public engagement officials from institutions within the Big Ten Academic Alliance (BTAA). There was one panel conversation and two focus group sessions, the first on Merit Review, and the second included a panel discussion on institutional commitment to Merit Review.

This listening session focused on how institutions support principal investigators in their efforts related to Broader Impacts and Intellectual Merit, and on feedback about how those criteria are used in evaluating faculty for tenure and promotion. These conversations helped Commission members understand the perspective of institutions in terms of the perception of Merit Review and its implementation at the institutional level.

Committee on Equal Opportunities in Science and Engineering Listening Session
June 13, 2024

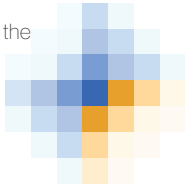
The Committee on Equal Opportunities in Science and Engineering (CEOSE) convened a public discussion with the NSB Commission on Merit Review leadership in June of 2024. The conversation focused on recent CEOSE reports and the status of recommendations made to NSF regarding specific areas of interest.

The first portion of the listening session concerned recent CEOSE reports and recommendations, including those made to increase attention to community driven projects and improving representation on panels and in ad-hoc reviewer pools. There was also a brief discussion on pursuing high-risk and high-reward projects. CEOSE members were asked to describe what success would look like in each area and offer specific implementation suggestions to achieve these goals. This discussion helped Commission members understand the Advisory Committee members' perspective on Merit Review policy and process, including expanded participation and high-risk, high-reward projects.

Vice Presidents for Research and Advisory Committee member Focus Groups
August/September 2024

The Commission contracted Mathematica to organize two focus groups. The first was comprised of individuals currently serving as Vice Presidents of Research or a similar role at an academic institution. The second group was comprised of individuals who served on an NSF Advisory Committee and Committee of Visitors panel. The study team used a semi-structured interview protocol, developed with input from Commission members and staff, to gather information related to Merit Review policies and processes from each population.[60] These focus groups helped Commission members understand the perspective of these cohorts on Merit Review policy and process.

60 This information collection was conducted with approval from the Office of Management and Budget under NSF's Generic Clearance for the Collection of Qualitative Feedback on Agency Service Delivery (approval no. 3145-0215).



Appendix I: NSF Committee of Visitors Analysis

Longitudinal Review of National Science Foundation Committee of Visitor Reports, from 2012-2023

June 2024

National Science Board (NSB)
National Science Foundation (NSF)

Introduction

To help inform the Commission's efforts to improve the National Science Foundation's (NSF) Merit Review (MR) process, National Science Board Office (NSB) staff surveyed comments from Committee of Visitor (COV) reports produced between 2012-2023.

For context, the purpose and scope of COVs are identified in the most recent COV Policy document:

> *NSF relies on the judgment of external experts to maintain high standards of program management, to provide advice for continuous improvement of NSF performance, and to ensure openness to the research and education community served by the Foundation. A COV is a committee of external experts, appointed as Special Government Employees, convened by NSF to conduct a review of the quality and integrity of the merit review process, operations, and technical and managerial matters pertaining to proposal determinations made within one or more programs under review. COVs shall not be used for outcome assessment and evaluation of the outcomes or long-term impacts of program investments.* (pg2)

In short, NSF relies on external experts to review the quality and integrity of the MR process, operations, and technical and managerial matters pertaining to proposal decisions.

The purpose of this study was to consolidate and evaluate COV comments regarding the MR process and criteria. The primary analysis focused on trends and patterns regarding the aggregated perception of the strengths and weaknesses in the MR criteria, including Intellectual Merit (IM) and Broader Impacts (BI), and MR process over time and across Directorates. The secondary analysis focused on general trends in COV comments and recommendations since the last MR examination, in 2011.

Methods

Twenty-eight COV Reports produced between 2012-2023, representing 25% of all COV Reports since the previous Merit Review Taskforce, were reviewed. One to two Divisions from each NSF Directorate, which held at least three COVs between 2012-2023, were selected at random for review. Since most Divisions hold COV panels every 3-4 years, reviewing the 3 most recent COV reports provided a glimpse at comments spanning the past decade, allowing for longitudinal analysis. Because of these parameters, our sample did not include COV reports from 2017 or 2021.

Originally, COV Reports from two divisions per Directorate were planned for review. However, time constraints forced the evaluation of only one Divisions worth of COV Reports per Directorate halfway through data collection. As a result, two Divisions within each of the Directorates for Engineering (ENG), Geosciences (GEO), and Math and Physical Sciences (MPS) were evaluated while only one Division was evaluated from the Directorate for Biological Sciences (BIO), Computer and Information Science and Engineering (CISE), STEM Education

(EDU), Social, Behavioral and Economic Science (SBE), and the Director's Office of Integrative Activities (OIA). Similarly, NSF *Responses to COV Reports* were also originally planned to be reviewed but were similarly omitted due to time constraints.

Reports were surveyed for comments pertaining to known Commission interests, emphasizing discussions on the MR criteria and process. The surveying process involved searching for 39 common MR-related terms and phrases, outlined in Appendix 1. These terms will be discussed in greater detail later in this report but were essentially binned following coding rubrics or lemmatized, which is a linguistic method of organizing variations on the same root word. For example, COV comments employing the phrases "Broaden Participation" and "Broadening Participation" were grouped together in the lemmatization process. Comments were considered for analysis based on their phrasing, argument, or message.

Comments were then categorized as "Positive" if commending the Directorate or Division, "Critical" if offering a critique, "Neither" if neutral, or "Recommendation" if offering specific advice or avenues to resolved discussed challenges. Each comment was defined by capturing a specific COV idea or perspective. COVs would commonly generate large paragraphs discussing specific MR, IM, or BI-related topics, and generally followed a "positive" beginning, "critical" midsection, and ended with a section dedicated to "recommendations." These larger paragraphs were saved entirely for reference, but broken into their respective "positive," "critical," and "recommendation" sections.

Positive, critical, and recommendation comments were analyzed for their average distribution through time and across Directorates. This allowed for the generation of an *Apparent Satisfaction Level* to help communicate aggregate COV perceptions of the MR criteria and process between Directorates, shown in *Table 1*. Similarly, the bins of common MR terms were evaluated for their average use across positive, critical, and recommendation comments, and average combined use within comments.

Results

A total of 543 comments were collected from COV reports between 2012-2023. From that, 150 (27.6%) were categorized as "positive", 203 (37.4%) were categorized as "critical", and 172 (31.7%) were categorized as "recommendations," shown in *Figure 1*. Aggregated for the entire previous decade, critical comments comprised just under a third of all MR-related comments, while positive comments comprised slightly more than a quarter. Again, our sample did not include COVs from 2017 or 2021.

Patterns Through Time

Using 2017 as a natural cut point to split the data roughly in half allowed for the analysis of changes in comments. Comparing comments from 2012-2016 (see *Figure 2*) to comments from 2018-2020 & 2022-2023 (see *Figure 3*), the proportion of positive comments decreased by 11 percentage points, and critical comments and recommendations increased by 6 and 5 percentage points, respectively. However, looking at the actual numbers, you can see that the number of positive comments remained fairly consistent, while the number of critical comments and recommendations increased. While reasons for these changes were beyond the scope of this study, one may speculate that it could be due to changes in NSF prompts for COV discussions. However, this large

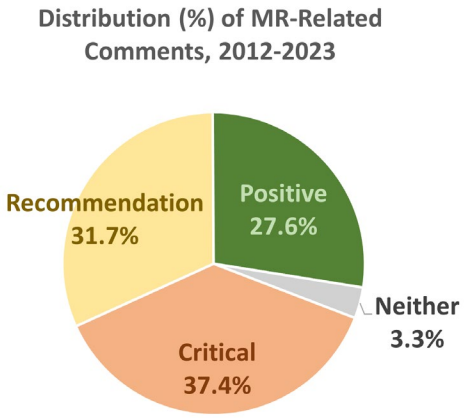


Figure 1. Distribution of positive comments, critical comments, and recommendations captured in COV Reports between 2012-2023.

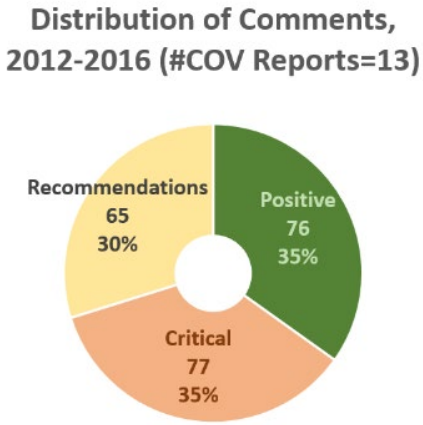


Figure 2. Distribution of positive comments, critical comments, and recommendations from COV reports between 2012-2016.

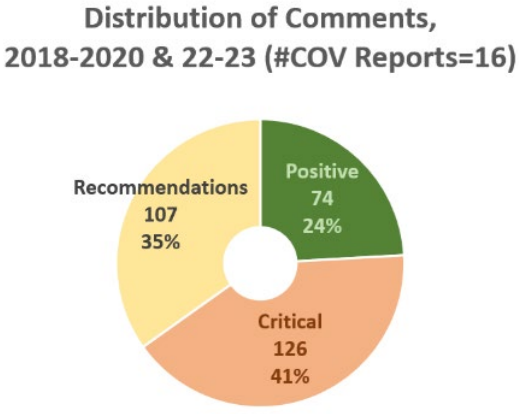


Figure 3. Distribution of positive comments, critical comments, and recommendations from COV reports between 2018-2020, 2022-2023.

shift in COV feedback is worth considering for future study.

Figure 4 displays the annual distribution over the past decade of the surveyed MR-related bins of terms, averaged by COV report, employed per positive, critical, and recommendation comments. Essentially, this chart identifies the average number of MR-related

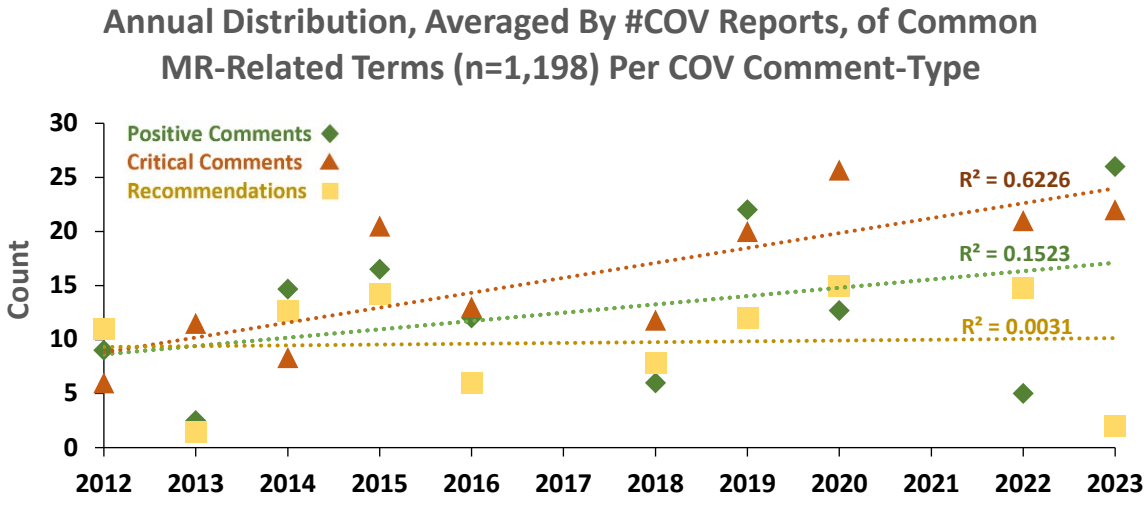


Figure 4. Annual distribution of MR-related terms across positive, critical, and recommendation comments, averaged by number of COV reports. This figure reduces the influence of varying numbers of COV reports per year. The vertical axis represents the number of terms being employed per comment, averaged annually. MR-related terms were, in aggregate, being increasingly employed in critical comments, outperforming those in positive comments, through the past decade.

terms used per comment. The distribution of MR-related terms in recommendations, although variable, remained consistent on average through time. The average distribution of MR-related terms in positive comments, while also variable, experienced a slight increase through the past decade. The average distribution of MR-related terms in critical comments appears to have substantially increased over time, particularly after 2017. These data reveal MR-related terms were employed by COVs at an increasing rate in critical over positive or recommendation comments.

Of additional note, most of the comments included multiple MR-related terms. The inclusion of multiple terms per comment impacted averages (i.e. average number of terms vs. entire comments) and the reader should take care to consider these differences. Additionally, the number of analyzed COV reports ranged between one to five COV reports per year, except for 2017 and 2021 during which no surveyed COV reports occurred.

Distribution of Positive, Critical, and Recommendation Comments Across Directorates

Figures 5, 6, and *7* display heatmaps showing the distribution of positive, critical, and recommendation comments, respectively, across the survey Directorates through the past decade. The average for all directorates is shown at the top of each figure. As shown in *Figure*

5, the average distribution of positive comments remained generally consistent across three Directorates (i.e. EDU, GEO's Atmospheric and Geospace Sciences (AGS) division, and MPS). By comparison, three others saw a decrease in positive comments (i.e. BIO, ENG's Emerging Frontiers and Multidisciplinary Activities (EFMA) division, and OIA).

Some Divisions and Directorates produced COV reports but did not produce any positive comments employing the surveyed terms, as indicated with zeros ("0"). Similarly, double astricts (i.e. "**") are employed to indicate a COV report was not accessible and therefore not analyzed for this report.

Positive COV Comments (n=150) Per Directorate Through Time												
Directorate	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**
Average:	4.3	1.5	5.7	8.8	4.0		3.2	16.0	4.3		2.0	8.0
BIO			7				3				3	
CISE	**						2				1	
EDU	3			1					1			
ENG: EFMA			5				0				2	
ENG: IIP		0			0				0			
GEO: AGS			5				7		3			
GEO: OCE			**				4				2	
MPS: DMS		3			8				8			
MPS: PHY				13				16				8
OIA	9			6					3			
SBE	1			15					11			

** issues accessing the COV Report

Figure 5. Distribution of positive comments in COV Reports, including average. Data was not collected from the GEO:OCE 2014 or CISE 2011 COV Report due to technical issues. COV reports published in 2017 and 2021 were not included in this review.

As shown in *Figure 6*, the average distribution of critical comments appears to generally increase over the past decade. Five Directorates (BIO, ENG's EFMA division, MPS's Mathematical Sciences (DMS) division, OIA, and SBE) saw increases while only one (MPS's

Critical COV Comments (n=203) Per Directorate Through Time												
Directorate	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**
Average:	3.0	6.5	2.7	9.8	4.0		4.4	7.0	9.5		8.3	7.0
BIO			1				3				12	
CISE	**						4				6	
EDU	2			2					2			
ENG: EFMA			4				7				12	
ENG: IIP		2			2				0			
GEO: AGS			3				5		2			
GEO: OCE			**				3				3	
MPS: DMS		11			6				26			
MPS: PHY				18				7				7
OIA	1			2					5			
SBE	6			17					22			

** issues accessing the COV Report

Figure 6. Distribution of positive comments in COV Reports, including average. Data was not collected from the GEO:OCE 2014 or CISE 2011 COV Report due to technical issues. COV reports published in 2017 and 2021 were not included in this review.

Physics, or PHY) saw a decrease. Three Directorates (EDU, ENG's Engineering Education and Careers (EEC) division, and GEO) saw no notable change in their rate of critical comments.

Figure 7 reveals the average distribution of recommendations remained generally consistent across the past decade. Three directorates (BIO, MPS's DMS division, and SBE) saw an increase in recommendations, two (MPS's PHY division and OIA) saw a decrease, and three (EDU, ENG, and MPS's PHY division) remained consistent.

COV Recommendations (n=172) Per Directorate Through Time												
Directorate	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**
Average:	4.0	0.5	6.0	7.0	2.5		4.0	4.0	7.3		9.3	2.0
BIO			3				3				17	
CISE	**						1				1	
EDU	4			5					5			
ENG: EFMA			13				4				18	
ENG: IIP		1			0				3			
GEO: AGS			2				12		1			
GEO: OCE			**				0				1	
MPS: DMS		0			5				20			
MPS: PHY				3				4				2
OIA	4			2					2			
SBE	4			18					13			

*** issues accessing the COV Report*

Figure 7. Distribution of positive comments in COV Reports, including average. Data was not collected from the GEO:OCE 2014 or CISE 2011 COV Report due to technical issues. COV reports published in 2017 and 2021 were not included in this review.

The distributions discussed above, of positive, critical, and recommendation comments produced by COVs, exacerbate the distribution of MR-related terms used within these comments, as discussed in *Figure 4*. The average use of positive, critical, and recommendation comments employing MR-related terms were used to generate an *Apparent Satisfaction Levels* metric to rank COV perceptions across Directorates, shown in *Table 1*. Directorates were categorized in *Greatest Satisfaction* if their COVs used MR-related terms at above average rates in positive comments and below average rates in critical or recommendation comments. COVs from the GEO Directorate and the Physics Division within MPS employed MR-related terms at above-average rates in positive comments, thus expressing greater satisfaction compared to the other Directorates. The inverse of this method led to categorization as *Least Satisfied* and

Table 1. Conceptualization of satisfaction-levels based on above-average use of MR-terms in positive, critical, and recommendation comments across NSF Directorates.

Communicated Satisfaction Levels	COVs within…
Greatest Satisfaction	MPS:PHY and GEO
Average Satisfaction	OIA, EDU, and BIO
Least Satisfaction	ENG, MPS:DMS, SBE, and CISE
Apparent *Increasing* Proportion of **Critical Comments**	BIO, ENG:EFMA, MPS:DMS, OIA, and SBE

employing these terms near-average led to categorization in *Average Satisfaction*. For example, CISE, ENG, SBE, and the DMS Division in MPS employed MR-related terms at above average rates in critical and recommendation comments, thus expressing less satisfaction compared to the other Directorates. The most pronounced Directorates and Divisions are listed first (i.e. on the left of each list) and decrease when moving left-to-right, down the list.

Also of note, BIO, OIA, SBE, the EFMA Division in ENG, and the DMS Division in MPS appear to have an increased proportional use in critical comments post-2017 compared to earlier in the decade. However, this trend may result from changes in COV prompts.

Bins of Lemmatized Terms and Term-Coding Rubrics

Originally, 39 common Merit Review-related terms were binned into either lemmatized groups, to consolidate various forms of root words, or within coding rubrics (see Appendices 1 and 2 for tables organizing the terms within each bin). The original bins include *Broadening Participation, Broader Impacts, Clarity, Intellectual Merit, Merit Review,* and *Weighting*. During Analysis, an additional 24 terms related to the *Response Length of reviewers in proposals* were binned into a coding rubric to evaluate COV comments discussing the length and substance of review feedback within reviews. The distribution of these terms in positive, critical, and recommendation comments are shown in *Figure 8.* As a reminder, *Figure 8* captures the

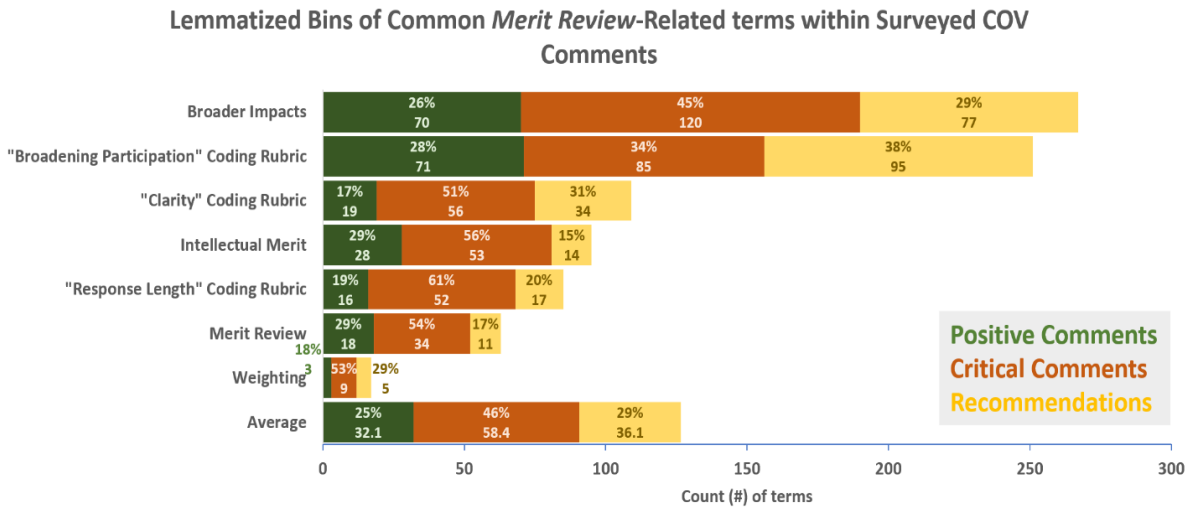


Figure 8. Bins of Lemmatized terms or Coding Rubrics for terms commonly employed within COV comments of NSF's Merit Review process and criteria. Terms within positive comments are colored green, critical comments are colored orange, and recommendations colored yellow. Multiple terms were often found within individual COV comments.

distribution of comments employing at least one term per bin, of which multiple bins were often used in tandem. This may appear to artificially inflate the total number of comments.

Another 60 terms were further binned into three coding rubrics to disaggregate the subject of Broadening Participation discussions (i.e. 4 terms related to the *Characteristics of reviewers*; 45 terms related to the *makeup of proposers/awardees*; and 11 terms related to *BP*

Table 2.Outline of MR-related terms found within positive, critical, and recommendations comments. The displayed percentage points quantify the above-average use of each respective bin within the type of comments.

Above-Average Employed in…	Searched Merit Review-Related Terms
Positive Comments	*Broadening Participation (+5%)*
Critical Comments	*Response Length (+16%)* *Weighting (+8%)* *Clarity (+6%)*
Both, Positive and Critical Comments	*Intellectual Merit (Pos +3%, Cri +11%)* *Merit Review (Pos +3%), Cri +9%)*
Recommendations	*Broadening Participation (+7%)* *Clarity (+2%)*

activities), shown in *Figure 13*. These additional bins expanded and clarified trends in the data, and bin-sizes reflect the variation in the related language found in COV comments.

Certain terms were utilized above average in positive, critical, or recommendation comments, shown in *Table 2*. The percentages shown denote how far above average each bin of terms was used. *Broadening Participation* was referred to most often in positive comments. Reviewer *response length*, criteria *weighting*, and *clarity* of criteria were most often mentioned in critical comments. *Intellectual Merit* and *Merit Review* were mentioned at above the average rate in both positive and critical comments, although at higher rates in critical than positive comments. *Broadening Participation* and criteria *clarity* were most often mentioned in COV recommendations. *Broader Impacts* was mentioned at an average rate for all comment categories.

Finally, using 2017 as a natural cut point to, again, bisect the data as previously done with *Figures 2* and *3*, certain bins of terms were found to be employed more often towards the end of the decade. *Figure 9* compares the rates of comments using four bins of terms. All four bins saw an increased, though proportionally inconsistent, usage in critical comments towards the end of the decade, with *Broader Impacts* and *Clarity* terms also seeing large proportional

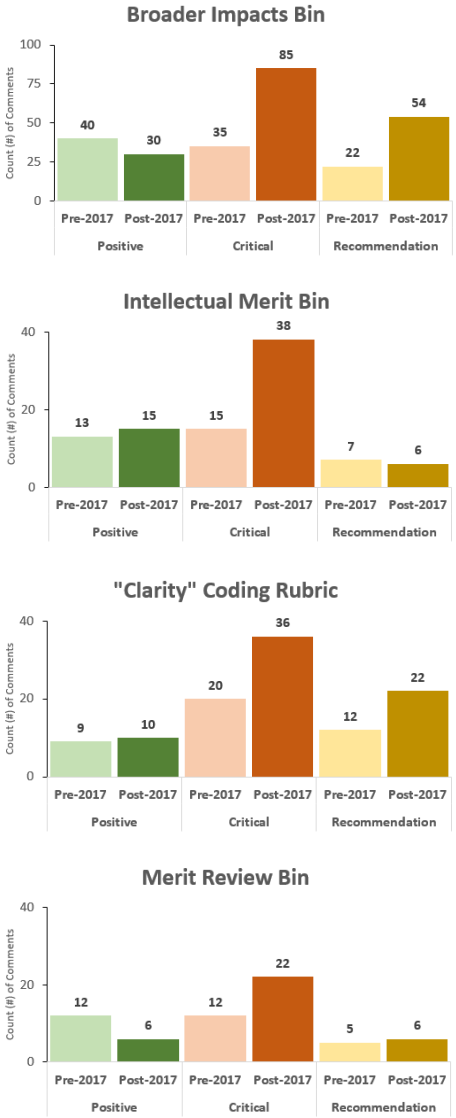


Figure 9. Four charts displaying pre- vs post-2017 rates of comments using at least one term per bin.

increases in recommendation comments. Positive comments more consistently employed terms from these four bins, although the usage of *Broader Impacts* terms fell by one-quarter, and *Merit Review* terms fell by half in the second half of the decade. *Weighting* and *Broadening Participation* bins of terms were used consistently throughout the decade.

Trends Across Bins

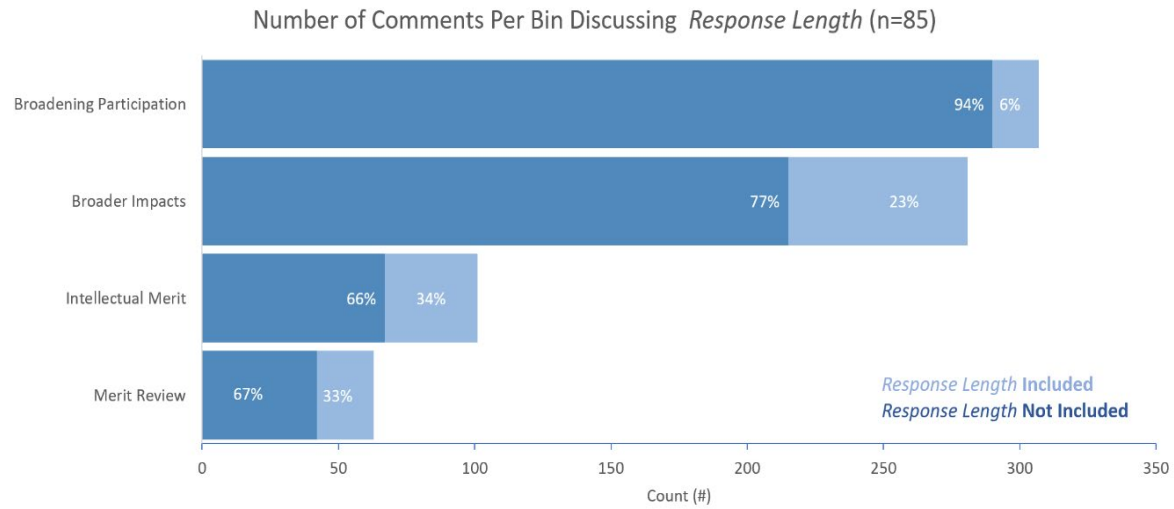


Figure 10. Distribution of comments incorporating various bin of terms which either further include terms within the "Response Length" bin (shown in light blue) or do not (shown in dark blue). Comments on "Intellectual Merit", "Merit Review", and "Broader Impacts" more often additionally discussed the length of reviewer responses to proposals than did comments on "Broadening Participation."

Figure 10 displays the proportion of COV comments across four categories (e.g. BP, BI, IM, and MR), of which the proportion that include *Response Length* terms are shown in light blue and those that do not are shown in darker blue. As previously mentioned, a table of the terms included in the *Response Length* coding rubric may be found in Appendix 1. References to response length are more common in comments on *Intellectual Merit*, *Merit Review*, and *Broader Impacts* than in comments on *Broadening Participation*.

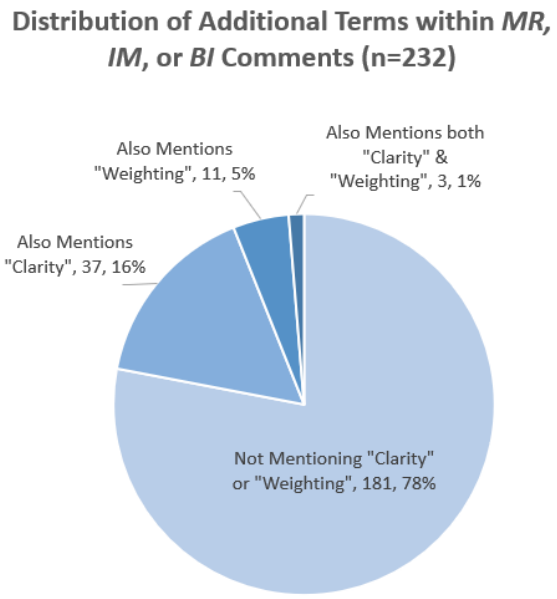


Figure 11. Distribution of COV comments employing MR, IM, or BI-related terms which are also employing "Clarity" or "Weighting" terms.

The COV perception of the *clarity* and *weighted* values of NSF's Merit Review criteria were of particular interest to this

study. Of the 232 COV comments captured from the past decade, 37 (16%) of comments employed terms within the *Clarity* coding rubric, 11 (5%) employed terms within the *Weighting* bin, and a further 3 (1%) employed both *Clarity* and *Weighting* as shown in *figure 11*.

Evaluating comments employing terms within the *Clarity* coding rubric, 111 comments were identified, 80 (72%) of which did not include terms within the *Response Length* coding rubric. Of these 80 comments, the majority, 43 (54%), included terms regarding BP, followed by BI, and any combination of MR, IM, or BI, as shown in *Figure 12*.

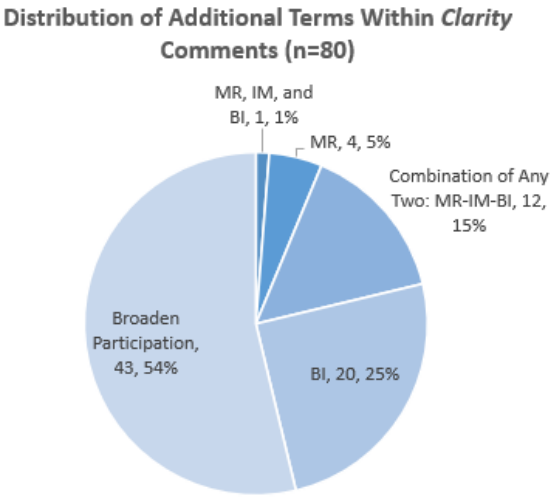


Figure 12. Distribution of COV comments employing terms within the "Clarity" coding rubric which are also employing "Merit Review", "Intellectual Merit", "Broader Impacts", or "Broadening Participation" terms, and combinations thereof.

Observing *Broadening Participation* being the only term mentioned at an above average rate in positive comments, additional analysis was conducted to determine the distribution of BP-related comments. These can be broadly categorized as (1) Proposer or Awardee Characteristics (i.e. proposer demographics; institutional geographics), (2) Reviewer characteristics, and (3) BP activities or actions. *Figure 13* shows reviewer characteristics were minimally discussed. However, comments on proposer or awardee characteristics and BP activities made up 70% of BP-related comments. The remaining large category, "Other," was comprised of BP-related comments not involving proposers, awardees, reviewers, or activities.

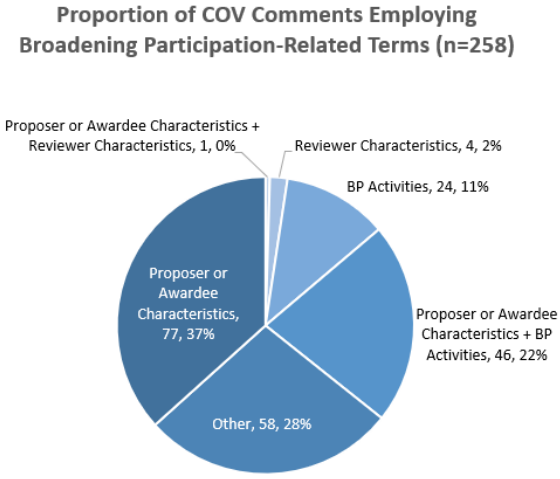


Figure 13. Distribution of terms, related to the subject of comments, discussing "Broadening Participation." Terms included within the coding rubrics for each subject is outlined in Appendix 2.

Discussion

In conclusion, COV reports provide expert perspectives on successes and challenges throughout the MR process. Taking a longitudinal view, we can see how those perspectives have changed since the last MR examination.

Across the agency, there was an overall increasing trend in critical perceptions of the Merit Review process over the past decade, especially after 2017. It is unclear why comments became more critical, but it is possible that improvements in COV panel composition and panelist guidance and training may be influencing factors. Additional study is required to make such a determination.

Merit Review and the IM criterion were perceived more positively than BI, which matches previous Board 2021 Statements noting COVs consistently calling out disparities between IM and BI in written reviews and the Board's subsequent Resolution to add BI experts to COV panels to ultimately improve proposal evaluations. NSF's report to the Committee on Oversight, in December 2023, presented the results of their pilot in response to Board's Resolution. Pilot results suggested an increase in COV quality due to the addition of BI experts on COV panels, and NSF planned to broaden the pilot across all COVs. It is possible that with the addition of BI experts across all COV panels, we could see more critical perceptions of the MR process and possibly more actionable recommendations. In addition, there may be an opportunity to revisit the Core Questions in the COV Template to determine if additional questions relevant to BI are needed.

Broadening Participation was mentioned at above average rates in positive comments, mostly within the context of proposer or awardee characteristics or BP activities. Meanwhile, reviewer response length, criteria weighting, and criteria clarity were most frequently mentioned in critical comments, possibly indicating areas in need of improvement. While additional study is needed to better understand these findings, this initial evidence may reflect concerns regarding the interpretation of the MR criteria to which future analysis of NSF COV responses may shed informative light.

We also saw differences by Division or Directorate. It is not yet clear why some COVs express more satisfaction with the MR process than others. However, this serves as a reminder that each Directorate, Division, and Program implements and assesses MR differently. Additional study is required to better understand these trends, including differences in disciplinary understanding of criteria, reviewer training, and best practices in MR implementation. Pilots are also frequently employed to address COV critiques and recommendations. OIA tracks agency-level MR pilots, but a centralized database tracking MR pilots below the agency level may also aid in understanding the diversity of issues within the MR process.

Understanding the limitations of COV reports and the challenges of text analysis, it is worth the effort to analyze the data collected through this mechanism. Again, COVs provide insight into "the quality and integrity of the merit review process, operations, and technical and

managerial matters pertaining to proposal determinations." This study helps us identify areas of further inquiry that we can suggest to NSF and/or NSB to maximize and improve the use of COVs.

Of additional note, the last time there was a study of the use of the COV mechanism was in 2013. This report included a recommendation to develop annual summary reports of COV reviews across NSF and be provided to program and management staff, shared with Advisory Committees, publicized internally, and made available publicly. However, that annual summary report does not seem to exist, but may be worth considering being initiated. A regular, holistic view of the state of NSF's Merit Review process may be helpful for both NSF and NSB in ensuring the quality and integrity of Merit Review.

Avenues for Further Analysis

The strength of this report surrounds the analysis of trends and patterns in the language employed within COV reports. The next level of analysis is required to further explain the findings found herein. While this report captures many perspectives from COV recommendations, analytics to ascribe the commonality of these recommendations were unattainable.

Areas of exploration could include NSF prompts for COVs and how NSF prompts may have changed since 2011. Are COVs producing greater amounts of critical proportions because NSF is asking for such inputs at increasing rates? Are NSF Directorates and Divisions prompting COVs with similar questions at similar rates or using unique ratios of prompts based on their community's needs? Likewise, a future study could review NSF Responses to COV reports, or discussions within COV reports regarding how COVs perceive NSF alterations based on previous COV recommendations. How do COVs across Directorates and Divisions perceive NSFs responsiveness to their input? Finally, the complete analysis of COV reports from two divisions per directorate, as was originally planned in this study, would be highly informative and is recommended for future analysis efforts.

NSF may elect to assess systematic changes to the Merit Review process based on the data found herein, initially on small scales. Directorates, Divisions, or programs may choose to employ similar structures to more consistently evaluate proposals against the NSF MR criteria, allowing for comparison or the collections of comparable performance metrics.

Finally, the raw data of COV comments collected, reviewed, and analyzed for this report will be made available for review and further analysis.

Appendix 1.

Organization of terms used to search COV reports and identify comments discussing the Merit Review process or criteria. Bins of terms consist of either coding rubrics or collections of variations on root words, or *lemmatization*.

Table 3. Table of terms within each lemmatized bin or coding rubric and used to identify comments in which COVs are discussing the Merit Review process or criteria.

Bins of Terms	Terms Included
Broader Impacts	*Broader Impact, BI, (BI)*
Underrepresented Minorities (URM) & Demographics [Coding Rubric]	*Underrepresented, Under-represented, URM, Minorities, Minority, Marginalized Groups, Diversity, Demographic*
Broadening Participation	*Broadening Participation, Broaden Participation, BP, Participation of*
Intellectual Merit	*Intellectual Merit, IM*
Merit Review	*Merit Review, Merit Review Criteria, Merit Criteria*
Weighting	*Weight, Weighted, Weighting*
Clarity [Coding Rubric]	*Concern, Consistency, Consistent, Inconsistent, Confusion, Clarity, Lack of Clarity, Clarification, Sufficient, Insufficient Information, Improvement, Room to Improve, Bias, Gold Standard*
Response Length (of BI comments from reviewers on proposals) [Coding Rubric]	*Short, Length, long, feedback, substantive, perfunctory, Substance, uneven, one sentence, significant variation, more attention, lack, quality of review, varied considerably, quantitative, explicitly addressed, adequately addressed, general statement, some inconsistencies, how much importance, more attention, included a sentence, reviewer comments, was lacking* **Evaluated terms that produce minimal or no value:** *more important, less consistent, highly variable, enough detail, more detail regarding, neglected to, provide comments, addressed both, quality of, inconsistencies were noted, quality of constructive, written critique, were brief, less insightful, not everyone paid, more developed, quality of comments*

Appendix 2.

Organization of terms within coding rubrics used to identify the subjects of COV comments discussing *Broadening Participation*. Numerous comments employing BP-related terms were reviewed to generate these coding rubrics.

Figure 4. Table of terms used to identify the subject of COV comments discussing Broadening Participation.

Characteristics of Reviewers	panelists and reviewer, pool of reviewer, identify reviewer, lacked diversity
Makeup of Proposers / Awardees	COUNT, Geographic, number of, serving institution, inclusion, HBCU, historically black, MSI, minority serving, of awards, collaborations, of researchers, successful proposals, education opportunities, collaborator, candidate pool, who gets to do it, those who can, awardee portfolio, accessibility, been excluded, reaching out to, membership society, collaborative opportunity, lags behind, mentor applicant, consider mechanism, outreach, improve, women, female, gender, men , male, involve, proposals, award, encourage, attract, small, grow, PIs , background, fewer, engagement, submitted by
BP Activities	education effort, implement, experts, students, pipeline, recruit, their career, developing workforce, capacity building, who apply to, engineering training, capacity

Appendix J: Mathematica Examination of Merit Review at the U.S. National Science Foundation

Mathematica
Progress Together

Examination of Merit Review at the U.S. National Science Foundation

Prepared for the National Science Board-National Science Foundation Commission on Merit Review

December 6, 2024

Kaleen Healey, Shira Haderlein, Tim Bruursema, Lizzie Gellman, Joyce Kim, and Evan Stowe Evans

Submitted to:

U.S. National Science Foundation
National Science Board Office
2415 Eisenhower Avenue
Alexandria, VA 22314
Attention: Faith Hixson, Ann Bushmiller, and Dr. Portia Flowers

Submitted by:

Mathematica
1100 First Street, NE, 12th Floor
Washington, DC 20002-4221
Phone: (202) 484-9220
Fax: (202) 863-1763

Contents

Exhibits

Executive Summary

The U.S. National Science Foundation (NSF) is charged with promoting the progress of science, advancing the nation's health, prosperity, and welfare, and securing the national defense. At the cornerstone of NSF's mission and its investments is its merit review process, which NSF uses to evaluate more than 40,000 proposals annually. The process is currently centered on two merit review criteria: (1) Intellectual Merit, the potential to advance knowledge, and (2) Broader Impacts, the potential to contribute to society and achieve specific, desired societal outcomes, and is guided by the merit review principles and elements. NSF program officers (POs) and external reviewers use the criteria, principles, and elements to evaluate proposals submitted to NSF.

The National Science Board-National Science Foundation (NSB-NSF) Commission on Merit Review (MRX) was convened to assess the efficacy of the current merit review policy, criteria, and processes in supporting NSF's mission. The Merit Review Examination study was designed to support MRX by generating evidence that can inform its recommendations on NSF's merit review policy, criteria, and processes. Data collection activities to obtain that evidence, conducted from mid-January through September 2024, included the following:

- A survey of select NSF staff including POs, division directors (DDs), deputy division directors (DDDs), deputy assistant directors (DADs), and deputy office heads (DOHs)
- Interviews with the NSF leadership—NSF director, NSB chair, and NSF assistant directors (ADs) and office heads (OHs)
- A Request for Information (RFI) published in the Federal Register and in an NSF Dear Colleague Letter, eliciting input from a broad set of people and groups who participate in or are affected by the merit review process, including principal investigators (PIs), reviewers, academic institutions, organizations that submit proposals to NSF, and the general public
- Focus groups with (a) vice presidents of research (VPRs) at academic institutions and (b) people who have served on both an NSF Advisory Committee (AC) and Committee of Visitors (COV); these results cannot be publicly disseminated per Office of Management and Budget (OMB) requirements and therefore are not included in this report.

Overall themes

Themes across responses to the NSF staff survey, NSF leadership interviews, and RFI include:

- **NSF staff survey respondents and RFI respondents were mixed on whether the merit review criteria are clear. A greater proportion of respondents from each group indicated that there was less clarity about the Broader Impacts criterion than the Intellectual Merit criterion.** Although almost all staff survey respondents reported that information on how to assess a proposal against the Intellectual Merit criterion is generally clear, only about half did so for the Broader Impacts criterion. Similarly, RFI respondents who indicated that the criteria were unclear elaborated that there were inconsistent understandings of the criteria among proposers and reviewers, especially on the Broader Impacts criterion. NSF staff survey respondents and RFI respondents offered suggestions for how to improve the clarity of the criteria, which included revising the definitions of the criteria, providing more

guidance and tools to reviewers in rating proposals against the criteria, making Broader Impacts more of a priority, and providing more training for PIs and reviewers.

- **Interviewed leaders said that Intellectual Merit and Broader Impacts are equally important for assessing a proposal's merit, but NSF staff survey respondents and RFI respondents indicated that, in practice, Broader Impacts is given less weight.** The majority of staff survey respondents said that they encountered times when Broader Impacts were given lesser weight relative to Intellectual Merit. RFI respondents cited the undervaluing of Broader Impacts as one reason that full consideration of both criteria is not achieved.
- **NSF staff survey respondents and interviewed leaders reported that there are factors that are important for assessing proposals that are not captured by the merit review criteria.** These respondents noted that solicitation-specific criteria are useful for achieving program-specific goals. They cited additional factors that could be important, for example, factors that could diversify a portfolio (such as institutional and applicant background characteristics) and that capture the relationship between a proposal's Broader Impacts and Intellectual Merit.
- **NSF staff survey respondents and interviewed leaders noted that reviewers' assessments of the criteria are factored somewhat or to a great extent into funding decisions and portfolio management, however staff survey respondents and RFI respondents reported that some reviewers lacked necessary knowledge and training.** Far fewer staff survey respondents said that *reviewers* have a high level of understanding of the merit review criteria compared to the share of respondents who said that *NSF staff* have a high level of understanding. RFI respondents recommended that NSF could improve merit review by improving the quality of panel discussions, diversifying the pool of reviewers, and enhancing reviewer training.
- **NSF staff survey respondents and RFI respondents reported that the merit review principles and elements are clear but indicated room for improvement in reviewers' understanding of the principles and elements and how to use them.** The share of NSF staff survey respondents who said that *PIs* and *reviewers* have a high level of understanding about the principles and elements was much lower than the share who said that *NSF staff* have a high level of understanding. More RFI respondents indicated that the principles and elements were clear than unclear. Some RFI respondents suggested that NSF could provide more guidance to reviewers on how to use the principles and elements.
- **Most NSF staff survey respondents and interviewed leaders agreed that merit review criteria, policy, and processes support a fair and accurate assessment of a proposal's merit.** Among staff survey respondents, White and Asian respondents were significantly more likely than Black or African American respondents to agree that individuals submitting proposals were treated fairly. Staff survey respondents, interviewed leaders and RFI respondents identified sources of potential unfair bias in the merit review process, including PI name recognition, institutional prestige, a lack of diversity among NSF staff and reviewers. RFI respondents identified several barriers that prevent PIs from submitting proposals to NSF, including intimidation about the process, limited guidance from NSF on how to navigate the process, administrative burden of preparing proposals, excessive jargon in NSF guidance, and historically low success rates.

- **NSF staff survey respondents, interviewed leaders, and RFI respondents offered a range of suggestions for how NSF could better assess whether funded projects support NSF's mission to advance scientific knowledge and benefit society.** Suggestions included tracking longer range outcomes, collecting more information about each project's Broader Impacts and about project staff beyond the PI, conducting additional analyses on the data that are already collected, and making final reports publicly available. RFI respondents provided additional suggestions for how NSF could better support awardees in reporting outcomes, including better aligning report sections to each merit review criterion, increasing accountability for reporting by making reports available to the public, reducing the administrative burden on investigators, and improving the tools for reporting that NSF offers PIs.

Key findings by guiding question

Seventeen guiding questions motivated the study and are listed below with applicable key findings.

1. How do PIs, reviewers, and NSF leaders and staff interpret the merit review criteria?

1.1. Although almost all staff survey respondents reported that information on how to assess a proposal against the Intellectual Merit criterion is generally clear, only about half did so for the Broader Impacts criterion. Staff cited unclear standards and insufficient training and expertise as specific barriers in assessing a proposal against the Broader Impacts criterion. Staff reported fielding questions from reviewers and PIs about how to assess Broader Impacts and how to achieve full consideration of both criteria.

1.2. The majority of interviewed leaders said that Intellectual Merit and Broader Impacts are equally important for assessing a proposal's merit.

1.3. The number of RFI respondents who indicated the merit review criteria are clear or unclear were about the same. Respondents who indicated that the criteria are unclear cited inconsistent understandings among proposers and reviewers about how to interpret the criteria, especially Broader Impacts.

2. How do PIs, reviewers, and NSF leaders and staff use the merit review criteria?

2.1. Most staff survey respondents said that, in practice, they encounter times when the two merit review criteria are weighted unevenly. In these cases, the Broader Impacts criterion is often given lesser weight and importance than the Intellectual Merit criterion.

2.2. Interviewed leaders said that full consideration is achieved by considering both merit review criteria at all stages of the merit review process and building a balanced and diverse portfolio across the directorate.

2.3. RFI respondents indicated that full consideration is not achieved due to factors such as reputational bias and the undervaluing of Broader Impacts. They suggested that NSF could improve implementation of the merit review criteria by improving transparency of the review process and recruiting more qualified reviewers. In a separate question, RFI respondents who had reviewed proposals shared their insights about how to improve implementation of the merit review criteria, policy, or processes. Respondents to this question suggested that NSF could improve the quality of panel discussions, diversify the pool of reviewers, and enhance reviewer training.

3. What published and ad hoc guidance does NSF offer to PIs and reviewers on interpreting and using the merit review criteria in preparing or evaluating proposals?

3.1. To make it clearer to reviewers how to assess proposals against the Broader Impacts criterion, staff survey respondents recommended that NSF clarify expectations for Broader Impacts, provide more training on the criterion to PIs and reviewers, make the criterion more of a priority, and strengthen the reporting mechanisms to track it over time.

3.2. The majority of interviewed leaders said NSF staff could participate in NSF trainings and seek guidance from their peers, supervisors, or agency leadership to help them interpret and apply the merit review criteria.

3.3. To improve use of the merit review criteria in preparing and evaluating proposals, RFI respondents suggested NSF provide clearer definitions of the criteria, for example, in the form of a rubric. To improve use of the Broader Impacts criterion, respondents suggested revising the definition and asking reviewers to rate the two criteria separately.

4. How do PIs, reviewers, and NSF leaders and staff understand and interpret the merit review principles and elements?

4.1. The vast majority of staff survey respondents said that NSF staff have a high level of understanding of the merit review principles and elements. The share of respondents who said that PIs and reviewers have a high level of understanding was much lower.

4.2. Interviewed leaders said that POs use reviewer comments on the merit review principles and elements in their funding decisions. The principles and elements are considered among other important priorities— for example, creating a diverse and balanced portfolio and solicitation-specific criteria—in funding recommendations and portfolio management.

4.3. The number of RFI respondents who reported that the principles and elements are clear was greater than those who reported that the principles and elements are unclear, but some respondents suggested NSF provide more guidance on how reviewers should use them.

5. What additional merit review criteria might NSF consider using in the merit review process to better achieve its mission to invest in research that advances scientific knowledge and benefits society?

5.1. About half of staff survey respondents noted additional factors that are important for evaluating proposals that the merit review criteria do not capture. These respondents most commonly cited solicitation-specific criteria, the applicant's background and institutional characteristics that could diversify the characteristics of awardees, more specific guidelines and priorities for Broader Impacts, and the relationships between Intellectual Merit and Broader Impacts as additional factors to consider.

5.2. Interviewed leaders cited solicitation-specific criteria, institutional diversity, and the relationships between Intellectual Merit and Broader Impacts as additional factors to consider.

6. How is reviewer and staff expertise in the merit review criteria assessed?

6.1. Almost all staff survey respondents said that they or their unit are somewhat or very effective at identifying which reviewers have the necessary expertise in the merit review criteria. However, when asked about the overall level of understanding of the merit review criteria among different groups, far fewer respondents said that reviewers and PIs have a high level of understanding of the merit review criteria, compared to the share of respondents who said that NSF staff have a high level of understanding.

6.2. Interviewed leaders commented that POs are responsible for determining which reviewers have the necessary expertise in the merit review criteria to evaluate proposals. They reported that POs make these decisions in consultation with division leaders, by examining reviewers' background and experience, and by relying on their own knowledge of the field.

7. How does reviewer and staff expertise on the merit review criteria factor into reviews, recommendations, and decisions?

7.1. Interviewed leaders described a layered process for how reviewer and staff expertise are evaluated to ensure that they are reviewing proposals and making recommendations that are aligned with the merit review criteria. Reviewers are expected to document the criteria in their reviews, which POs and DDs oversee. POs are expected to document their funding recommendations using the criteria, and these are overseen by division leaders and confirmed via the "DD concur" function.

8. How do reviewers balance the Intellectual Merit and Broader Impacts criteria in assessing proposals?

8.1. RFI respondents reported that Broader Impacts is not assessed comparably to Intellectual Merit, citing the lack of training on Broader Impacts and the lack of expertise among reviewers in aspects of Broader Impacts.

9. How do reviewers' assessments of each merit review criterion factor into NSF program officers' award recommendations?

9.1. Nearly all PO staff survey respondents reported that reviewers' assessments of the Intellectual Merit and Broader Impacts criteria are somewhat or to a great extent factored into POs' funding recommendations.

10. How do reviewers' assessments of each merit review criterion factor into NSF division directors' award decisions for individual proposals? At the portfolio level?

10.1. Most DD, DDD, DAD, and DOH staff survey respondents reported that reviewers' assessments of each merit review criterion are factored somewhat or to a great extent into portfolio management.

10.2. Interviewed leaders specified that POs and directorate leadership generally review funding decisions to understand how proposals are rated and why they are recommended (or not) for funding based on reviewer comments.

11. To what extent does use of reviewers' assessments of each merit review criterion vary within and across NSF directorates?

11.1. There was no statistically significant variation across directorates in the proportion of PO staff survey respondents who indicated that reviewers' assessments of each merit review criterion factor at least somewhat into award recommendations. There was also no variation across directorates in the proportion of DD, DDD, DAD, and DOH respondents who indicated that reviewers' assessments of each merit review criterion factor at least somewhat into portfolio management.

12. To what extent do NSF constituencies perceive the merit review policy and processes to be unfairly biased? How do these perceptions differ by constituency?

12.1. Most staff survey respondents agreed that submitted proposals are evaluated fairly, and the merit review criteria support a fair and accurate assessment of a proposal's merit; however, White and Asian staff were more likely than Black or African American staff to agree that individuals submitting proposals are treated fairly.

12.2. The majority of interviewed leaders agreed that the merit review policy supports a fair and accurate assessment of a proposal's merit, citing NSF policies such as rules on managing conflicts of interest, policies on diversity in reviewers and on projects, and the use of external reviewers. A few leaders disagreed, however, citing implicit bias as a potential reason that a proposal might not receive a fair and accurate assessment.

13. What aspects of the merit review policy and processes are perceived to be biased?

13.1. The majority of staff survey respondents noted aspects of the merit review processes that could introduce unfair bias, citing PI name recognition or background and institutional research infrastructure and prestige as factors that could bias reviews. Moreover, only about half of staff agreed that NSF staff and reviewers are sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process. Some staff also noted aspects of the merit review policy that could introduce unfair bias, explaining that some NSF policies favor proposals from institutions with adequate infrastructure to respond to the proposal requirements and that the lack of anonymizing of proposals can introduce bias from PI or institution name recognition.

13.2. The majority of interviewed leaders named aspects of the merit review processes that could introduce unfair bias, citing institutional research infrastructure and prestige as factors that could introduce bias.

14. What barriers to participation do NSF constituencies perceive in the merit review policy and processes?

14.1. RFI respondents identified several factors that could discourage people from submitting proposals to NSF. These included perceived bias in the merit review process; intimidation due to administrative burden, excessive jargon in NSF guidance, and historically low success rates; limited guidance from NSF on how to navigate the process; and lack of reviewer expertise needed to make an informed assessment of the proposal.

15. To what extent do the reported outputs and outcomes of funded research align with NSF's mission?

15.1. Over half of staff survey respondents reported that the data NSF collects are somewhat or very effective at helping NSF assess whether funded projects support NSF's mission to advance scientific knowledge; fewer than half reported that the data are somewhat or very effective at helping NSF assess whether funded projects benefit society.

15.2. Interviewed leaders reported using a range of sources to assess whether funded projects help support NSF's mission to advance scientific knowledge and benefit society, including interim and final reports, external evaluations, and external review and advisory committees.

16. What guidance does NSF offer to PIs on documenting outcomes that address the merit review criteria in annual and final reports?

This question was only addressed in the focus groups with VPRs and AC members. Per OMB requirements, these results cannot be publicly disseminated.

17. How might NSF better measure the outcomes of funded research with respect to each merit review criterion?

17.1. Staff survey respondents suggested that NSF could better measure the outcomes of funded research for Intellectual Merit by tracking longer-range outcomes of funded projects and going beyond metrics related to publications and citations. For Broader Impacts, staff also suggested tracking longer-range outcomes of funded projects, as well as requiring more information on Broader Impacts from PIs in annual and final reports.

17.2. Interviewed leaders suggested that NSF could better measure the outcomes of funded research by conducting additional analyses on data that NSF already collects and collecting data on project personnel beyond the PI and co-PI.

17.3. RFI respondents suggested that NSF could better support awardees in reporting outcomes of their awards with respect to each criterion. Suggestions included tracking more and longer-range outcomes, adding report sections for outcomes related to each merit review criterion, increasing accountability for reporting by increasing public access to reports, reducing the administrative burden of reporting on investigators, and improving the suite of tools for reporting that NSF offers PIs.

I. Introduction

The U.S. National Science Foundation (NSF) is charged with promoting the progress of science, advancing the nation's health, prosperity, and welfare, and securing the national defense. At the cornerstone of NSF's mission and its investments is its merit review process, which NSF uses to evaluate more than 40,000 proposals annually. The merit review process is considered by the National Science Board (NSB) to be "an international 'gold standard' for review of science and engineering research proposals" (NSB 2005). NSF program officers (POs) with technical and programmatic expertise lead the process with assistance from external experts who help evaluate submitted proposals for the two merit review criteria currently in use: (1) Intellectual Merit, the potential to advance knowledge, and (2) Broader Impacts, the potential to contribute to society and achieve specific, desired societal outcomes.

It is critically important to NSF, NSB, and the nation that NSF implements its merit review process in a fair, thorough, competitive, and transparent way and that those within and outside NSF recognize the process as such. NSB last reviewed the merit review policy in 2011 (NSB 2011). In December 2022, NSB established the National Science Board-National Science Foundation (NSB-NSF) Commission on Merit Review (MRX) assess the efficacy of the current merit review policy, criteria, and processes at supporting NSF's mission.

The Merit Review Examination study was designed to support MRX by generating evidence that can inform its recommendations on NSF's merit review policy, criteria, and processes. The study team worked with the National Science Board Office's (NSBO) contracting officer's representative, technical lead, staff lead, and MRX to develop guiding questions for the study that reflect MRX's interests. With input from NSBO and MRX, the study team developed a plan for collecting and analyzing data from multiple constituencies on their perceptions and implementation of the merit review policy, criteria, and processes. Data collection activities, conducted from mid-January through September 2024, included both internal and external constituencies as noted below and described in Box 1:

- A survey of NSF staff including POs, division directors (DDs), deputy division directors (DDDs), deputy assistant directors (DADs), and deputy office heads (DOHs; late-January through mid-February 2024)
- Interviews with the NSF leadership–NSF director, NSB chair, and NSF assistant directors (ADs) and office heads (OHs; mid-January through mid-February 2024)
- A Request for Information (RFI) published in the Federal Register and in an NSF Dear Colleague Letter eliciting input from a broad set of people and groups including current, past, and prospective NSF proposers, reviewers, and staff; sponsored research administrators and support professionals; representatives of organizations and communities working in or supporting the science and engineering research and education enterprise; members of other communities of practice in the science and engineering research and education fields; and members of the general public (August through September 2024)
- Focus groups with (a) vice presidents of research (VPRs) or similar role at an academic institution, and (b) people who have served on an NSF Advisory Committee (AC) and Committee of Visitors (COV; August 2024); these results cannot be publicly disseminated per Office of Management and Budget (OMB) requirements and therefore are not included in this report.

This document summarizes the findings from the Merit Review Examination study. Box 1 describes the study participants and analyses. The next section lists the study's guiding questions and respective results for each question. Appendix A includes tables and figures displaying results from the survey of NSF staff. Appendix B maps the questions from each data collection instrument to the study's guiding questions. Appendix C contains the data collection instruments.

Box 1. Study overview

Who participated in the study?

NSF staff survey

- All NSF program officers (POs), deputy division directors (DDD), division directors (DDs), deputy assistant directors (DADs) and deputy office heads (DOHs) from the directorates listed below and with permanent federal employee status were invited to complete a 30-minute web survey.
  - Intergovernmental Personnel Act employees, also known as rotators, were not included in the sample as they are not considered federal employees.
  - DADs and DOHs who participated in interviews were removed from the survey sample.
- Staff from the following directorates or offices which fund science, technology, engineering, and mathematics (STEM) research were invited to participate in the survey: Biological Sciences (BIO); Computer and Information Science and Engineering (CISE); Engineering (ENG); Geosciences (GEO); Mathematical and Physical Sciences (MPS); Social, Behavioral and Economic Sciences (SBE); Science, Technology, Engineering, and Mathematics Education (EDU); Technology, Innovation and Partnerships (TIP); Office of Integrative Activities (OD/OIA); and Office of International Science and Engineering (OISE).
- The study team emailed 472 NSF staff an invitation to participate in the web survey.
  - During the four weeks of web survey administration, 241 NSF staff (51 percent) participated in the web survey.
  - Seven NSF staff screened out of the web survey because they had never had relevant responsibilities with merit review, or it had been more than three years since they participated in merit review. We also removed one staff who participated in the leadership interview and was not removed from the survey sample before invitations were sent.
  - See Exhibit 0.1 for staff survey respondent characteristics.

NSF and NSB leadership interviews

- We invited 12 leaders to participate in a 60-minute interview conducted on Zoom.
- We recruited the assistant director (AD) or office head (OH) of each of the directorates and offices funding STEM research (that is, BIO, CISE, ENG, EDU, GEO, MPS, OIA, OISE, SBE, and TIP). We also recruited the NSF director and NSB chair. Nine (75 percent) responded and completed an interview.
  - Interviewed leaders were identified from NSF's public staff directory: https://www.nsf.gov/staff/index.jsp
  - If the AD or OH indicated that they were not able or did not wish to participate in the interview, we asked them to recommend another leader in their directorate or office so that each directorate and office that funds STEM research was represented in the sample.
- The study team used a semi-structured interview protocol, developed with input from NSBO and MRX, for the interviews to gather information related to the guiding questions while being flexible in tailoring questions appropriate to each respondent.

Request for Information (RFI)

- From August 26 to September 20, 2024, the study team collected input from individuals and groups external to NSF, via an RFI published as a Notice in the Federal Register (*Request for Information (RFI) on National Science Board-National Science Foundation Merit Review Commission Review of NSF's Merit Review Policy and Processes*) and disseminated in an NSF Dear Colleague Letter.
- The RFI consisted of six open-ended questions developed by MRX. Responses were collected via Qualtrics.
- The RFI received 130 responses. Responses were counted if they commented on at least one item on the RFI.
- RFI respondents represent a convenience sample, as there were no restrictions on who could respond to the RFI. As a public comment opportunity, key characteristics of the population are unknown. Some respondents identified themselves as proposers to NSF, PIs of NSF awards, reviewers for NSF and other federal agencies, academic institutions, and research and professional organizations.

Focus groups with vice presidents of research (VPRs) and NSF Advisory Committee (AC) members

- In early through mid-August 2024, the study team invited (a) VPRs or equivalent role at an academic institution and (b) people who had served on an NSF AC and Committee of Visitors (COV) to participate in 90-minute focus groups conducted on Zoom (one focus group for each position). This information collection was conducted with approval from the Office of Management and Budget (OMB) under NSF's Generic Clearance for the Collection of Qualitative Feedback on Agency Service Delivery (approval no. 3145-0215). Per OMB requirements, these results cannot be publicly disseminated and therefore are not included in this report.

How was the study conducted?

Descriptive quantitative analyses

- The study team conducted descriptive quantitative analyses of the responses to multiple-choice items on the NSF staff survey.
- After closing the web survey, the study team cleaned the survey data file and prepared the data for analyses. Data cleaning included range verification and skip logic verification to confirm that each variable only contained allowable values.
- As needed, we constructed variables to address small cell sizes. For example, few respondents indicated "Other" as their racial background, so we created a variable for reporting purposes that collapses those responses with another response option.
- The study team calculated an overall survey response rate. Each record was categorized as a complete case, partial response, or nonresponse.
- The study team computed frequency tabulations for each multiple-choice item on the NSF staff survey. Select survey responses are reported by position type, directorate, gender, race, ethnicity, and disability status. We conducted significance testing by using a chi-square statistic on disaggregated results.

Descriptive qualitative analyses

- The study team conducted descriptive qualitative analyses of the interview notes and verbatim responses to 15 free-response items from the staff survey and six free-response items from the RFI.
- We analyzed the interview, staff survey, and RFI data in NVivo using framework analysis (Srivastava and Thomson 2009).
  - This approach entailed first coding the data using a set of broad descriptive codes aligned to the guiding questions (that is, parent codes) then assessing the coded data for data-driven themes (that is, child codes).
  - For many codes, child codes were developed inductively—that is, based on the responses received. After applying the parent codes, the task lead and lead coder independently identified child codes using a sample of responses, met to reach consensus on the child codes, and added them to the codebook.

- A small team of coders received codebook training. Coders each coded one common interview, a set of survey responses, and a set of RFI responses, and then met and compared coding for consistency.
- Once the data were coded, we identified themes within each parent code to answer the guiding questions by counting the number of cases coded to each child code. For each theme, we describe the theme and number of supporting cases and include quotes or examples that illustrate the supporting cases. ◢

II. Findings by Guiding Question

In this section, we list each guiding question followed by relevant key findings from the NSF staff survey, leadership interviews, and RFI. Survey findings from close-ended survey items refer to exhibits in Appendix A showing the results by respondent characteristics; in the key findings we only highlight differences that are statistically significant. Themes from open-ended survey items, interviews, and RFI responses are supplemented with illustrative quotes from respondents.[1]

1. How do PIs, reviewers, and NSF leaders and staff interpret the merit review criteria?

Key findings

1.1. Although almost all staff survey respondents reported that information on how to assess a proposal against the Intellectual Merit criterion is generally clear, only about half did so for the Broader Impacts criterion. Staff cited unclear standards and insufficient training and expertise as specific barriers in assessing a proposal against the Broader Impacts criterion. Staff reported fielding questions from reviewers and PIs about how to assess Broader Impacts and how to achieve full consideration of both criteria.

1.2. The majority of interviewed leaders said that Intellectual Merit and Broader Impacts are equally important for assessing a proposal's merit.

1.3. The number of RFI respondents who indicated that the merit review criteria are clear or unclear were about the same. Respondents who indicated that the criteria are unclear cited inconsistent understandings among proposers and reviewers about how to interpret the criteria, especially Broader Impacts. ◢

According to your understanding of the merit review policy, how important are the following factors for receiving a high rating on a proposal? Staff survey respondents were asked to rate the importance of each merit review criteria, principle, and element for receiving a high rating on a proposal. For the criteria, 99 percent of staff reported that the Intellectual Merit criterion is somewhat or very important for receiving a high rating and 95 percent of staff said the Broader Impacts criterion is somewhat or very important (Exhibit 1.1).

Is it generally clear how a proposal's merit is to be assessed against the Intellectual Merit criterion and the Broader Impacts criterion? Almost all staff survey respondents said that it is generally clear for the Intellectual Merit criterion (96 percent), compared to only about half for the Broader Impacts criterion (53 percent) (Exhibit 1.4).

- POs and DDs were less likely to report that it is clear how to assess a proposal's merit against the Intellectual Merit criterion (96 percent and 71 percent, respectively) compared to DDDs, DADs/DOHs, and other NSF staff positions (100 percent, 100 percent, 100 percent, respectively; $p < 0.05$).

What is unclear about how a proposal's merit is to be assessed against the Intellectual Merit criterion? Nine staff survey respondents received this question because they indicated it was not generally clear how a proposal's merit is to be assessed against the Intellectual Merit criterion. Among these responses, the most commonly cited reason was that the interrelated nature of the two criteria makes it difficult to assess a proposal's merit against the Intellectual Merit criterion (7 responses). For

[1] Examples from the interviews are paraphrased, as the interviews were not recorded.

example, one staff member wrote, "There are instances that are dependent on the goals of the project. The lack of strength of IM [Intellectual Merit] influences the BI [Broader Impacts]. The strength of the BI strengthens the IM of the proposed project."

What is unclear about how a proposal's merit is to be assessed against the Broader Impacts criterion? There were 103 staff survey respondents who were asked this question because they indicated it was not generally clear how a proposal's merit is to be assessed against the Broader Impacts criterion. Themes and examples from these responses include the following:

- Unclear standards make it difficult to assess a proposal's merit against the Broader Impacts criterion (63 responses). Some staff noted that there are not clear standards for assessing Broader Impacts, which makes it difficult to assess clearly and consistently across proposals and reviewers. For example, one commented, "The definition of BI is not explicitly clear to most reviewers (or even to proposers). Over the years, NSF has provided many definitions/examples of what BI is. The fact that this has evolved over time is confusing. Also, reviewers do not have clarity on how to assess the merits of a proposal's BI. Most proposers do not present the BI section of their proposals with the same rigor as they present the IM section, e.g., they do not state gaps/needs, goals, methods, expected outcomes, assessment metrics."
- Lack of training or expertise in Broader Impacts make it difficult to assess a proposal's merit against the Broader Impacts criterion (37 responses). Staff also noted that reviewers and panelists do not typically have formal training in Broader Impacts, which can result in misunderstandings about how to evaluate a proposal against the Broader Impacts criterion. For example, one staff member said, "While most NSF staff understand how to assess Broader Impacts, many reviewers come to the process with a murky concept of what Broader Impacts means. Some think it refers only to outreach or broadening participation, and some think certain specific activities (such as K–12 outreach) are mandatory. The PDs have the opportunity to correct any misconceptions and clarify the meaning of Broader Impacts at the panel, and this often steers the discussion toward a better assessment of the merits but does not always result in updated language in the written reviews."

What, if any, are the types of questions you get from [reviewers and PIs/POs] related to the merit review criteria?[2] A total of 177 staff survey respondents answered this question. Themes and examples from these responses include the following:

- Questions about how to assess Broader Impacts (77 responses). Some staff reported receiving questions about how to assess Broader Impacts, including how Broader Impacts is defined, whether specific activities count as Broader Impacts, requests for examples of Broader Impacts, and how creative or "new" the Broader Impacts activities must be. Some reported receiving questions about whether discussion of societal impacts or scientific impacts were sufficient, or if the proposer needed to have "educational and/or social outreach components."
- Questions about how to achieve full consideration (60 responses). Some staff reported questions about how to balance Intellectual Merit and Broader Impacts, whether one was more important than the

[2] POs were asked about questions they receive from reviewers and PIs; DDs, DDDs, DADs, and DOHs were asked about questions they receive from POs.

other, or if there is a formula for how to weigh them. Examples of questions staff shared include "Can a proposal with excellent BIs but weaker IMs be recommended for funding?" and "Do we have to give equal weight to both Intellectual Merit and Broader Impacts?"

- Few or no questions related to the merit review criteria (19 responses). Some staff noted that they rarely or never receive questions about the criteria. Not all staff elaborated, but those who did noted that reviewers and panelists do not feel the need to ask questions, either because NSF guidance is clear or because they are trained so that they can address the criteria on their own. Some DD, DDD, DAD, and DOH respondents stated that POs do not need to ask supervisors how they should do their work, and that the process relies on POs being empowered to fund projects and build portfolios that are balanced and meritorious.

What do you think NSF is trying to accomplish with the Intellectual Merit and Broader Impacts criteria? Interviewed leaders most commonly cited advancing scientific knowledge and impacting society:

- The criteria aim to advance scientific knowledge (8 responses). Almost all leaders mentioned that a core goal of the criteria is to fund high-quality research that contributes to advancing scientific knowledge. For example, one leader said, "(paraphrased) The goal of using the criteria is for NSF to fund the best proposals that advance knowledge and understanding."
- The criteria aim to impact society (8 responses). Similarly, almost all leaders mentioned the goal of funding science that has positive impacts on society. For example, one leader said, "(paraphrased) To me, the BI criterion is asking PIs to think about the impact of their work, which could be about broadening participation, addressing societal challenges, making international connections, development of students/early career researchers… looking at society and impact."

Do you consider one to be more important than the other, and why or why not? Most interviewed leaders said that they considered both criteria to have equal weight (6 interviewed leaders); however, one leader said they place more weight on Broader Impacts. Two leaders did not provide a response.

- Both criteria are about equal (6 responses). Most leaders said that both criteria arere approximately equal, and that, together, they helped NSF achieve its mission. For example, one leader said, "(paraphrased) No, I don't think one is more important than the other. Together, that's how we realize our mission. As a federal agency, we have a responsibility to provide a benefit to the public, and in particular, our agency has a mission to spur innovation and knowledge, so those things have to come together."

  Broader Impacts is more important (1 response). One leader said that Broader Impacts is more important in their directorate. They discussed how the Broader Impacts criterion is embedded in the mission of their directorate.[3]

The MRX is interested in identifying opportunities to improve NSF's current Merit Review criteria, policy, and processes. Importantly, this includes documenting and understanding any areas of misunderstanding, gaps, or lack of clarity regarding (a) the three Merit Review Principles which are the foundations of the Merit Review Process, (b) *the two statutory Merit Review Criteria which are*

[3] We do not include a quote from this interviewed leader to preserve their anonymity.

***used to evaluate all proposals to NSF*, and (c) the five Merit Review Elements NSF uses to assess each criterion; for instance: Are the Principles, *Criteria*, and Elements clear? Could they be improved upon?**[4] We received 114 responses to this RFI question, but fewer than half of responses addressed the merit review criteria. Among these responses, 11 indicated they had submitted a proposal to NSF, 2 indicated they had served as a reviewer, and 21 had experience both as a proposer and reviewer. Themes in the responses related to the merit review criteria include the following:

- Criteria are clear (19 responses). Multiple respondents indicated that both the Intellectual Merit and Broader Impacts criteria are clear. As one respondent explained, "In my experience submitting and reviewing proposals to multiple federal agencies, the NSF guidelines and criteria are the most clear and fair. I don't see a need to revise them."
- Expectations for meeting the criteria are unclear (19 responses). Other respondents explained that proposers and reviewers are often unclear on what is needed to meet the criteria. The majority of responses here (14 responses) were in reference to the Broader Impacts criterion; these respondents suggested NSF could define Broader Impacts more explicitly and within the context of specific disciplines. As one respondent explained, "In my opinion, it would be helpful to define Broader Impacts more precisely, perhaps to provide some examples. Clearly, Broader Impacts of a proposal depends on the discipline: it is easily described if one works, say, on long COVID, and harder if one is a pure mathematician. Speaking of the latter, does outreach activity qualify? What about service to the profession such as editing journals, organizing conferences, serving on scientific advisor boards, etc.?" A few respondents lamented what they saw as reviewers' focus on innovation in Broader Impacts. One explained, "I view NSF's advocacy of Broader Impacts as vague and unproductive. The criteria with which activities should be evaluated are unclear. It is *very* common for panels to undervalue realism, rewarding grandiose but unworkable plans, and there seems to be a near-total lack of follow-up from NSF about implementation."
- Expectations for balancing Intellectual Merit and Broader Impacts are unclear (13 responses). Some respondents shared that although the individual criteria are clear, it is not always clear how they should be balanced in proposals and reviews. One respondent explained, "As a preparer of proposals I do not think I ever understood the balance expected between IM and BI. As a reviewer of NSF proposals, I had the same uncertainty. I knew BI had to be there, but given that most proposals did not develop this aspect of the work as [much] as the IM aspects, it was unclear what to hold the PI(s) responsible for in terms of BI."
- Metrics for assessing the criteria would be helpful (9 responses). Several respondents offered suggestions for how the criteria could be improved. Particularly for Broader Impacts, some indicated NSF could create metrics for assessing the criteria and require more systematic reporting about criteria-related outcomes. One respondent explained, "We encourage the Foundation to consider clarifying in the proposal process that proposals include as part of Broader Impacts a specific outcome, or proxy outcome, and plan for measuring that outcome(s) within the proposed activities that must then be reported on in the award's Final Report."

[4] Italicized text indicates the parts of the item most relevant to the guiding question.

2. How do PIs, reviewers, and NSF leaders and staff use the merit review criteria?

Key findings

2.1. Most staff survey respondents said that, in practice, they encounter times when the two merit review criteria are weighted unevenly. In these cases, the Broader Impacts criterion is often given lesser weight and importance than the Intellectual Merit criterion.

2.2. Interviewed leaders said that full consideration is achieved by considering both merit review criteria at all stages of the merit review process and building a balanced and diverse portfolio across the directorate.

2.3. RFI respondents indicated that full consideration is not achieved due to factors such as reputational bias and the undervaluing of Broader Impacts. They suggested that NSF could improve implementation of the merit review criteria by improving transparency of the review process and recruiting more qualified reviewers. In a separate question, RFI respondents who had reviewed proposals shared their insights about how to improve implementation of merit review criteria, policy, or processes. Respondents to this question suggested that NSF could improve the quality of panel discussions, diversify the pool of reviewers, and enhance reviewer training. ◢

In your experience, how do reviewers rank the importance of the following when assigning a high rating on a proposal? Staff survey respondents were asked to rate how *reviewers* rank the importance of each merit review criteria, principle, and element when assigning a high ranking on a proposal (Exhibit 2.1). More staff rated the Intellectual Merit of a proposal as somewhat or very important to reviewers (99 percent), and fewer staff rated the Broader Impacts of a proposal as somewhat or very important to reviewers (76 percent).

- On average, more staff from BIO, SBE, and EDU (83 percent, 100 percent, and 96 percent, respectively) reported that reviewers treat Broader Impacts as somewhat or very important than staff from GEO, TIP, and OIA (56 percent, 63 percent, and 60 percent, respectively; Exhibit 2.2).

The NSF Proposal & Award Policies & Procedures Guide notes that the two merit review criteria (Intellectual Merit and Broader Impacts) are to be given full consideration during the review and decision-making processes. Each criterion is necessary, but neither, by itself, is sufficient. In practice, how is full consideration achieved to produce a single [recommendation/decision] for an award?[5] There were 187 staff survey respondents who answered this question; themes from these responses include the following:

- Full consideration is achieved by considering both merit review criteria at all stages of the merit review process (72 responses). Numerous staff indicated that both criteria are considered at all stages in the merit review process, from reviewer feedback to DD concur. For example, one respondent commented, "Full consideration of IM and BI is requested of the individual reviewers in their written reviews, and of the panel in their discussions. Program directors then give full consideration to these in their recommendations, which the DD and DDD consider in their concurrence."
- Full consideration is achieved by accounting for both criteria in panel discussions (52 responses) and reviewer feedback (50 responses). Similarly, some staff mentioned the importance of both criteria in reviewer feedback and panel discussions, which are the basis of funding recommendations and

[5] POs were asked about a recommendation; DDs, DDDs, DADs, and DOHs were asked about a decision.

decisions. For example, one staff member said, "Panelists and reviewers play a major role in the review process. Instruction about both review criteria is provided early in the review process and clearly to the panelists. Panel briefing provides a clear explanation and instruction of BI. The BI of each proposal is also discussed by the panel following the discussion of IM. After the panel, POs of a program consider the panel recommendation in view of both IM and BI to make recommendations."

- As a counter example, 35 staff indicated that they did not feel that Broader Impacts is prioritized, suggesting that at times, full consideration is not achieved. For example, one staff member commented, "BI is generally ignored, and readers (reviewers or panelists) focus primarily on the IM. POs here also weigh the IM WAY more heavily than the potential for a proposal to benefit society or the economy, or at least that is how it is in [directorate]."

In your experience, are there situations in which the importance of the Intellectual Merit criterion and the Broader Impacts criterion are weighted unevenly? The vast majority (91 percent) of staff survey respondents responded affirmatively, meaning that there are times when the criteria are weighted unevenly (Exhibit 2.3).

In your experience, how frequently are the merit review criteria weighted unevenly? Response options included never, rarely, sometimes, and always. Overall, over one-quarter of staff survey respondents (27 percent) said that the criteria are *always* weighted unevenly, two-thirds (66 percent) said that they are *sometimes* weighted unevenly, 8 percent said that they are *rarely* weighted unevenly, and 0 percent said they are *never* weighted unevenly (Exhibit 2.4).

- There were differences in response patterns by directorate. For example, 63 percent and 44 percent of staff from TIP and MPS, respectively, said the criteria are *always* weighted unevenly, compared to 19 percent and 20 percent of staff from GEO and OD/OIA, respectively.

What weight do you typically place on the importance of the Intellectual Merit criterion compared to the Broader Impacts criterion in your use of the merit review criteria? Many staff survey respondents (61 percent) reported they place more weight on Intellectual Merit than on Broader Impacts. Only a small number (5 percent) reported placing more weight on Broader Impacts (Exhibit 2.5 and Exhibit 2.6).

How challenging do you generally find it to assess a proposal's merit against each criterion? More staff survey respondents reported that it is not challenging to assess a proposal's merit against each Intellectual Merit criterion (82 percent) compared to Broader Impacts (61 percent; Exhibit 2.7).

Have you had any recurrent challenges assessing a proposal's merit against each criterion? Over one-quarter of staff survey respondents (28 percent) indicated that they have had recurrent challenges and nearly all of these respondents (60 of 64 who indicated that they have had recurrent challenges) (Exhibit 2.8) provided an answer to the follow-up question: **What recurrent challenges have you had assessing a proposal's merit against each criterion?** Sixty staff received this question because they answered yes to the previous question. The most commonly cited challenges include the following:

- Lack of clarity about Broader Impacts is a challenge (27 responses). Almost half of staff who answered this question mentioned that the standards for assessing the Broader Impacts are unclear, making it challenging to assess a proposal's merit against the criterion. Respondents discussed several pain points related to this, including reviewers' misunderstandings of or inconsistent standards for Broader Impacts; a lack of accountability for projects to achieve their proposed Broader Impacts; and that, generally, Broader Impacts is not prioritized in proposals and reviews. For example, one respondent commented, "The reviewers tend to provide very divergent discussion and ratings of the Broader Impacts in a proposal. This make[s] it difficult to judge them. I often receive Broader Impact ratings that range from E to F on the same proposal."
- Finding qualified reviewers is a challenge (17 responses). Numerous staff wrote to the challenge of recruiting reviewers with the necessary expertise—either on their topic or on the two criteria—to review proposal and serve on panels. For example, one staff member said, "For the intellectual merit, the challenge faced in some panels is that, despite my best efforts in recruiting panelists, some panelists do not have the depth of understanding necessary to fully evaluate the IM. For the broader impacts, the problem is that some reviewers, particularly in their individual reviews, do not assess the BI against the review elements, but rather summarize the proposed activities or provide vague statements about the merit. Such reviewers must be reminded in the panel discussion that the BI must be assessed using the review elements."

What, if anything, about the scientific field(s) of your directorate, division, or program makes it challenging to apply the Intellectual Merit criterion? Fifty-five staff survey respondents received this question because they answered yes to a previous question indicating there is something about the scientific field of their directorate, division, or program that makes it challenging to apply the Intellectual Merit criterion (Exhibit 2.9). Themes from these responses include the following:

- It is challenging to apply the Intellectual Merit criterion to broadening participation projects (7 responses). Several staff commented that it can be difficult to assess the Intellectual Merit on a project that is more focused on Broader Impacts, such as broadening participation projects. For example, one person wrote, "For programs that are focused on BI-specific priorities—like building research capacity or broadening participation—then it can be challenging to weigh extremely strong BI against less strong IM."
- It is challenging to apply the Intellectual Merit criterion to projects where IM and BI are intertwined (5 responses). Staff also wrote about how the two criteria are often intertwined, especially in some directorates. In these cases, it can sometimes be challenging to differentiate a proposal's merit against each criterion distinctly. For example, one person commented, "IM can be reliant on BI and BI can be reliant on IM in [directorate]. Strengths under IM could just as easily be under BI and vice versa."

What, if anything, about the scientific field(s) of your directorate, division, or program makes it challenging to apply the Broader Impacts criterion? Sixty-two staff survey respondents received the question only if they answered yes to a previous question indicating there is something about the scientific field of their directorate, division, or program that makes it challenging to apply the Broader Impacts criterion (Exhibit 2.9). Themes from these responses are discussed below:

- It is challenging to apply the Broader Impacts criterion when Broader Impacts are not valued (16 responses). Some staff wrote to the challenge of applying the Broader Impacts criterion when they don't feel that Broader Impacts are valued. Respondents discussed the pressure to give a strong rating on Broader Impacts for proposals that have strong Intellectual Merit, misconceptions about Broader Impacts from reviewers and other NSF staff, and the difficulty of evaluating Broader Impacts authentically within the structure of merit review. For example, one person wrote, "The treatment of BI as a checklist can be alarming, especially when outreach efforts are not meaningful and mutually benefiting or the integration of research and education is lacking details to determine the appropriateness of the effort/activities."
- It is challenging to apply the Broader Impacts criterion when the outcomes of Broader Impacts are harder to assess and track (8 responses). Several staff commented on the difficulty of assessing and tracking the outcomes of Broader Impacts, either because the outcomes might be harder to measure or be realized many years after the award has been completed, or because it can be difficult to assess whether the proposed Broader Impacts are well suited to the project to achieve maximum impact. For example, one person said, "Scientists tend to fall back on 'checklist' items in Broader Impacts, rather than assessing how, if the project is successful, society at large will be the beneficiary. They therefore look for checklist items like recruiting URMs [underrepresented minorities], K–12 outreach, etc. These are certainly meritorious, but not necessarily the best suited impacts to every project."

In practice, how is full consideration achieved in your leadership of the [directorate's/office's] portfolio? Similar to responses from the staff survey, interviewed leaders commented that the two criteria are considered at all stages of the merit review. In addition, they discussed working to create a balance of the criteria across the directorate's/office's portfolio. These themes are elaborated on below:

- Full consideration is achieved by considering both merit review criteria at all stages of the merit review process (5 responses). Leaders talked about how the two criteria guide all stages of the merit review process, from reviewer feedback to portfolio management. For example, one person said, "(paraphrased) We think about these two aspects at all levels of portfolio, starting with projects we fund and how we scope programs, what the portfolio of programs should be. We also think about level of decision making too in terms of programs—looking at what opportunities there are to advance science and engineering, what opportunities we're making available to enable a positive impact."
- Full consideration is achieved by building a balanced and diverse portfolio across the directorate (3 responses). Interviewed leaders also talked about the need to build a diverse and dynamic portfolio within their directorates. For example, one person said, "(paraphrased) The beauty of the NSF system is the POs have the capability of building a portfolio. So, we want to balance institution types that have potential for group investments in a field. That may be weighting one criterion more in one proposal than the other. But we never have one that's great on Intellectual Merit and terrible on Broader Impacts: it has to have both."

Tell me a bit about applying the merit review criteria in your [directorate/office] or program specifically. In response, five interviewed leaders said that the Broader Impacts criterion is easier to apply in their directorate relative to other disciplines, and two leaders said that the Broader Impacts criterion is more difficult to apply. These themes are expanded upon below:

- The Broader Impacts criterion is easier to apply (5 interviewed leaders). Five leaders said that the Broader Impacts criterion is easier to apply in their directorate or office, often because it is directly built into the focus of their scientific field. For example, one person said, "(paraphrased) I would say for us in [directorate name], it's not a problem at all because so much of what we invest in impacts people's daily lives … the community we support really gets the societal relevance side, so we see really interesting things people propose … they are very well-steeped in the importance of what they do, so for us we don't have that problem because people look outside and see the impacts."
- The Broader Impacts criterion is more difficult to apply (2 interviewed leaders). In contrast, two leaders commented that the Broader Impacts criterion is more difficult to apply in their directorate or office. One person, for example, said that the broadness of their directorate or office meant that they sometimes have trouble finding reviewers with the necessary expertise in Broader Impacts, because they spend most of their focus finding experts with topical expertise (which more closely corresponds with Intellectual Merit). Another leader said, "(paraphrased) Broader Impacts is very broad. I think that's intentional…. It can be a challenge for people therefore to understand what's expected of them. In [directorate name], a lot of the field is intrinsically use-inspired…. [It] has a built-in Broader Impact there. People might think, I'm doing [topic of study], that's a Broader Impact, that should be good enough. One thing we want people to understand, for example with PIs, it's do you have Broader Impacts? Not is it acceptable? Is it good enough?"

What, if anything, about the scientific field of your directorate, division, or program makes it challenging to interpret and apply the merit review criteria? Three interviewed leaders said the multidisciplinary nature of their field makes it difficult to interpret and apply the merit review criteria, and two said that the criteria are overly broad, which makes it difficult to apply in their field. Examples of these themes are as follows:

- The multidisciplinary nature of the directorate can be a challenge (3 responses). Several leaders mentioned that the multidisciplinary nature of their field is a challenge because, for example, it makes it more difficult to find reviewers and panelists with expertise on such a broad range of topics, or because people from different disciplines might have different conceptions of what the criteria mean. One respondent said, "(paraphrased) We cover the entire breadth of NSF disciplines, so that's a challenge that other directorates don't have to face. In order to cover what we consider our mission of innovation for Intellectual Merit, and the impact on society, we are bringing in new types of reviewers, people who have expertise in building innovative ecosystems, translating research, even from the venture capital community … the terminology around IM and BI is pretty academically focused … so, part of the challenge is teaching reviewers about what we mean."
- The broadness of the merit review criteria can be a challenge (2 responses). A few leaders also noted that the criteria are purposely broad to all of the fields that are funded by NSF, but that this can sometimes make it difficult to apply them to a specific discipline. For example, one interviewed leader said, "(paraphrased) So the criteria are broad because we want to cover all of science and engineering… But at same time, trying to give proposers opportunity to pitch the ideas they want, that they feel like motivates them, that they're experts on, what they think they can contribute…. So they're intended to be broad, which could create a challenge on expectations for people."

NSF strives to conduct a fair, competitive, transparent Merit Review process for the selection of projects. To accomplish this, NSF relies on a process that considers both the technical aspects of a proposed project and its potential to contribute more broadly to advancing NSF's mission using the statutory Intellectual Merit and Broader Impacts Merit Review Criteria. MRX invites suggestions on the implementation of the Merit Review Criteria. We especially invite feedback that would (a) clarify how the Merit Review Criteria can be used in preparing and reviewing proposals, (b) *ensure proposals, reviews, and funding decisions demonstrate full consideration of both criteria while maintaining openness to the full spectrum of potential activities under each*, and (c) *better recognize and support potentially transformative and high-risk/high-reward activities*. We received 94 responses to this RFI question, where respondents shared their thoughts on full consideration as well as recommendations on how to achieve it.

- Full consideration is not achieved (23 responses). Numerous respondents indicated that full consideration of both criteria is not achieved in proposals, reviews, or funding decisions. About half of these responses came from respondents with experience as a proposer (11 responses), and another significant portion came from those with experience as both a proposer and a reviewer (9 responses). For example, one respondent shared, "I find the current Merit Review Criteria clear and effective in theory. However, in practice, the implementation of the review process often suffers from issues like favoritism, review bias, and the selection of inappropriate reviewers." Others expressed concern over the imbalance between the consideration of Intellectual Merit and Broader Impacts. One respondent said, "We are concerned that Broader Impacts are not taken seriously by review panels and are often treated as perfunctory." Another said, "With quality and advancing the frontiers as the very first principle, Intellectual Merit is foregrounded as the primary criterion."

Themes in these responses related to how NSF can ensure full consideration of both criteria included the following:

- NSF could make the merit review process more transparent (12 responses). Some respondents discussed the lack of and need for more transparency within the merit review process. Without this transparency, they said, it is difficult to know whether proposals, reviews, and funding decisions demonstrated full consideration of both criteria. Of these respondents, the majority reported experience as both a proposer and reviewer (9 responses), and the remainder reported experience as a proposer only (3 responses). Respondents offered a range of suggestions for how to improve transparency. Some suggested that NSF could release declined and accepted research proposals "to the public (of course, sans intellectual property content), so that the whole world and history [could] be the witness to what or who deserves Merit or not." Another respondent shared that releasing proposals to the public could "help educate researchers writing proposals and reviewers by publicizing best practices in broader impacts and case studies of exemplars from past NSF-funded initiatives."
- NSF could improve reviewer quality (9 responses). Several respondents, including four with experience as a proposer and three with experience as both a reviewer and proposer, suggested NSF consider ways to improve the effectiveness and quality of the reviewers involved in merit review. Six respondents specifically raised concerns that reviewers are unqualified or lacked the expertise to rank proposals, often with "zero educational background or relevant work experience in the subject matter." As one respondent flagged, although some reviewers might recuse themselves for lacking subject matter

expertise in a field like computer science, others participate, only for proposals to be "validated and rejected repeatedly by reviewers who have doctorates in unrelated fields such as metallurgy or chemistry."

- Other suggestions for how NSF could improve merit review. Respondents suggested several other ways that NSF could improve merit review, including adding or revising existing questions for reviewers to respond to, adding checks to ensure the fairness of reviews, using AI to reduce burden and increase the efficiency of the process, implementing a double-blind review system, and implementing a scoring rubric to help reviewers rate proposals. For example, one respondent suggested "adding a question for reviewers to rate their own level of knowledge with the subject matter … as a way for panelists to assess and weight ad hoc reviews." Another said that NSF should "incorporat[e] merit review guidance that evaluates the extent to which a proposed project is inclusive, accessible, and grounded in and responsive to community concerns."

Multiple respondents also offered suggestions for how NSF could better recognize and support potentially transformative and high-risk or high-reward activities:

- NSF could be more welcoming to nontraditional researchers (8 responses). Some respondents recommended NSF could take steps to ensure a wider variety of institutions, regions, and scholars of different academic, demographic, or experiential backgrounds have the opportunity to receive funding and advance transformative research. As one respondent said, "By intertwining the concepts of transformative ideas and inclusive representation, the NSF can create a more equitable landscape for breakthrough science." Two respondents also mentioned the need for NSF to diversify panel reviewers by "incorporating cross-disciplinary reviewers" and "rotat[ing] panel reviewers more," which could help expand the reach of NSF funding.
- NSF could take more risks (8 responses). Some respondents recommended NSF consider changes to better enable and account for high-risk projects. Most of these respondents had experience both as a proposer and reviewer (6 responses), while the rest each had experience as just a proposer or reviewer (1 response each). For example, one respondent suggested NSF should "broaden its definition of 'moonshot' to encompass [both] high-risk ideas … and high-potential individuals who have been historically underrepresented in their fields." Another said that NSF could provide guidance "regarding how to manage a proposal that reviewers find transformative but maybe does not rank well in other areas (BI for instance)." Clarity on this front could help reviewers balance the priorities between the two criteria and other aspects of the review process.

MRX is interested in the experiences and perspectives of those who have reviewed proposals submitted to NSF. We invite you to share your insights and describe any opportunities you believe would improve implementation of the Merit Review criteria, policy, or processes based on your experience reviewing NSF proposals. We received 71 responses to this RFI question. Themes related to how NSF could improve implementation of the merit review process include the following (Exhibit 2.10 and Exhibit 2.11):

- NSF could improve the quality of panel discussions (11 responses). While some respondents shared positive experiences serving as a panel or ad hoc reviewer, others indicated that the reviewer experience—especially on panels—could be improved. Seven of these respondents indicated they had

served on review panels and four were ad hoc reviewers; five had experience reviewing within the past five years. Specifically, respondents suggested holding panel discussions in person, implementing evaluation rubrics, and implementing better training on Broader Impacts for reviewers. A few respondents said that in their experience, panel discussions can sometimes be derailed by one or two reviewers with unsubstantiated arguments, and that this can unfairly influence the ratings of the proposals. Other respondents shared that because reviewers are not required to update their initial reviews after panel conversations, proposers can sometimes feel that the panel did not fully understand their proposal.

- NSF could increase the diversity of reviewers (10 responses). Respondents also highlighted the lack of diversity of reviewers as a way that NSF could improve the implementation of merit review. Of the respondents who spoke to this theme, five served on panels and five served as ad hoc reviewers; four had experience reviewing within the past five years. Most commonly, respondents cited racial and ethnic diversity, but some also mentioned diversity of professional experience and expertise. For example, one respondent suggested that NSF should "consider exploring ways to broaden who serves as a reviewer, for example, organize opportunities where experienced reviewers of various backgrounds meet with less experienced reviewers on their campuses or at associations to explain the process, mentor, and answer questions." Other respondents said that NSF should recruit reviewers with specific expertise in Broader Impacts and experience working with interdisciplinary teams and community partners.
- NSF could improve reviewer training (7 responses). Some respondents said that reviewers need better training, especially in Broader Impacts and implicit bias. Four of these respondents reported serving on a panel and three served as an ad hoc reviewer; four had experience reviewing within the past five years. One respondent said, "Reviewers need quality training on how to review BI prior to engaging in the review process. ARIS [Center for Advancing Research Impact in Society] would be an excellent place to develop that training, which should be required by all reviewers." Another respondent explained that because panelists are experts in their field, they can proficiently assess Intellectual Merit, but panelists do not necessarily have the expertise to assess Broader Impacts. A couple respondents also recommended improved implicit bias training for panelists and ad hoc reviewers, especially because reviewers are often less diverse than proposers. This, according to one respondent, "leads to the same researchers and institutions receiving funding." Although most respondents who discussed reviewer training indicated it was unsatisfactory, one said it was adequate.
- NSF could reduce favoritism and bias among reviewers (7 responses). Relatedly, some respondents expressed concern about reviewers favoring certain factors—like the institution where the proposer worked—in their reviews. Of respondents who spoke to this theme, three reported serving as a panelist and four reported serving as an ad hoc reviewer. Respondents worried that reviewers' implicit bias could lead to proposers from some institutions being funded over others, particularly those from Historically Black Colleges and Universities (HBCUs). For example, one respondent said, "We have heard panelists bring in personal biases and project their own experiences on the proposers, thus voting down proposals because they are 'not doable' or too risky."

- NSF could improve consistency between reviews (7 responses). Some respondents highlighted that reviewers can be inconsistent, which leads to feelings of confusion and unfairness among some proposers. Of the respondents who spoke to this theme, three reported serving as a panelist and four as an ad hoc reviewer. As one respondent said, "Reviews are not ineffective most of the time, but there could be more consistency in the reviews at least based on what I see in my program." Another respondent suggested standing panels could improve consistency because reviewers would serve through multiple review cycles.

3. What published and ad hoc guidance does NSF offer to PIs and reviewers on interpreting and using the merit review criteria in preparing or evaluating proposals?

Key findings

3.1. To make it clearer to reviewers how to assess proposals against the Broader Impacts criterion, staff survey respondents recommended that NSF clarify expectations for Broader Impacts, provide more training on the criterion to PIs and reviewers, make the criterion more of a priority, and strengthen the reporting mechanisms to track it over time.

3.2. The majority of interviewed leaders said NSF staff could participate in NSF trainings and seek guidance from their peers, supervisors, or agency leadership to help them interpret and apply the merit review criteria.

3.3. To improve use of the merit review criteria in preparing and evaluating proposals, RFI respondents suggested NSF provide clearer definitions of the criteria, for example, in the form of a rubric. To improve use of the Broader Impacts criterion, respondents suggested revising the definition and asking reviewers to rate the two criteria separately. ◢

In your opinion, what could be done to make it clearer how a proposal's merit is to be assessed against the Intellectual Merit criterion? Nine staff survey respondents received this question only if they answered no to the previous question: **Do you think it is generally clear how a proposal's merit is to be assessed against the Intellectual Merit criterion and the Broader Impacts criterion?** Themes from these responses include the following:

- NSF could clarify expectations or be more prescriptive about how a proposal's merit is to be weighted against the Intellectual Merit criterion (3 responses). As one staff member commented, NSF could "Explain to the panelists and ad hoc reviewers the goals of any given program and the role IM plays, and to incorporate that into their review." Another noted that NSF should acknowledge *all* ways that knowledge can be created, not just those that are "creative, original, or transformative."
- NSF could require reviewers to support their assessments with evidence (2 responses). Two staff noted that requiring reviewers to provide evidence from the proposal in support of their assessments would underscore the importance of each criterion. As one explained, "Assessment of merit must be justified—that is, it is the responsibility of the reviewer to make claims and to support those claims with evidence."
- NSF could better draw on PO's field awareness (2 responses). Similarly, two staff members noted the importance of relying on PO understanding of issues and trends in their respective field. One stated, "Maybe the best asset of the agency are experienced program officers with knowledge in the field."

- No additional guidance is needed (2 responses). Alternatively, two staff members noted that there is no need for more clarity in how to assess a proposal's merit against the Intellectual Merit criterion, because PO discretion is an important aspect of the process. One said, "One of the strengths of the NSF merit review process is that it calls on reviewer and PO expertise. So there is discretion in the system." These respondents emphasized that reviewers are responsible for making claims and supporting them with evidence, but that POs should use their knowledge of the field to assess these claims and make decisions.

In your opinion, what could be done to make it clearer how a proposal's merit is to be assessed against the Broader Impacts criterion? A total of 102 staff survey respondents received the question because they answered no to the previous question. Themes from these responses include the following:

- NSF could clarify expectations or be more prescriptive (47 respondents). Some staff noted a need for brief yet clear information for reviewers and PIs to reference on what NSF considers to be a Broader Impact. One staff member noted, "Describe clearly, in the context of an academic researcher, what is meant by BI in terms of nature of activities in contrast to the usual faculty work, potential outcomes, ways to assess success of outcomes, notions of novelty and innovation, and the difference between various types of impacts, such as societal impact, research impact, education impact, and outreach impact." Another suggested NSF could "create categories of Broader Impacts and include two or three bullets that describe reasonable expectations for each; require reviewers and program directors to identify which category(ies) of Broader Impacts a proposal addresses and to what extent (perhaps a Likert scale)."
- NSF could make Broader Impacts more of a priority for PIs and reviewers (20 responses). Staff emphasized that Broader Impacts should be prioritized to a greater extent, with varying suggestions for how to do so. For example, one staff member said, "Maybe it should be required that a certain number (or percentage) of proposal pages be devoted to Broader Impacts. It should be clear that funds should be requested to support this activity. More proposals should be declined based on inadequate Broader Impacts." Other staff suggested rating Intellectual Merit and Broader Impacts separately to encourage more attention to Broader Impacts.
- NSF could provide more training for PIs and reviewers (18 responses). Some staff emphasized the importance of better outreach and training on Broader Impacts to both reviewers and PIs. According to one staff member, "We need more robust training and communication with the PI community, especially the PIs and institutions who are not our "frequent flyers": new PIs, new institutions, MSIs [Minority Serving Institutions], teaching-centered institutions, and EPSCoR [Established Program to Stimulate Competitive Research] states."
- NSF could track Broader Impacts outcomes over time (14 responses). Staff suggested tracking outcomes over time to ensure follow-through on proposed Broader Impacts activities. Others suggested reporting on outcomes so that others see what PIs do in the Broader Impacts space: "I think NSF needs to start emphasizing and reporting on the Broader Impact activities that PIs are accomplishing. Right now, all that seems to come out is 'science in the news.' We do not seem to have articles promoted by NSF that acknowledge amazing Broader Impact activities that PIs are doing with NSF funding."

What instructions or training do you provide to Program Officers and Division Directors to help them interpret and apply the merit review criteria? Themes from interviewed leaders' responses include the following:

- Staff can participate in NSF training (6 responses). Leaders emphasized the NSF training, Merit Review Basics, as a main resource for POs and (if they are interested) DDs. Leaders described the Basics course as a four-day series about how the process works. They added that there is also a shortened version of the training focused on division leadership for DDs and DDDs. Some leaders referred to additional NSF trainings as well. For example, one noted, "(paraphrased) There are other opportunities that PO's can do as well. Different groups will have some internal things, whether it's mentoring or formal training. Over the last 20 years, we have done a much better job of preparing POs for how to do the gold standard merit review. The benefit to the trainings is that you are mixing folks from all over the foundation, so they can share best practices, and when people try new things and innovate in some way, they are able to share those advances."
- Staff can seek guidance from their supervisors and agency leadership (4 responses). Leaders noted that staff come to leadership and supervisors with questions. One commented, "(paraphrased) There are also supervisors who may monitor whether people are attending that training, as well as provide additional context themselves. Or junior/newer staff may have more senior or experienced staff they can talk to for support and help in understanding and getting necessary expertise to attend merit review."
- Staff can seek guidance from their peers (3 responses). Leaders noted that POs and DDs with questions can turn to their more experienced peers. One leader said, "(paraphrased) If they've got questions, they start with colleagues."

What instructions or training do you provide to reviewers to help them interpret and apply the merit review criteria? Themes from interviewed leaders' responses include the following:

- Reviewers can participate in NSF training (6 responses). Leaders noted that reviewers receive a merit review training, and some noted that POs will hold additional sessions to train reviewers on directorate-specific considerations. One leader said, "(paraphrased) For [my directorate], POs will do a pre-panel training with the panelists: here's how it will work, here's what we expect of you, here's a bit about bias and how to balance biases. The mail reviewers also have online training; there is some instruction on how best to provide a useful review." Another added, "Reviewers get a variety of trainings—some in writing, videos, meetings, one-on-one meetings with panel."
- Reviewers can seek guidance from POs (4 responses). Leaders stated that POs serve as a resource for reviewers who have questions and that POs also cover merit review criteria at the beginning of panels. One leader said, "(paraphrased) Reviewers are not assigned mentors; they should consult POs if they have any questions."

NSF strives to conduct a fair, competitive, transparent Merit Review process for the selection of projects. To accomplish this, NSF relies on a process that considers both the technical aspects of a proposed project and its potential to contribute more broadly to advancing NSF's mission using the statutory Intellectual Merit and Broader Impacts Merit Review Criteria. MRX invites suggestions on the implementation of the Merit Review Criteria. We especially invite feedback that would

(a) *clarify how the Merit Review Criteria can be used in preparing and reviewing proposals*, (b) ensure proposals, reviews, and funding decisions demonstrate full consideration of both criteria while maintaining openness to the full spectrum of potential activities under each, and (c) better recognize and support potentially transformative and high-risk/high-reward activities. This RFI question received 94 responses. Themes related to how NSF can clarify use of the merit review criteria in preparing and reviewing proposals are described below. The majority of the responses related to areas of improvement for Intellectual Merit and Broader Impacts were from people who indicated they had experience as either a reviewer, or a proposer, or both.

- NSF could provide better descriptions of the criteria (8 responses for Intellectual Merit; 15 responses for Broader Impacts). The most common response about improving the criteria related to providing better descriptions of them. More specifically, respondents suggested NSF provide a clear rubric and definition to understand the criteria. One respondent explained that a first step may be to identify metrics: "You cannot establish criteria without defining a set of metrics by which these criteria can be applied. These metrics at present are missing." Another noted that "developing detailed rubrics that provide clear definitions and examples could significantly enhance the review process of each level of performance relative to the Intellectual Merit and Broader Impacts criteria." Other respondents suggested NSF revise the review process to elicit more explicit responses from reviewers. As one respondent noted, this could involve adding specific questions for reviewers such as "Please describe weaknesses and strengths of the proposal with respect to its potential to benefit society or societal outcomes." Another respondent recommended modifying or removing the question related to assessing proposed Broader Impacts activities, because assessment might not be entirely relevant or possible for some activities during the project period.
- NSF could revise the Broader Impacts criterion (8 responses). Respondents shared ideas or opinions related to modifying the Broader Impacts criteria or review process. Several respondents advocated for the Broader Impacts criterion to be amended to consider whether proposals demonstrated engagement with community partners to support work that would further Broader Impacts. For example, one respondent noted that these community partners could be engaged and work alongside the research team to "evaluate what, if any, harms could occur" as a result of the proposed project.
- NSF could require reviewers to rate Broader Impacts separately from Intellectual Merit (8 responses). Several respondents suggested Broader Impacts should be rated separately from Intellectual Merit; their justifications ranged from a desire for Intellectual Merit to be established first, or that different reviewers should be assigned to assess the Intellectual Merit and Broader Impacts aspects of the proposals. One respondent pointed out that by rating Intellectual Merit and Broader Impacts separately, NSF could "provide greater guidance to the applicants on which section needed greater focus should they decide (or be able to) re-submit an application." However, another countered that "the two criteria should not be 'scored' separately as this would inappropriately separate them."

4. How do PIs, reviewers, and NSF leaders and staff understand and interpret the merit review principles and elements?

Key findings

4.1. The vast majority of staff survey respondents said that NSF staff have a high level of understanding of the merit review principles and elements. The share of respondents who said that PIs and reviewers have a high level of understanding was much lower.

4.2. Interviewed leaders said that POs use reviewer comments on the merit review principles and elements are in their funding decisions. The principles and elements are considered among other important priorities— for example, creating a diverse and balanced portfolio and solicitation-specific criteria—in funding recommendations and portfolio management.

4.3. The number of RFI respondents who reported that the principles and elements are clear was greater than those who reported that the principles and elements are unclear, but some respondents suggested NSF provide more guidance on reviewers should use them. ◢

How would you rate the overall level of understanding of the merit review *principles* by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators? Overall, most staff survey respondents reported they have a high level of understanding of the merit review principles (95 percent) and that POs (91 percent) and DDs, DDDs, and section heads (86 percent) have a high level of understanding (Exhibit 4.1). Staff were less likely to report that reviewers (55 percent) and PIs (46 percent) have a high level of understanding of the merit review elements. DDs were least likely to report that reviewers (14 percent) and POs (14 percent) have a high level of understanding of the merit review principles.

How would you rate the overall level of understanding of the merit review *elements* by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators? Similar to Exhibit 4.1, most staff reported they have a high level of understanding of the merit review elements (96 percent) and that POs (93 percent) and DDs, DDDs, and Section Heads (86 percent) have a high level of understanding (Exhibit 4.2). Staff were less likely to report that reviewers (65 percent) and PIs (51 percent) have a high level of understanding of the merit review elements.

Interviewed leaders were asked a series of questions about the merit review principles and elements. In many cases, leaders discussed the principles and elements interchangeably with the criteria. The themes we describe below thus draw on responses both about the principles and elements, as well as the criteria themselves.

How are the merit review principles and elements used to identify which projects to fund? Most (6 interviewed leaders) cited that the principles and elements are incorporated into reviews and recommendations, and one person said that this is done in the form of a checklist:

- Principles and elements are incorporated into reviews and recommendations (6 responses). One leader noted the challenges of this, saying "(paraphrased) I would say when those principles and elements were instituted, it started getting more complicated…. I think reviewers stumbled on these. They felt like

all of a sudden, they were filling out a form…. My recollection is these show up when they fill out their reviews."

- Proposals are most successful if they tell a compelling narrative (1 response). On the other hand, one leader said that the biggest driver of a proposal's success is a compelling narrative. They said, "(paraphrased) The proposal guide language evolves every year, and people look carefully to make sure they're addressing those items. So people read them at that level, but if you were to poll a group of panel reviewers and asked if they systematically read that list and checked each box etc., they'd say probably not. I tell people writing proposals that in review, people look for an interesting story. Honestly think that's how most people read proposals. If there's a surplus of good proposals, some of these secondary criteria come into play, but often it's about—is this someone I want to talk to, who would excite me about what they're doing. That's the high-level litmus test that usually drives reviews."

How are the principles and elements used by Program Officers to make funding recommendations in your directorate? Interviewed leaders often cited reviewer comments and assessing the directorate portfolio. These themes are expanded upon below:

- POs use reviewer comments to make funding recommendations (4 responses). Almost half of leaders talked about reviewer comments as a core part of how POs use the principles and elements to make funding decisions. For example, one leader said, "(paraphrased) [Reviewers] each write their own review, and then there will be a panel review, and it's in that group conversation that most of the decision making occurs… As a practical matter, given large number of proposals submitted, and being a lean institution, [there is] not a lot of time for program officers to engage in that type of thoughtful scrutiny after panel reviews because of the burden of writing out that rationale."
- POs make funding recommendations by assessing the needs of their portfolio (3 responses). Some leaders also mentioned scoping and assessing the program's portfolio and goals as a key aspect POs use to make funding recommendations. For example, one leader said, "(paraphrased) Like I said, these actually inform the way we scope our programs themselves. This is what we are trying to achieve, but depending on what the program is, some of these are more relevant and important than others, so to me these are kind of principles that underlie what NSF and [directorate] are trying to accomplish, so always keeping them in mind when evaluating which proposals to fund is the important thing."

How are the principles and elements used by Division Directors in their portfolio management? Themes captured in interviewed leaders' responses are described below:

- DDs seek to fund a portfolio balanced against multiple goals (3 responses). Several leaders mentioned the importance of seeking a balanced portfolio and named multiple aspects of balance, including against the merit review criteria, NSF's broader goals, and the directorate's or office's more specific goals.
- DDs seek to develop a diverse portfolio (2 responses). Two leaders talked about the importance of the DD in building a diverse portfolio. Where reviewers and panelists evaluate a specific proposal against the merit review criteria, the role of the DD is to keep an eye on the bigger picture. For example, one leader said, "(paraphrased) Program officers and division leaders may have overarching understanding

of field and NSF and based on assessments…. [They] can think about it in context of a bigger effort and fit those pieces together so there's a diverse and optimized portfolio."

In your experience, are these uses similar to or different from how the principles and elements are used in other directorates and offices? All five of the interviewed leaders who responded to the question said that the uses are largely similar, but four noted that the application of the principles and elements are sometimes supplemented to meet the needs of specific programs or directorates. This can occur in the form of solicitation-specific criteria or additional considerations like short-term Broader Impacts or portfolio management. For example, one leader said, "(paraphrased) I would say we take [them] seriously and consider both in every review. It's addressed by every reviewer and discussed in every panel. But to me, it's a starting point for a review of a proposal. We add additional criteria as necessary."

How confident are you that the principles and elements are used consistently within your [directorate/office]? Can you tell me about where you think there are areas of consistency and inconsistency? Most of the interviewed leaders who responded said that they were very or moderately confident that the principles and elements are used consistently within their directorate (5 interviewed leaders), and two of these individuals cited Committee of Visitors (COV) reports as supporting evidence. In contrast, one leader said that they were not confident that the principles and elements are applied consistently within their directorate. When asked about areas of consistency and inconsistency, the following theme emerged:

- Inconsistencies might come from process differences (2 responses). Two leaders mentioned that process differences could be the cause of inconsistencies in how the principles and elements are applied. These might include programs that use individual reviewers versus panel reviews and differences in the length and specificity of reviews. Similarly, the flexibility given to POs could introduce some process differences. For example, one respondent said, "(paraphrased) I think principles and elements are consistently applied, but the inconsistencies are a little bit more on the process side, are people writing review analyses, do we need a review? Probably not, you can address principles and elements in less space, but since we are having that discussion, it allows program staff to talk about what they are really applying, whether it's principles and elements or how they do decision making."

The MRX is interested in identifying opportunities to improve NSF's current Merit Review criteria, policy, and processes. Importantly, this includes documenting and understanding any areas of misunderstanding, gaps, or lack of clarity regarding (a) *the three Merit Review Principles which are the foundations of the Merit Review Process*, (b) the two statutory Merit Review Criteria which are used to evaluate all proposals to NSF, and (c) *the five Merit Review Elements NSF uses to assess each criterion*; for instance: Are the *Principles*, Criteria, and *Elements* clear? Could they be improved upon? We received 114 responses to this RFI question, but few spoke specifically to the principles or elements. Of these responses, 11 indicated they had submitted a proposal to NSF, 2 indicated they had served as a reviewer, and 21 had experience both as a proposer and reviewer. Themes in these responses related to the merit review principles and elements include the following:

- Principles are clear (15 responses). Some respondents indicated that the principles are well defined and clear. Within this group, three respondents noted that although the principles themselves were clear, they were not always interpreted uniformly. For example, one respondent said, "While Principle 3

('Meaningful assessment and evaluation of NSF funded projects...') is sound, it is focused on funders rather than individual proposers.... To aid proposers, [organization] recommends NSF provide guidance on the ways in which a proposer should articulate how their project contributes to broader initiatives and addresses specific knowledge gaps."

- Elements are clear (17 responses). Similarly, respondents indicated that the descriptions of the five elements are clear, but that there is some variation in how the elements are interpreted and applied. For example, one respondent wrote, "[The elements] are clear but are not interpreted uniformly in some solicitations." Another respondent noted the challenge of applying the elements to Broader Impacts: "Element 2 ('To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?') is problematic in its application to the Broader Impacts criterion. There is no need for a BI approach to be novel or creative to be impactful. There is no need for the exact same questions to be applied to each criterion. Let them be different."
- Principles (5 responses) and elements (4 responses) are subjective or unclear and require better metrics to evaluate. Some respondents reported that phrases used in the principles, such as "highest quality" and "societal goals," are too subjective to evaluate accurately. They suggested that NSF simplify the principles and provide examples that proposers and reviewers can refer to. As one respondent explained, "The discussion about how a project should contribute to societal goals is problematic, because what are those goals, who sets them, how will a particular research outcome impact society, and how would anyone prove it?... This guidance would benefit from specific guardrails; for example, it could point to [Office of Science and Technology Policy] initiatives or other clear reference points that help the PI center their work around national priorities." Similarly, a few respondents suggested NSF clarify how the elements are to be assessed. One respondent noted that without such guidance, reviewers may choose their own approach: "In some cases the metrics that should be used to make an assessment are not well defined and in some instances reviewers are using questionable measures. For example, to the question whether the PI is well qualified to execute the work, instead of being informed by the CV, reviewers are using social media to receive their information, falling victim many times to faulty information. Other times they are using questionable metrics such as citations.... NSF needs to define the metrics which the reviewers should or should not use."

5. What additional merit review criteria might NSF consider using in the merit review process to better achieve its mission to invest in research that advances scientific knowledge and benefits society?

Key findings

5.1. About half of staff survey respondents noted additional factors that are important for evaluating proposals that the merit review criteria do not capture. These respondents most commonly cited solicitation-specific criteria, the applicant's background and institutional characteristics that could diversify the characteristics of awardees, more specific guidelines and priorities for Broader Impacts, and the relationships between Intellectual Merit and Broader Impacts as additional factors to consider.

5.2. Interviewed leaders cited solicitation-specific criteria, institutional diversity, and the relationships between Intellectual Merit and Broader Impacts as additional factors to consider. ◢

Based on your experience, are there any factors important for evaluating proposals that are not captured by the two merit review criteria? Almost half of staff survey respondents (49 percent) responded yes to this question (Exhibit 5.1).

- DDs were most likely to report that there are important factors that are not captured (100 percent), compared to 59 percent of DDDs, 46 percent of POs, 40 percent of DADs and DOHs, and 64 percent of other NSF staff ($p < 0.05$).

Please describe the factor(s) and why you think it would be important for reviewing proposals. What could this capture that is currently missed by the merit review criteria? A total of 108 staff survey respondents received the question because they answered yes to the previous question. Themes of these responses include the following:

- Include solicitation-specific criteria (23 responses). Some staff noted that solicitation-specific review criteria are important for reviewing proposals, as they can capture program-specific aspects that are important for evaluation beyond the standard criteria. One noted, "This is a program-specific issue. I am surprised at how few solicitations include solicitation-specific review criteria. This is the correct mechanism to address this need."
- Consider other factors with the potential to diversify characteristics of awardees (23 responses). Some staff emphasized that they already consider factors related to applicants' background, including demographic, institutional, and geographic factors, with the intention to broaden participation, but that it would be valuable to make this a separate criterion from Broader Impacts. A staff member noted that, "PIs/institutions don't have access to the same resources. R1 institutions often have more resources to assist with their proposals. Other types of institutions don't have that, but they are judged as if they do." Another staff member added, "POs need to have better training and recognize when the same players are getting money from their programs to the detriment of new PIs, underrepresented PIs, and PIs with limited institutional support." Another staff member noted that not all scientists communicate best in writing and suggested "a merit review system that allows for discussion," because "some of our best science came from groups of scientists who discussed with each other and helped to elucidate the path forward through these discussions."

- Create more specific guidelines and consider additional priorities for Broader Impacts (18 responses). Some staff suggested clearer questions for Broader Impacts. One staff respondent said, "It's really that the five questions elicit clearer responses from reviewers related to IM. The five questions come across as less clear and less relevant with respect to BI." Some suggested more explicit criteria to measure reach of projects beyond their funding period. Other staff suggested stronger consideration as to whether the proposed scientific research is environmentally sustainable and potential risks of carrying out the research or risks of the emergent knowledge, as well as requiring applicants to describe how the project will impact the whole society, not just the scientific community.
- Consider the relationship between Intellectual Merit and Broader Impacts (8 responses). Some staff asked for greater distinctions between the two merit review criteria, while others asked for greater emphasis on how to connect the two criteria. One staff respondent noted, "Please make greater and more meaningful distinctions between the two review criteria. The sub-questions under each criteria are the same. Why do we do this? The criteria are different and deserve different emphases. This would greatly help reviewers in making judgments about IM and BI." One staff respondent noted, "Stronger emphasis on how to connect IM and BI in such a way that both criteria can be viewed equally would be important." Another suggested, "A specific discussion linking IM and BI."

In your opinion, are there any other important factors for evaluating proposals that are not captured by the merit review criteria? Tell me a bit more about this factor and why you think it would be important for reviewing proposals. What could this capture that is missed by the current merit review criteria? All nine interviewed leaders answered this question. Four indicated that they believe there *are* important factors not captured by the merit review criteria, three indicated that they do *not* believe there are additional factors, and two provided an ambiguous response. Themes from their responses regarding other important factors include the following:

- Include solicitation-specific criteria (6 responses). Some interviewed leaders noted that other factors important for evaluating proposals are program specific. One leader commented, "(paraphrased) I think this is a program-by-program issue; we add additional criteria to specific programs as needed. I don't think there's one criteria missing for all of NSF."
- Consider the relationship between Intellectual Merit and Broader Impacts (2 responses). Two leaders noted the interdependencies between the two criteria, and in their responses, they weighed the possibility of more clearly defining each criterion or combining them. One leader stated, "(paraphrased) There are benefits and challenges associated with the way they're phrased in this broad way. One of the things that comes up occasionally is broadening participation in education and workforce activities. [I] think those are activities that would certainly be a part of Broader Impact and could also be Intellectual Merit if the explicit purpose of the project is to advance knowledge on how do we broaden participation, why is our participation the way it is, or what are better ways we can deliver education and workforce development. That's something that could be covered by both Intellectual Merit and Broader Impacts already." Another respondent noted, "(paraphrased) I think Broader Impacts is within Intellectual Merit, and Intellectual Merit is within Broader Impacts. And that is to say they are both; if we are going to keep it, it has to be in both; with that being said, if we just had Intellectual Merit and Broader Impacts as one criterion, and then you have sub under that; maybe knowledge mobilization is

one, maybe it's methodology, something about the personnel/experiences, ability to execute the project, these kind of things that I think would enhance it."

- Consider a factor that explicitly addresses institutional diversity (2 responses). One leader noted that community colleges and small universities have an innately more difficult time competing with large universities with more resources and suggested a criterion that examines whether proposals from smaller institutions will be high impact in the space they are in. Another noted, "(paraphrased) We need a criteria that does not blur where a person comes from. I have seen as a PO where an investigator comes from an institution that may not be highly esteemed, but the proposal was stellar, and the investigators say the individuals won't be able to execute what they planned out."

6. How is reviewer and staff expertise in the merit review criteria assessed?

Key findings

6.1. Almost all staff survey respondents said that they or their unit are somewhat or very effective at identifying which reviewers have the necessary expertise in the merit review criteria. However, when asked about the overall level of understanding of the merit review criteria among different groups, far fewer respondents said that reviewers and PIs have a high level of understanding of the merit review criteria, compared to the share of respondents who said that NSF staff have a high level of understanding.

6.2. Interviewed leaders commented that POs are responsible for determining which reviewers have the necessary expertise in the merit review criteria to evaluate proposals. They reported that POs make these decisions in consultation with division leaders, by examining reviewers' background and experience, and by relying on their own knowledge of the field. ◢

How effective do you think [you are/your unit is] at identifying which reviewers have the necessary expertise of the merit review criteria? Overall, 98 percent of staff survey respondents indicated that they are somewhat or very effective at identifying which reviewers have the necessary expertise (Exhibit 6.1).

How would you rate the overall level of understanding of the merit review criteria by you, other NSF staff, reviewers, and principal investigators? Staff survey respondents rated themselves highly on this measure (Exhibit 6.2). They rated their fellow NSF staff highly as well, with reviewers and PIs receiving lower ratings.

- Asked about themselves ("you"), 98 percent indicated they have a high level of understanding and 2 percent indicated they have a moderate level of understanding of the criteria. There were differences in response patterns by position type. A substantial proportion of POs (99 percent), DDDs (100 percent), DADs and DOHs (100 percent), and other NSF staff (90 percent) indicated they have a high level of understanding ($p < 0.05$). No staff indicated they have a low level of understanding.
- Staff largely indicated that POs have a high level of understanding (93 percent) of the criteria, with 6 percent indicating POs have a moderate level of understanding, and less than 1 percent rating POs level of understanding as low.
- Most staff also reported that DDs, DDDs, and section heads have a high level of understanding (88 percent), with 9 percent indicating this group have a moderate level of understanding, and 3 percent rating their level of understanding as low. There were differences in response patterns by

directorate or office. A greater proportion of respondents from BIO (94 percent), CISE (100 percent), ENG (96 percent), MPS (92 percent), and TIP (100 percent) reported that these leaders have a high level of understanding of the merit review criteria. Fewer respondents in GEO (83 percent), EDU (60 percent), OD/OIA (60 percent), and respondents who did not report a directorate (85 percent) reported that leaders have a high level of understanding of the merit review criteria.

- Sixty-three percent of staff respondents indicated that reviewers have a high level of understanding of the criteria, 33 percent rated reviewers' level of understanding as moderate, and 3 percent rated reviewers' level of understanding as low.
- About half of staff respondents (49 percent) indicated that PIs have a high level of understanding, 45 percent indicated they have a moderate level of understanding, and 6 percent indicated they have a low level of understanding.

Who determines which *NSF staff* have the necessary expertise of the merit review criteria to evaluate proposals? How do they make that determination? Four of nine interviewed leaders answered this question. Themes from their responses include the following:

- NSF staff do not evaluate proposals and are not directly hired for their expertise in the merit review criteria (2 responses). One interviewed leader noted that the evaluation is done by the review panel, not NSF staff, and another leader noted that they are not looking for staff with expertise in merit review criteria; staff are hired for other expertise and build expertise in the criteria through trainings. This leader said, "(paraphrased) We bring in people to be POs who have relevant expertise in the topic area and investments we are trying to make. We look for people who have a holistic perspective and a range of experience, then we send them to merit review criteria courses, the Basics. DDs and others do that training, so that to me gives them expertise in the merit review criteria. So, they are hired for their other expertise, their topical expertise in the criteria, but then they are trained in the criteria."
- Other staff make the determination (2 responses). Leaders identified supervisors and leadership (1 respondent) as well as hiring managers (1 respondent) as responsible for assessing staff expertise. As one leader explained, "(paraphrased) That's in the hiring process really. So the mentoring of new people, it's the group mentoring that's done, the nice thing is the way our electronic system works; it's almost like a wizard, it walks through it on the technical side. As far as the intellectual expertise, it's in the hiring."

Who determines which *reviewers* have the necessary expertise of the merit review criteria to evaluate proposals? All interviewed leaders noted that POs make this determination. When asked how they make that determination, respondents described how POs consult with division leaders (3 responses), examine reviewers' background and experience (2 responses), and rely on their [the PO's] understanding of the field (2 responses) to select reviewers. For example, one leader explained, "(paraphrased) In terms of figuring out who has expertise, this is one of the key elements of the job. Knowledge of what's going on in community, having familiarity with the literature, attending conferences, as well as asking folks in community to volunteer, or might involve prospective reviewers themselves describing expertise. These are some of the ingredients that go into matching technical expertise of a reviewer with technical expertise of a proposer." One respondent also noted that they "(paraphrased) discourage POs from just picking people they know," because they "really want to diversify the review community."

7. How does reviewer and staff expertise on the merit review criteria factor into reviews, recommendations, and decisions?

Key finding

7.1. Interviewed leaders described a layered process for how reviewer and staff expertise are evaluated to ensure that they are reviewing proposals and making recommendations that are aligned with the merit review criteria. Reviewers are expected to document the criteria in their reviews, which POs and DDs oversee. POs are expected to document their funding recommendations using the criteria, and these are overseen by division leaders and confirmed via the "DD concur" function. ◢

How do you evaluate whether *Program Officers* are reviewing proposals and making recommendations that are aligned to the merit review criteria? Seven of nine interviewed leaders answered this question. Themes from their responses include the following:

- The "DD concur" function provides oversight (4 responses). Interviewed leaders noted that DDs concur or do not concur with the recommendation of a PO, and if they do not concur on something, it gets pushed back to a PO. One leader noted, "(paraphrased) How do we evaluate POs? That's the DD concur step, where the DD reads all of the supporting materials related to the decision and is able to evaluate whether the PO has taken the input they have received and made an appropriate decision, or if they have justified their decision."
- Conducting other oversight of POs (3 responses). Interviewed leaders noted that division leaders conduct oversight of POs to ensure they are making decisions based on the merit review criteria: "(paraphrased) Program officers are responsible for recommending which projects to fund, which should be framed in the merit review criteria. Often there may be a leadership meeting involving program officers and division leaders discussing strengths and weaknesses in proposals, and projects they're preparing to fund, which will be framed in the merit review process. Program officers' bosses will hear their recommendations and see whether they are making those recommendations based on the merit review process."
- PO documentation of recommendations (2 responses). Respondents noted that DDs examine POs' reviews and look for justification of the scores they give and incorporation of the merit review criteria into the reviews. One stated, "(paraphrased) There's some documentation expected of POs and reviewers that's captured in our systems. That's also framed in the merit review process. A division leader will review that documentation in that process, and that's how a division director or a deputy division director will see if a reviewer or PO is talking about the merit review criteria."

How do you evaluate whether reviewers are *reviewing* proposals and making recommendations that are aligned to the merit review criteria? Three of nine interviewed leaders answered this question. Themes from their responses include the following:

- Evaluation of reviewer comments (2 responses). Two leaders noted that they evaluate reviewer summaries or reviews to determine whether they are using the merit review criteria. One noted that, "(paraphrased) With a mail reviewer, you get what you get, you won't send it back. If it isn't helpful, then you generally don't use much of what's there. Then you don't usually use that reviewer again. Good reviewers do want to look at online resources to help guide them."

- Oversight by POs and division leaders (2 responses). Leaders noted that oversight by POs and division leaders is used to evaluate whether reviewers are making recommendations aligned with the merit review criteria. One respondent noted, "(paraphrased) It depends whether it's panel or mail review. They are very different. On a panel, the PO has more ability to look over the shoulder of panelists; this is where the word reviewer is confusing; in my head a reviewer is ad hoc, for one proposal. A panelist is where they get a group of proposals, of which they may review them as well. On the panelist side, a PO has much more ability to look over their shoulder and kind of say 'you are repeating a lot of what's in the proposal, we have it, we don't need that, can you focus more on what you think is intellectually strong or Broader Impacts?'"
- Evaluation through COV reports (1 response). One leader pointed to the COV as a method of evaluating reviewers' recommendations. "(paraphrased) There's this committee that takes place every few years. This is another check about whether our reviewers are talking about merit review criteria."

How do you evaluate whether *Division Directors* are making decisions about whether to award or decline proposals in a way that is aligned to the merit review criteria? Five of nine interviewed leaders answered this question, with four stating that this type of evaluation does not happen. These respondents noted that DDs do not evaluate proposals in this way, and that a DD's role is not to make the decisions, but to concur or not concur. One leader noted, "(paraphrased) They are not making decisions, they are determining whether decisions are sound." Another added, "As executives, they're empowered to sign off on those decisions themselves."

8. How do reviewers balance the Intellectual Merit and Broader Impacts criteria in assessing proposals?

Key finding

8.1. RFI respondents reported that Broader Impacts is not assessed comparably to Intellectual Merit, citing the lack of training on Broader Impacts and the lack of expertise among reviewers in aspects of Broader Impacts. ◢

MRX is interested in the experiences and perspectives of those who have reviewed proposals submitted to NSF. We invite you to share your insights and describe any opportunities you believe would improve implementation of the Merit Review criteria, policy, or processes based on your experience reviewing NSF proposals. We received 71 responses to this RFI question; however, most did not address how reviewers balance the Intellectual Merit and Broader Impacts criteria, and those who did primarily focused on how Broader Impacts is assessed. Themes in the responses related to how reviewers balance the two merit review criteria include the following:

- Broader Impacts is not comparably assessed as Intellectual Merit (7 responses). Several respondents said that reviewers are more likely to have expertise in Intellectual Merit than Broader Impacts. Four of these respondents indicated they had served as ad hoc reviewers and two as panelists; three of these respondents were reviewers in the past five years. To evaluate Broader Impacts as rigorously as Intellectual Merit, NSF should train reviewers in Broader Impacts and recruit reviewers with expertise in it. One respondent summed up this sentiment: "One of the key barriers to implementation of the BI criteria, as well as recognition by their own institutions of the BI work that researchers do … is the difficulty of reviewers understanding how to evaluate the quality of proposed BI activities." In addition

to including more training and BI experts as reviewers, one respondent suggested that proposers' past impact on community partners should be assessed as deeply as their publication record. Another respondent raised a concern that "panelists target the BI of a proposal for undue criticism if they do not like the IM."

9. How do reviewers' assessments of each merit review criterion factor into NSF program officers' award recommendations?

Key finding

9.1. Nearly all PO staff survey respondents reported that reviewers' assessment of the Intellectual Merit and Broader Impacts criteria are somewhat or to a great extent factored into POs' funding recommendations. ◢

How much do reviewers' assessments of each merit review criterion factor into funding recommendations within your division? Staff survey respondents identifying themselves as POs were asked this question (Exhibit 9.1). Nearly all of the responding POs reported that reviewers' assessment of the Intellectual Merit (99 percent) and Broader Impacts (95 percent) criteria are somewhat or to a great extent a factor in funding recommendations.

10. How do reviewers' assessments of each merit review criterion factor into NSF division directors' award decisions for individual proposals? At the portfolio level?

Key findings

10.1. Most DD, DDD, DAD, and DOH staff survey respondents reported that reviewers' assessments of each merit review criterion are factored somewhat or to a great extent into portfolio management.

10.2. Interviewed leaders specified that POs and directorate leadership generally review funding decisions to understand how proposals are rated and why they are recommended (or not) for funding based on reviewer comments. ◢

How much do reviewers' assessments of each merit review criterion factor into portfolio management within your division? DD, DDD, DAD, and DOH staff survey respondents were asked this question (Exhibit 10.1). For both Intellectual Merit and Broader Impacts, 93 percent of responding DDs, DDDs, DADs, and DOHs reported that reviewers' assessments are somewhat or to a great extent a factor in portfolio management.

How are reviewers' assessments of each merit review criteria factored into portfolio management within your [directorate/office]? Six of nine interviewed leaders answered this question. Most indicated that POs and directorate leadership review funding decisions (5 responses). As one explained, "(paraphrased) That's done at the section level. It's often presented to me and there's a question of how did these proposals review, and why are we funding them? We have a practice of presenting on these and asking about funding decisions."

11. To what extent does use of reviewers' assessments of each merit review criterion vary within and across NSF directorates?

Key finding

11.1. There was no statistically significant variation across directorates in the proportion of PO staff survey respondents who indicated that reviewers' assessments of each merit review criterion factor at least somewhat into award recommendations. There was also no variation across directorates in the proportion of DD, DDD, DAD, and DOH respondents who indicated that reviewers' assessments of each merit review criterion factor at least somewhat into portfolio management. ◢

To answer this guiding question, we looked for variation by directorate in staff survey responses reported under other guiding questions. These items related to how reviewers' assessments of each merit review criterion factor into POs' award recommendations (Exhibit 9.1) and DDs', DDDs', DADs', and DOHs' portfolio management (Exhibit 10.1). There is little variation by directorate, and none of the differences between directorates were statistically significant. At least 97 percent of responding POs in each directorate indicated that, for their division, reviewer assessments of the Intellectual Merit criterion factor into funding recommendations to some or a great extent; at least 82 percent indicated this for Broader Impacts. Among DD, DDD, DAD, and DOH respondents, between 50 percent (SBE) and 100 percent (BIO, CISE, ENG, GEO, and EDU) indicated that reviewer assessments of the Intellectual Merit criterion factor into portfolio management to some or a great extent; this ranged from 50 percent (SBE) and 100 percent (BIO, CISE, ENG, GEO, and EDU) for Broader Impacts.

12. To what extent do NSF constituencies perceive the merit review policy and processes to be unfairly biased? How do these perceptions differ by constituency?

Key findings

12.1. Most staff survey respondents agreed that submitted proposals are evaluated fairly and the merit review criteria support a fair and accurate assessment of a proposal's merit; however, White and Asian staff were more likely than Black or African American staff to agree that individuals submitting proposals are treated fairly.

12.2. The majority of interviewed leaders agreed that the merit review policy supports a fair and accurate assessment of a proposal's merit, citing NSF policies such as rules on managing conflicts of interest, policies on diversity in reviewers and on projects, and the use of external reviewers. A few leaders disagreed, however, citing implicit bias as a potential reason that a proposal might not receive a fair and accurate assessment. ◢

Based on your experience with the merit review process, to what extent do you agree or disagree with the following statements? Individuals submitting proposals are treated fairly. The merit review criteria support a fair and accurate assessment of a proposal's merit. Among all staff survey respondents, 87 percent agreed or strongly agreed that individuals submitting proposals are treated fairly, and 85 percent agreed or strongly agreed that the merit review criteria support a fair and accurate assessment of a proposal's merit (Exhibit 12.1 and Exhibit 12.2).

- DDDs (100 percent) and DADs and OHs (100 percent) were more likely than POs (87 percent), DDs (67 percent), and other NSF staff (60 percent) to agree that individuals submitting proposals are treated fairly.
- Staff in CISE and BIO were most likely to agree that individuals submitting proposals are treated fairly (100 percent and 97 percent, respectively; respondents from TIP and OD/OIA were least likely to agree (50 percent and 20 percent, respectively). This pattern also held for agreement that the merit review criteria support a fair and accurate assessment of a proposal's merit (100 percent for CISE, 94 percent for BIO, 40 percent for OD/OIA, and 38 percent for TIP).
- White staff (90 percent) and Asian staff (88 percent) were more likely than Black or African American staff (60 percent) to agree or strongly agree that individuals submitting proposals are treated fairly.

How much does the merit review process increase or decrease the diversity of ideas, principal investigators, and institutions represented in *proposals submitted for review*? Staff survey respondents were most likely to report that the merit review process increases the diversity of ideas (41 percent), neither decreases nor increases the diversity of PIs (40 percent), and decreases the diversity of institutions (41 percent) in proposals submitted for review (Exhibit 12.3).

How much does the merit review process increase or decrease the diversity of ideas, principal investigators, and institutions represented in awarded projects? Staff survey respondents were most likely to report that the merit review process increases the diversity of ideas (52 percent), increases the diversity of PIs (49 percent), and increases the diversity of institutions (45 percent) in awarded projects (Exhibit 12.4).

- Staff from BIO were most likely to report that the merit review process increases the diversity of PIs (79 percent), while staff from TIP (25 percent) and EDU (28 percent) were least likely to report that the merit review process increases the diversity of PIs. Similarly, 71 percent of BIO staff indicated the merit review process increases the diversity of institutions compared to 25 percent for TIP and 20 percent for EDU and OD/OIA, respectively.

Do you believe that the merit review policies support a fair and accurate assessment of a proposal's merit? Why or why not? Seven interviewed leaders agreed and three disagreed that the merit review policies support a fair and accurate assessment of a proposal's merit (leaders could both agree and disagree) (Exhibit 12.5 and Exhibit 12.6). Those who agreed named several aspects of NSF policies related to merit review (3 responses), including rules on managing conflicts of interest, policies on diversity in reviewers and on projects, and the use of external reviewers, that help to limit the potential for the process to be biased. Among those who disagreed, the reasons they gave included that implicit biases are part of human nature (2 responses).

What, if anything, does your [directorate/office] implement to mitigate potential biases? Interviewed leaders named multiple actions taken to mitigate bias. Themes include the following:

- Raising awareness of and screening for implicit biases among NSF staff and reviewers (4 responses). Some leaders described how their directorate or office seeks to increase attention to the potential for implicit biases to shape reviews through training or guidance to POs. One leader explained,

"(paraphrased) The other piece is making people aware of this question of whether you're bringing bias to the table in the review. That's the job of a program officer and other members of panel to say, wait a minute, that's an assumption. Where's it grounded in fact, so we have a basis in judgment?"

- Leaders described other actions including seeking a group of reviewers with diverse perspectives and backgrounds (3 responses); being explicit about the merit review criteria, including any solicitation-specific criteria (3 responses); conducting outreach to institutions that have less experience with NSF awards (3 responses); and using programs intended to build the capacity of institutions and groups underrepresented in NSF funding (3 responses).

Do you think the merit review process increases or decreases the diversity of *proposals submitted for review*? How so? Interviewed leaders' responses were mixed:

- The merit review process has no clear effect on diversity in proposals (5 responses). Some leaders did not see a connection between the merit review process and the proposals that are submitted. One noted that, if there is a lack of diversity in proposals, it is more likely due to factors external to NSF. "(paraphrased) The merit review criteria to me does not and should not shift. If there's a lack of diversity, it should be the things around merit review that we need to build up.... what we need to focus on is how to give the opportunity, training, muscle building, motivation. That's what we should focus on in NSF. How do we make sure they're far more capable to transcend the bar?"
- The merit review process decreases diversity in proposals (3 responses). Some leaders noted that submission requirements and questions about the fairness of the process can lead PIs to decide not to apply for funding. One leader explained, "(paraphrased) I will say there are folks out there that believe the process is subjective, and if they believe that, they are less inclined to submit proposals to NSF, and those are folks from underrepresented groups and populations who maybe don't have the resources at their institutions to facilitate the development of proposals or to help translate what we mean by merit review criteria. These are exactly the types of communities we are trying to engage with in more under-resourced institutions, so that could inform whether they choose to submit or not."
- The merit review process increases diversity in proposals (2 responses). Two leaders asserted that the process could increase the diversity of proposals. One leader attributed this to outreach that their directorate or office conducted to provide underrepresented institutions with the information. "(paraphrased) This is an anecdote: people said out loud that they felt that our directorate was more responsive and attentive to [resource differences between institutions].... It could be that they have more comfort and that the playing field is more even [in my directorate] because the institutions are the same."

Do you think the merit review process increases or decreases the diversity of *awarded projects*? How so? As with the diversity in proposals, interviewed leaders' responses were mixed:

- The merit review process has no clear effect on diversity in projects (5 responses). Five leaders maintained that the merit review process is not related to diversity in awarded projects. As one leader stated, "(paraphrased) Not directly—each individual has to make their case. The proposal meets the criteria, or it doesn't."

- The merit review process decreases diversity in projects (3 responses). Leaders who saw the merit review process as decreasing diversity in awarded projects tended to attribute it to differences in institutional resources. One leader explained, "(paraphrased) If we see projects concentrated at certain institutions, and we ask ourselves if that's how we want the distribution of projects we look like, it could be that it's skewed towards institutions that have resources to comply with our expectations for how to steward taxpayer resources."
- The merit review process increases diversity in projects (3 responses). Leaders identified several aspects of the merit review process that serve to increase the diversity of projects. As one leader stated, these include "(paraphrased) having explicit criteria so people know how they're supposed to rate a proposal on, and what [the] criteria/principles are, making sure people in [the] process are making recommendations couched in those criteria, that we have people participating [in the] process without a conflict of interest, seeking diversity in our reviewers… those are all intended to increase diversity of the projects."

13. What aspects of the merit review policy and processes are perceived to be biased?

Key findings

13.1. The majority of staff survey respondents noted aspects of the merit review *processes* that could introduce unfair bias, citing PI name recognition or background and institutional research infrastructure and prestige as factors that could bias reviews. Moreover, only about half of staff agreed that NSF staff and reviewers are sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process. Some staff also noted aspects of the merit review *policy* that could introduce unfair bias, explaining that some NSF policies favor proposals from institutions with adequate infrastructure to respond to the proposal requirements and that the lack of anonymizing of proposals can introduce bias from PI or institution name recognition.

13.2. The majority of interviewed leaders named aspects of the merit review processes that could introduce unfair bias, also citing institutional research infrastructure and prestige as factors that could introduce bias. ◢

NSF *staff* and NSF *reviewers* are sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process. Over half of staff survey respondents agreed that NSF staff (59 percent) and NSF reviewers (56 percent) are sufficiently diverse to achieve the goals of the merit review process (Exhibit 13.1).

Are there any aspects of the merit review *process* that you think introduce unfair bias into the evaluation of proposals? If so, how? Among the 163 staff survey responses, 135 indicated yes, there are aspects of the process that introduce unfair bias, and 35 indicated no, there are not aspects of the process that introduce unfair bias (responses could be coded to both yes and no). Among the yes responses, themes include the following:

- Reviews can be less critical based on PI name recognition or background (41 responses). Some staff survey respondents indicated that knowing the identity of the PI can lead some reviewers to give the benefit of the doubt to PIs whose work they are familiar with. Some staff referred to this as a halo effect, and others noted that it can also make it more difficult for early career scientists to be awarded funding. One staff member explained, "In many research communities, everyone knows everyone else, so even if

a reviewer doesn't have a conflict of interest with the PI, they may know of and have interacted with the PI. This often leads to certain PIs getting preferential ratings from reviewers (i.e., 'I am confident the PI can overcome these shortcomings') that new PIs or unknown PIs would not get. By allowing reviewers to know the PIs they are reviewing, there is definitely an opportunity for substantial familiarity bias. Double blind reviews would alleviate this issue."

- Reviews can be biased towards PIs from institutions with adequate research infrastructure (38 responses). Some staff noted that institutional resources can make a difference in proposal quality and project plans. For example, one staff member explained, "Some large universities (R1) have more support for PIs both for proposal preparation, for supporting students and postdocs, and for encouraging DEI [Diversity, Equity, and Inclusion]. Those that don't have that support may have a harder time with proposal sections such as mentoring plans, outreach activities recruiting minority students, designing mechanism to assess success, etc. Non-R1 universities should be given more flexibility in judging these criteria."
- Reviews can be more favorable toward PIs from prestigious institutions (33 responses). Similar to the leniency sometimes afforded to proposals from well-known PIs, staff reported that reviews can be less critical of proposals from institutions with R1 status or other name recognition. As one staff member described, "There is a tendency to "trust" well-known researchers and R1 institutions and question both the Intellectual Merit and Broader Impacts of emerging research institutions and MSIs. That is why it is important to have panels comprised of diverse individuals from diverse institutions."

Are there any aspects of the merit review *policies* that you think introduce unfair bias into the evaluation of proposals? If so, how? We received 132 responses to this staff survey question; 56 indicated yes, there are aspects of the policies that introduce unfair bias, and 65 indicated no, there are not aspects of the policies that introduce unfair bias (responses could be coded to both yes and no). Among the yes responses, themes include the following:

- The policies favor proposals from institutions with adequate research infrastructure (16 responses). Staff indicated that the merit review element related to resources to carry out the project can introduce bias. One staff member said, "For under resourced institutions, the merit review criteria about adequate resources, and the policies that limit purchases of equipment or PI time, certainly introduce unfair bias."
- Lack of anonymizing can introduce bias (15 responses). Staff reported that reviewers' awareness of proposers' identities can sometimes lead them to base their review on their perceptions of the PI or institution, rather than the proposal. As one staff member noted, "PIs with a longer (or more productive) track record seem to be treated with more deference by panelists."

Are there any aspects of the merit review *process* that you think introduce unfair bias into the evaluation of proposals? If so, how? Seven interviewed leaders named aspects of the process that they thought could introduce unfair bias, and two stated that they did not think there were aspects of the process that introduce unfair bias. Themes in the responses about how bias could be introduced include the following:

- Reviews can be swayed by the prestige of the PI's institution (4 responses). Some leaders spoke from their experiences with or observations of review panels and gave examples of proposals that were

viewed more or less favorably based on the institution where the PI was based. As one leader explained, "(paraphrased) We need a criteria that does not blur where a person comes from. I have seen where an investigator comes from [an] institution that may not be highly esteemed, but the proposal was stellar, and the reviewers say the individuals won't be able to execute what they planned out. No one should be trying to become Inspector Gadget to see whether the institution has the infrastructure, because one assumption we make is if a sponsored research program allows a proposal to go forward, we are saying they can effectively execute this project."

- The process can be biased towards institutions with more robust research infrastructure (4 responses). Some leaders described the challenges of maintaining the fairness of the process and increasing the diversity of awards. For example, one leader stated, "(paraphrased) If you are at a community college or small university and have a hell of a teaching load, then you are trying to build a research portfolio without a lot of resources. There may be one person in the sponsored research office, and putting together a proposal is an incredible burden on you that fundamentally is going to make you fail. So these groups don't have time to write 15 pages and do the budget, and the administrative burden and complexity of what we are asking for hinders our ability to reach those groups that we really want to encourage…. If you're at [an R1] you don't need it, you have hundreds of people helping you get this done; in fact you probably have people writing your proposal for you. That to me is the biggest challenge we have; how do you lower the bar for one group while still keeping standards?"

14. What barriers to participation do NSF constituencies perceive in the merit review policy and processes?

Key finding

14.1. RFI respondents identified several factors that could discourage people from submitting proposals to NSF. These included perceived bias in the merit review process; intimidation due to administrative burden, excessive jargon in NSF guidance, and historically low success rates; limited guidance from NSF on how to navigate the process; and lack of reviewer expertise needed to make an informed assessment of the proposal. ◢

MRX is interested in the experiences and perspectives of those who have considered submitting and/or submitted proposals in the past. We invite you to share your insights and describe any opportunities you believe would improve implementation of the Merit Review criteria, policy, and processes based on your experience as a proposer or investigator. This includes any experiences that may have encouraged or dissuaded you from submitting proposals to NSF. We are especially interested in learning (a) how NSF guidance (e.g., as provided in the NSF PAPPG, program solicitations, or other funding opportunity announcements), may have played a part in your decision(s) whether to submit proposals, and (b) how NSF might best support investigators interested in submitting a proposal to NSF. We received 84 responses to this RFI question. Eight respondents indicated that they decided not to submit a proposal to NSF, 20 respondents indicated they had submitted a proposal in the past five years, and 18 respondents indicated they had reviewed a proposal in the past five years.

Themes related to experiences that dissuaded respondents from submitting a proposal to NSF include the following:

- Process is biased or arbitrary (13 responses). RFI respondents shared beliefs that the merit review process is biased, unfair, or arbitrary. Six respondents described the process as biased; main themes of these responses included natural biases and uneven assessments that come from the process of relying on judgments of POs and reviewers. Others noted that reviewers and POs discriminate based on ethnicity, program type, or proposers' education backgrounds. One respondent shared that scientists from smaller institutions or MSIs or from underrepresented populations have encountered "inappropriate or biased comments from reviewers." Two respondents noted an issue of favoritism and bias toward proposers well-connected to the process and reviewers. One stated, "It seems like an arbitrary system at best or at worst one that benefits those who are connected. I know colleagues at my institution who routinely receive reports from 'inside' the panel." Another added, "Reviewers often prioritize funding for individuals within their professional networks." Lastly, three respondents described inter-reviewer variation as discouraging to proposers, with one noting, "There is a great deal of ambiguity in the review process that may lead to two different decisions on the same proposal by two slightly different panels. This ambiguity hurts the community's trust in the process."
- Process is intimidating (11 responses). Some respondents noted that proposers are intimidated by the merit review process, including how to navigate it and its administrative burden, and by POs, funding rates, jargon in NSF guidance, and the general evaluation process. Seven respondents noted that the administrative burden of the process is daunting, with two stating that changing requirements are difficult to keep up with. One respondent shared, "Thinking back to my early career perspectives on writing proposals, I think program officers were very intimidating and asking questions of them seemed like it might have detrimental impacts on my chances, so I did not ask." Another added that as a new PI, "learning how to navigate OSP [Office of Sponsored Programs] and the budgeting process was the most difficult and demoralizing part for me, by a large margin." One noted that the evaluation practice of labeling proposals as "poor" is "not only in poor taste but actively discourages" researchers NSF aims to attract. One stated that low success rates are intimidating to potential proposers.
- NSF provides limited guidance (11 responses). Some respondents noted that guidance on the process is limited (not available, difficult to access, or insufficient) and that some researchers do not understand whether they are a fit for submissions. Respondents struggled with "generic boilerplate guidance." One respondent said they were told to highlight the important aspects of the RFP and to directly address the merit review criteria in their proposal, which resulted in an award, "But there is no guidance from NSF to follow this format." Three respondents were particularly confused by guidance or lack thereof regarding Broader Impacts, with one stating, "New PIs do not understand 'Broader Impacts,'" and another noting, "NIH folks were very confused by the BI criterion; my first few years at NSF after leaving NIH/biomedicine, I had multiple conversations about it with former colleagues who weren't sure if they could propose to NSF." On the other hand, one respondent noted the value of NSF's investment in Advancing Research Impact in Society (ARIS) in terms of providing guidance for both researchers and proposal support professionals to thoughtfully address the Broader Impacts criterion.
- Reviewers are not qualified to evaluate submissions (9 responses). Respondents noted that some NSF reviewers do not have relevant expertise to evaluate proposals. One stated, "Researchers have reported

developing a proposal for a specific disciplinary panel only to find out that the proposal was reviewed by a panel of diverse and unexpected expertise." Some respondents shared concerns about "unqualified reviewers" and "inconsistent expertise among reviewers" leading to "inadequate evaluations."

- Merit review policy and processes do not discourage submissions (4 responses). Some respondents stated that NSF guidance is not a deciding factor in whether to submit a proposal. One stated, "I always submit proposals that I consider innovative or transformative to NSF and have never decided against a submission to NSF because of the NSF guidelines." One added that they have always "felt encouraged" to submit proposals. Another stated, "Overall I believe the merit review process is solid in that it provides program officers with sufficient flexibility to fund intellectually risky research…. I don't see the merit review process as being anywhere close to the top of any list concerning disincentives to submit to NSF."

Themes related to how NSF guidance played a part in their decision whether to submit a proposal include the following:

- Clear guidance supports submissions (6 responses). Some respondents noted that NSF's Proposal & Award Policies & Procedures Guide (PAPPG) is very clear, with one stating that the "transparency" of the guidance in the PAPPG, program solicitations, and other announcements have played a "key role" in encouraging them to submit proposals and developing an understanding of NSF's expectations on both merit review criteria. Another respondent added, "I find NSF's guidance to be among the best in the federal funding agencies." However, this same respondent suggested that NSF work more with research development and administration communities to vet new guidance and funding opportunity announcements: "We live in those documents every day—and we invariably spot things that could've been fixed (ambiguous language, contradictory statements, etc.) if they'd just asked us to read it over before publishing."
- Guidance does not support submissions: it is unclear and inaccessible (4 responses). One respondent noted that NSF could provide clearer solicitation guidance and suggested providing examples of strong proposals to "help proposers understand the nuances of a solicitation." This respondent added that the nuanced and demanding review process may unintentionally decrease the diversity of proposers. Another respondent echoed this, sharing that NSF should "ensure that expectations for both review criteria are clearly laid out in program announcements, along with direct inclusions of helpful resources." This respondent added that language about Broader Impacts should be centered in solicitations and should not require prospective investigators to consult external materials and websites. A third respondent stated every solicitation should include specific guidance for the Broader Impacts criterion that proposers should focus on: "If you provide generic boilerplate guidance, we don't know what you're looking for." Finally, one respondent noted that there are "substantial barriers to understanding NSF's rules and regulations" and that NSF "might consider strategies to making the PAPPG more accessible." The respondent stated that these barriers are especially prevalent for staff at smaller institutions, community colleges, MSIs, or other institutions that do not have an Office of Sponsored Research or an equivalent.
- Guidance does not support submissions: it changes too frequently (2 responses). One respondent noted that NSF guidance "changes too frequently, requiring new formats or additional information for very little value," and that "most new solicitations have an onerous number of 'additional criteria' beyond the

two merit review criteria," noting that this "dissuades some researchers from submitting as they don't have the right team or their university doesn't have the capacity to do all the 'extra' things." Another respondent shared concerns with changing guidance on Broader Impacts, adding that it's "unfair to proposers to have changed what they're [NSF] looking for regarding Broader Impacts without prior notice."

Themes related to how NSF might best support investigators interested in submitting a proposal to NSF include the following:

- More outreach, feedback, and communication (16 responses). Respondents noted that more communication with and feedback from POs would be helpful, suggesting a variety of methods of accomplishing this. These included "basic acknowledgement" and more responsiveness to inquiries to keep proposers engaged in the process, mentoring by retired or successful PIs, and structured workshops and outreach to individuals and institutions. Three respondents noted that reducing "dwell time," or time between submission of proposals and awards, would encourage participation, and one said this would be particularly true for early-career researchers. Two respondents recommended mentoring, particularly mentoring aimed at newer PIs and PIs at non-R1 institutions. Three respondents recommended workshops aimed at demystifying application and funding processes, including explaining funding opportunities across different NSF directorates and divisions, and clarifying the nuances of different funding mechanisms, their focal audiences, and specific evaluation criteria. Lastly, a respondent noted that members from marginalized groups may need "specific follow-up and encouragement to reapply," and that members of these groups or institutions less familiar with the process may benefit most from direct outreach.
- Changes to the merit review process (15 responses). Respondents suggested a variety of changes to the merit review process that they believe would support investigators interested in participating and address some of the barriers to participation. These included the following:
  - Changes to improve feedback. One respondent recommended introducing a rebuttal mechanism allowing proposers to respond to errors or misunderstandings in reviews. Similarly, a respondent suggested a panel that follows a single submission through two review and resubmission cycles to motivate reviewers to provide useful and action-oriented feedback and enable proposers to address criticism. One respondent noted that sometimes reviewers "miss critical things, even things that were in the submission, and there is no opportunity to address them." Another respondent suggested "standing panels and reviewers" where reviewers serve for multiple cycles to enhance the consistency of reviews over time and ensure that prior feedback is meaningfully considered in resubmission evaluations. One respondent suggested abandoning the letter and number rating system in favor of a holistic evaluation approach that would "encourage a more comprehensive assessment of a project's merits and possible outcomes." Another respondent encouraged NSF to "evaluate proposals as pieces of the puzzle rather than a complete story."
  - Changes to reduce burden. Other suggestions centered on reducing administrative burden. One respondent suggested NSF keep requirements for background documents constant for at least five years at a time and provide flexibility in how they are prepared, specifically suggesting NSF not require the use of ScienCV. Another respondent suggested streamlining the application process to limit burden on researchers and institutions with fewer resources. Similar to the concern about the

requirements burden, one respondent suggested a return to the simple two criteria of Intellectual Merit and Broader Impacts, as "many of the changes of the past 5-10 years have significantly increased the burden of preparing and reviewing proposals without appreciably improving the quality of research funded."

  - Changes to reduce bias. Two respondents proposed solutions to reducing bias in the process. One recommended implementing a double-blind review process. The other recommended developing a large language model specifically tailored for the NSF review process to "ensure reviewers fully understand the proposals they assess, assist in flagging potential biases, and ensure consistency in the application of review criteria."

- More qualified reviewers and POs (8 responses). Respondents had a variety of suggestions to address what they saw as inconsistency among reviewer expertise and implementation of the evaluation process. They suggested reviewer pre-trainings run by POs and aligned to specific solicitations to develop consistency in how reviewers interpret and apply the merit review criteria, reduce biases, improve the quality of feedback provided to applicants, and help reviewers understand NSF's broader mission and goals. One respondent suggested enhancing the selection process for reviewers to ensure they have the relevant expertise. Another respondent suggested NSF publicize reviewer rosters and provide channels for applicants to report inappropriate feedback in order to "hold reviewers accountable." One respondent suggested hiring more senior scientists or managers with scientific backgrounds as POs and recommended that NSF "avoid filling these roles with individuals who advanced through political means, such as former deans or executives with little scientific experience" to "ensure the focus remains on advancing science."

15. To what extent do the reported outputs and outcomes of funded research align with NSF's mission?

Key findings

15.1. Over half of staff survey respondents reported that the data NSF collects are somewhat or very effective at helping NSF assess whether funded projects support NSF's mission to advance scientific knowledge; fewer than half reported that the data are somewhat or very effective at helping NSF assess whether funded projects benefit society.

15.2. Interviewed leaders reported using a range of sources to assess whether funded projects help support NSF's mission to advance scientific knowledge and benefit society, including interim and final reports, external evaluations, and external review and advisory committees. ◢

How effective do you think these data are for helping NSF to assess whether funded projects support NSF's mission to advance scientific knowledge? How effective do you think these data are for helping NSF to assess whether funded projects support NSF's mission to benefit society? For advancing scientific knowledge, 61 percent of staff survey respondents reported that the data NSF collects are somewhat or very effective; only 39 percent of respondents reported the data collected are somewhat or very effective at helping NSF assess whether funded projects support NSF's mission to benefit society (Exhibit 15.1).

- DDs (14 percent) were least likely to report that collected data are somewhat or very effective at assessing whether funded projects benefit society, compared to DDDs (69 percent) and DADs and DOHs (75 percent).

What types of evidence does your [directorate/office] collect and examine to assess whether the projects that received funding helped support NSF's mission to advance scientific knowledge and benefit society? Interviewed leaders reported using a range of sources to assess projects. The most commonly cited sources include interim and final project reports (4 responses), external evaluations (3 responses), and external review and advisory committees (2 responses).

16. What guidance does NSF offer to PIs on documenting outcomes that address the merit review criteria in annual and final reports?

This question was only addressed in the focus groups with VPRs and AC members. Per OMB requirements, these results cannot be publicly disseminated.

17. How might NSF better measure the outcomes of funded research with respect to each merit review criterion?

Key findings

17.1. Staff survey respondents suggested that NSF could better measure the outcomes of funded research for Intellectual Merit by tracking longer-range outcomes of funded projects and going beyond metrics related to publications and citations. For Broader Impacts, staff also suggested tracking longer-range outcomes of funded projects, as well as requiring more information on Broader Impacts from PIs in annual and final reports.

17.2. Interviewed leaders suggested that NSF could better measure the outcomes of funded research by conducting additional analyses on data that NSF already collects and collecting data on project personnel beyond the PI and co-PI.

17.3. RFI respondents suggested that NSF could better support awardees in reporting outcomes of their awards with respect to each criterion. Suggestions included tracking more and longer-range outcomes, adding report sections for outcomes related to each merit review criterion, increasing accountability for reporting by increasing public access to reports, reducing the administrative burden of reporting on investigators, and improving the suite of tools for reporting that NSF offers PIs. ◢

What, if any, additional information do you think NSF should be collecting to monitor progress toward *advancing scientific knowledge* that is not being collected? Please include both quantitative and qualitative outputs and outcomes if relevant. Themes among the 119 responses received on this staff survey question include the following:

- Track longer range outcomes of funded projects (35 responses). Staff indicated that many outcomes related to advancing scientific knowledge are not evident within the typical award period. Some staff suggested that tracking a project's publications and contributions after the award would help NSF better understand progress toward its goals. As one staff member commented, "It would be more useful if the timing were shifted: for a three-year project, require a report after two years, three years, and five years, to give time to see the impact of the research after the funding is completed. The

impacts are frequently not visible when the money runs out, but there's no reporting beyond that point."

- Go beyond counts of publications and citations (15 responses). Staff noted the limitations of currently used metrics related to publications in that they provide little information on how intellectually influential a project was, especially given the limited period for reporting. Some staff offered ideas for alternative measures; for example, one noted, "At this point, NSF collects information on Scientific Publications, it would be helpful to have a way to measure how influential a paper has been (e.g., H factor, which is imperfect) but also on how disruptive the ideas in a particular publication have been at changing the status quo. This may be very hard to quantify, but perhaps if the measurements were done by tracking discoveries done in a particular field, say in the last few decades, and how these get associated with NSF (or NASA [National Aeronautics and Space Administration] or DoE [Department of Energy], etc.) awards may be a way to assess NSF's investment influence on advancing Science."

What, if any, additional information do you think NSF should be collecting to monitor its progress toward *benefitting society* that is not being collected? Please include both quantitative and qualitative outputs and outcomes if relevant. We received 129 staff survey responses. Themes include the following:

- Require more information on Broader Impacts from PIs in annual and final reports (27 responses). Staff survey respondents noted that a project's Broader Impacts receive little attention in NSF's reporting templates. One staff member explained, "Specific questions about BI outcomes would be helpful. As is, these are not spelled out well in the current annual report forms. PIs should be prompted to relate their BI activities to what they said they wanted to accomplish and, further, how these outcomes have societal benefit. Making this aspect of annual reporting more explicit might have a positive effect on how PIs propose BI activities in the first place."
- Track longer range outcomes of funded projects (22 responses). As with NSF's goal to advance scientific knowledge, staff reported that progress toward benefitting society is not always evident within an award period. As one respondent explained, "The benefits to society, especially those stemming from a new product or technology, may come many years or decades after the initial award."

Is there additional information that NSF could be collecting to monitor its progress to advance scientific knowledge and benefit society? Themes from interviewed leaders' responses include the following:

- Conduct additional analyses of data NSF already collects (3 responses). Three leaders indicated that NSF could dig deeper into the data they already collect from PIs and institutions. As one leader explained, "(paraphrased) There are data analysis mechanisms that we're not currently fully leveraging to be able to assess outcomes and short-term impacts…. How many students went here and there, how many did you bring into your lab? That's a hard thing to measure, and I'm also wary about adding burden to PIs. So I'm wondering if there's stuff we can do through text analysis, and drawing info out that way."
- Collect data on project personnel (2 responses). Some leaders noted that NSF does not track the staff that work on funded projects and explained that this information could help the agency monitor progress toward broadening participation. One leader stated, "(paraphrased) We know who the PI,

co-PIs are, the institution. But we don't actually know who in the lab actually gets supported…. We also don't have good data on the demographics on who is actually paid, so [we] don't know anything about gender, ethnicity, a whole host of criteria that would be helpful to understand how equitable the actual distribution of funding is. Because there's deep socioeconomic status, ethnicity, gender inequities that more data could give us insight into and help us think about how to tweak programs, which could help us think about Broader Impacts."

Is there anything about the *quality* of the data that NSF is collecting that might make it difficult to use to monitor its progress? Due to time constraints, most leaders were not asked this question. Among the three leaders that were asked, all identified challenges in identifying indicators and metrics for NSF's goals to advance knowledge and benefit society. As one leader explained, "(paraphrased) We are hoping to take all of the great ideas coming out of NSF-funded foundational research and hoping to accelerate the translation of that research. We are finding there are some challenges in identifying what the metrics are and identifying not just what is the right data, but how do we access that data so that we can use it for evaluation? So that's currently the challenge: what is the right data, and how do you get it?"

MRX is interested in exploring how NSF could better support awardees in demonstrating and documenting outcomes of their awards in advancing knowledge (Intellectual Merit) and benefiting society and contributing to the achievement of desired broader or societal outcomes (Broader Impacts). We invite you to share your insights on how NSF might better support awardees in demonstrating and documenting outcomes of their awards without unnecessarily increasing awardees' administrative burden of reporting. We received 67 responses to this RFI question. Of these responses, 22 reported they had experience as an investigator, 3 as a user of public outcomes reports, and 4 indicated another perspective. Most commonly, respondents suggested that NSF could better support awardees by doing the following:

- Track more outcomes (11 respondents). Some respondents suggested revising outcomes to emphasize measures that are real, meaningful, and reproducible, or tracking additional outcomes that could help PIs to better showcase their accomplishments. Some of these additional outcomes included longer-range or broader outcomes, the number and demographics of funded graduate students and postdocs, the demographics of PIs, or the outcomes originally described in the project proposal. For example, one respondent shared, "I'd … suggest having a larger scope of what is considered Intellectual Merit: teaching, papers, visiting/guest lecturer, graduate students, conference talks and posters, workshops run/attended, development programs and training attended. Or service based like mentoring (not just grad student mentoring but outside of that), conference roles (organizer, liaison, session coordinator, etc.), or engagement activities like media interviews, media publications ([including] YouTube, Radio, newspaper, online blogs, etc.), outreach events like science festivals and in-school engagement programs."
- Strengthen the link between the merit review criteria and reported outcomes (7 responses). Relatedly, some respondents suggested modifying the current reporting mechanisms to strengthen their alignment with the merit review criteria. For example, one respondent wrote, "NSF reporting does not match up with the two merit review criteria, which just compounds the unequal weighting of the two criteria. There is no specific BI section in the reports, and the section on impact has really nothing to do

with the BI of the proposal. Annual reports can be submitted and approved without any mention of the BI. This is a function of the disparate ways program officers view the merit criteria."

- Increase accountability in reporting by increasing public access to reports (7 responses). Some respondents called for more public accountability in NSF reporting, for example, by deidentifying reports and making them available to the public or by making the types of analyses that NSF conducts more transparent to the public. For example, one respondent stated, "There is a gap in how well the outcomes of funded projects are shared with the public and broader scientific community. Low-quality papers and results that remain behind closed doors hurt the credibility of NSF's funded research.… NSF should formalize a process for requiring researchers to share their findings, datasets, and methodologies on public platforms like GitHub or Zenodo. This transparency not only supports the reproducibility of research but also serves as a real-time way to verify whether research has achieved its intended Intellectual Merit and Broader Impacts."
- Reduce the administrative burden for reporting (6 responses). Some respondents highlighted the substantial administrative burden that is required to comply with NSF reporting requirements. For example, one respondent shared, "The Merit Review Criteria and other reporting requirements have become far too onerous. Concocting paragraph upon paragraph to attempt to address whatever NSF has in mind (which is far from clear) is now rapidly replacing effort on writing relevant and carefully considered proposals."
- Improve the suite of tools for reporting that NSF offers PIs (5 responses). Some respondents suggested ways that NSF could improve the suite of tools that they offer PIs, either by creating additional tools for reporting (for example, templates or dashboards with key outcomes) or by integrating NSF's reporting platform with existing tools. One respondent wrote, "NSF could further reduce the reporting burden by integrating with tools and platforms that awardees already use for project management and research documentation, such as ORCID, GitHub, and Google Scholar. This would allow awardees to easily link their project outcomes to their ongoing work without requiring them to duplicate efforts in separate reports.… This would streamline the documentation process by reducing the need for additional manual reporting, while still ensuring that NSF has access to all relevant outcomes data."
- Other. A few respondents shared additional insights about how NSF could better support awardees in their reporting. Four people highlighted that the timeline for reporting is often misaligned with when the work is actually conducted, for example, reports that are due at nine months of activity when the award is scheduled to take a full year. Other respondents suggested that NSF could benefit from soliciting input from additional groups, for example, community partners that contribute to NSF projects (3 respondents) or members of interdisciplinary teams who may not be listed as the PI (2 respondents). Finally, two respondents shared that it would be helpful for awardees to receive feedback on their annual reports, or at least for NSF to clarify how they plan to use the information submitted.

MRX welcomes any other comments on or suggestions for improving NSF's current Merit Review criteria, policy, and processes. It also welcomes information about aspects of Merit Review criteria, policy, and processes that are currently working well. Most responses to this question were relevant to other questions on the RFI. However, several respondents had additional comments or suggestions for

MRX that did not fall under any of the other questions in the RFI. Common themes from these responses include the following:

- Merit review is working well (8 responses). Some respondents commented on the important job that NSF is doing, and that the merit review process is working well. As one person said, "My overall impression of NSF proposal review is positive. It's *far* better than NIH in terms of clarity, fairness, high standards, feedback to PIs, and quality of reviews. I've served on over a dozen panels, and I have *always* thought the PO and the panel did a very good job!"
- NSF needs more accountability and oversight (4 respondents). Some respondents suggested that NSF could benefit from more financial oversight and accountability, for example, in how funding decisions are made and to ensure that proposals are reviewed in a timely manner.

References

National Science Board. 2005. *Report of the National Science Board on the National Science Foundation's Merit Review System* (NSB-05-119). Arlington: National Science Foundation.

National Science Board. 2011. *National Science Foundation's Merit Review Criteria: Review and Revisions* (NSB/MR-11-22). Arlington: National Science Foundation.

Srivastava, A., and S.B. Thomson. 2009. "Framework Analysis: A Qualitative Methodology for Applied Research Note Policy Research." *Journal of Administration and Governance, 4*(2): 72–79.

Appendix A.

Exhibits

Respondent characteristics

Exhibit 0.1. Number of completed NSF staff surveys (overall and by position type, directorate or office, gender, race, ethnicity, and disability)

Respondent characteristic	Frequency	
	Percentage of sample	Responses (N)
Overall		
All respondents	100%	234
Position type		
Program Officer	83%	194
Deputy Division Director	7%	17
Division Director	3%	7
Deputy Assistant Director and Deputy Office Head	2%	5
Other NSF staff position	5%	11
Directorate or office		
BIO	15%	35
CISE	6%	15
ENG	11%	25
GEO	13%	31
MPS	16%	38
SBE	6%	15
EDU	11%	26
TIP	3%	8
OD/OIA	2%	5
OISE	<1%	1
Not reported	15%	35
Gender		
Female only	40%	93
Male only	39%	91
Other category and not reported	21%	50
Race		
American Indian or Alaska Native	0%	0
Asian	7%	16
Black or African American	6%	15
Native Hawaiian or other Pacific Islander	0%	0
White	59%	137
Other or multiple racial categories selected	3%	6
Not reported	26%	60

Respondent characteristic	Frequency	
	Percentage of sample	Responses (N)
Ethnicity		
Hispanic or Latino	6%	13
Not Hispanic or Latino	71%	166
Not reported	24%	55
Disability		
Identifies as having a disability	6%	14
Does not identify as having a disability	17%	40
Not reported	77%	180

Source: Merit Review Examination Staff Survey, administration log; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"; Merit Review Examination Staff Survey, question 47: "How do you currently describe yourself?"; Merit Review Examination Staff Survey, question 48: "Are you Hispanic or Latino?"; Merit Review Examination Staff Survey, question 49: "What is your racial background?"; Merit Review Examination Staff Survey, question 50: "Do you identify as having a disability?"

Note: In this exhibit, we show the number of completed NSF staff surveys. Along with overall responses, we show the number of responses by position type, directorate or office, gender, race, ethnicity, and disability. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 0.2. Characteristics of RFI respondents

Respondent characteristic	Frequency	
	Percentage of sample	Responses (N)
Overall		
All respondents	100%	130
Experience		
Proposer	70%	91
Proposer in the last 5 years	16%	21
Reviewer	49%	64
Panel reviewer	25%	32
Ad hoc reviewer	15%	20
Reviewer in the last 5 years	15%	20
Did not submit	6%	8
Investigator	18%	23
Public reports user	2%	3
Other perspective	3%	4

Source: Merit Review Examination RFI, administration log; Merit Review Examination RFI, question 1 (excerpted) and 2 (excerpted): "Individuals responding to this request are encouraged to indicate whether their perspectives are informed by experience(s) preparing and/or reviewing proposals to NSF."; Merit Review Examination RFI, question 3 (excerpted): "Individuals responding to this request are encouraged to indicate whether they submitted or decided not to submit a proposal, and whether these experiences occurred within the past five years."; Merit Review Examination RFI, question 4 (excerpted): "Individuals responding to this request are encouraged to indicate whether they served on a panel and/or as ad hoc reviewers, and whether these experiences occurred within the past five years."; Merit Review Examination RFI, question 5 (excerpted): "Individuals responding to this request are encouraged to indicate whether their suggestions are based on experiences as investigators, users of public outcomes reports, or another perspective."

Note: In this exhibit, we show the number of responses to the RFI. Along with the overall number, we show the number of responses that indicated experience as a proposer, reviewer, investigator, public reports user, or other perspective, or if they decided not to submit a proposal to NSF. If a respondent's answer to any of the RFI questions indicated relevant experience, the response was coded to that category. Responses were not coded to the category if they stated that they did not have the relevant experience, or if they did not say.

RFI = Request for Information.

Select survey, RFI, and interview responses

1. How do PIs, reviewers, and NSF leaders and staff interpret the merit review criteria?

Exhibit 1.1. Percentage of NSF staff who indicated that the following factors are somewhat or very important for receiving a high rating on a proposal

Item	Percentage
Merit review principle: All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.	99%
Merit review principle: NSF projects, in the aggregate, should contribute more broadly to achieving societal goals.	97%
Merit review principle: Meaningful assessment and evaluation of NSF-funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of Broader Impacts and the resources provided to implement projects.	90%
Intellectual Merit criterion: The potential to advance knowledge.	99%
Broader Impacts criterion: The potential to benefit society and contribute to the achievement of specific, desired societal outcomes.	95%
Merit review element: What is the potential for the proposed activity to advance knowledge and understanding within its own field or across different fields (Intellectual Merit)?	98%
Merit review element: What is the potential for the proposed activity to benefit society or advance desired societal outcomes (Broader Impacts)?	97%
Merit review element: To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?	94%
Merit review element: Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?	98%
Merit review element: How well qualified is the individual, team, or organization to conduct the proposed activities?	97%
Merit review element: Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?	91%
Other important factor for receiving a high rating (Please specify)	24%
Responses (N)	234

Source: Merit Review Examination Staff Survey, question 4: "According to your understanding of the merit review policy, how important are the following factors for receiving a high rating on a proposal?"

Note: In this exhibit, we show the percentage of NSF staff indicating that the listed factors are "Somewhat important" or "Very important" for receiving a high rating on a proposal. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Because substantially fewer respondents indicated a response to the "Other important factor for receiving a high rating" question stem, the percentage shown for that item represents the number of respondents who selected "Somewhat important" or "Very important" divided by the total number of respondents to the survey. For all other question stems, the number of respondents who did not answer the question is not included in the denominator.

Exhibit 1.2. Percentage of NSF staff who indicated that the following factors are somewhat or very important for receiving a high rating on a proposal (by position type)

Item	Program Officer	Deputy Division Director	Division Director	Deputy Assistant Director and Deputy Office Head	Other NSF staff position	*p*-value
Merit review principle: All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.	98%	100%	100%	100%	100%	0.960
Merit review principle: NSF projects, in the aggregate, should contribute more broadly to achieving societal goals.	97%	94%	100%	100%	91%	0.765
Merit review principle: Meaningful assessment and evaluation of NSF-funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of Broader Impacts and the resources provided to implement projects.	91%	88%	100%	80%	82%	0.686
Intellectual Merit criterion: The potential to advance knowledge.	98%	100%	100%	100%	100%	0.960
Broader Impacts criterion: The potential to benefit society and contribute to the achievement of specific, desired societal outcomes.	96%	88%	86%	100%	91%	0.438
Merit review element: What is the potential for the proposed activity to advance knowledge and understanding within its own field or across different fields (Intellectual Merit)?	98%	100%	100%	100%	100%	0.933
Merit review element: What is the potential for the proposed activity to benefit society or advance desired societal outcomes (Broader Impacts)?	97%	94%	86%	100%	100%	0.486
Merit review element: To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?	95%	94%	86%	100%	82%	0.295

Item	Program Officer	Deputy Division Director	Division Director	Deputy Assistant Director and Deputy Office Head	Other NSF staff position	*p*-value
Merit review element: Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?	97%	100%	100%	100%	100%	0.902
Merit review element: How well qualified is the individual, team, or organization to conduct the proposed activities?	97%	100%	100%	100%	100%	0.867
Merit review element: Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?	91%	94%	86%	80%	100%	0.682
Other important factor for receiving a high rating (Please specify)	25%	12%	29%	20%	36%	0.652
Responses (N)	194	17	7	5	11	

Source: Merit Review Examination Staff Survey, question 4: "According to your understanding of the merit review policy, how important are the following factors for receiving a high rating on a proposal?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"

Note: In this exhibit, we show, by position type, the percentage of NSF staff indicating that the listed factors are "Somewhat important" or "Very important" for receiving a high rating on a proposal. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review during the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Because substantially fewer respondents indicated a response to the "Other important factor for receiving a high rating" question stem, the percentage shown for that item represents the number of respondents who selected "Somewhat important" or "Very important" divided by the total number of respondents to the survey. Chi-squared tests were used to assess whether response patterns for each item differ by position type; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

Exhibit 1.3. Percentage of NSF staff who indicated that the following factors are somewhat or very important for receiving a high rating on a proposal (by directorate or office)

Item	BIO	CISE	ENG	GEO	MPS	SBE	EDU	TIP	OD/ OIA	OISE	Directorate not reported	*p*-value
Merit review principle: All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.	97%	100%	100%	100%	100%	100%	96%	100%	100%	D	D	0.921
Merit review principle: NSF projects, in the aggregate, should contribute more broadly to achieving societal goals.	94%	100%	100%	97%	95%	100%	96%	88%	100%	D	D	0.883
Merit review principle: Meaningful assessment and evaluation of NSF-funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of Broader Impacts and the resources provided to implement projects.	91%	77%	88%	97%	89%	93%	92%	88%	100%	D	D	0.811
Intellectual Merit criterion: The potential to advance knowledge.	97%	100%	100%	100%	100%	100%	100%	100%	100%	D	D	0.605
Broader Impacts criterion: The potential to benefit society and contribute to the achievement of specific, desired societal outcomes.	97%	93%	100%	94%	89%	93%	96%	88%	100%	D	D	0.821
Merit review element: What is the potential for the proposed activity to advance knowledge and understanding within its own field or across different fields (Intellectual Merit)?	97%	100%	100%	97%	100%	100%	100%	100%	100%	D	D	0.771
Merit review element: What is the potential for the proposed activity to benefit society or advance desired societal outcomes (Broader Impacts)?	97%	100%	100%	94%	97%	93%	96%	100%	100%	D	D	0.950
Merit review element: To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?	89%	93%	100%	83%	100%	93%	96%	100%	100%	D	D	0.150
Merit review element: Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?	97%	93%	100%	94%	97%	100%	100%	100%	100%	D	D	0.737

Item	BIO	CISE	ENG	GEO	MPS	SBE	EDU	TIP	OD/ OIA	OISE	Directorate not reported	*p*-value
Merit review element: How well qualified is the individual, team, or organization to conduct the proposed activities?	91%	100%	100%	97%	97%	100%	100%	100%	100%	D	D	0.666
Merit review element: Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?	91%	80%	92%	87%	89%	100%	92%	100%	100%	D	D	0.825
Other important factor for receiving a high rating (Please specify)	29%	7%	16%	35%	18%	27%	35%	25%	60%	D	D	0.228
Responses (N)	35	15	25	31	38	15	26	8	5	D	D	

Source: Merit Review Examination Staff Survey, question 4: "According to your understanding of the merit review policy, how important are the following factors for receiving a high rating on a proposal?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show, by directorate or office, the percentage of NSF staff indicating that the listed factors are "Somewhat important" or "Very important" for receiving a high rating on a proposal. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Because substantially fewer respondents indicated a response to the "Other important factor for receiving a high rating" question stem, the percentage shown for that item represents the number of respondents who selected "Somewhat important" or "Very important" divided by the total number of respondents to the survey. Chi-squared tests were used to assess whether response patterns for each item differ by directorate or office; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 1.4. Percentage of NSF staff who indicated that it is generally clear how to assess a proposal's merit against the Intellectual Merit criterion or the Broader Impacts criterion (overall, by position type, and by directorate or office)

Respondent characteristic	Intellectual Merit	Broader Impacts	Responses (N)
Overall			
All respondents	96%	53%	232
Position type			
Program Officer	96%	52%	192
Deputy Division Director	100%	59%	17
Division Director	71%	29%	7
Deputy Assistant Director and Deputy Office Head	100%	80%	5
Other NSF staff position	100%	55%	11
p-value	0.022*	0.489	
Directorate or office			
BIO	94%	69%	35
CISE	100%	67%	15
ENG	100%	52%	25
GEO	97%	39%	31
MPS	97%	53%	38
SBE	100%	57%	14
EDU	92%	46%	26
TIP	100%	38%	8
OD/OIA	100%	80%	5
OISE	D	D	D
Not reported	D	D	D
p-value	0.544	0.296	

Source: Merit Review Examination Staff Survey, question 11: "Do you think it is generally clear how a proposal's merit is to be assessed against the Intellectual Merit criterion and the Broader Impacts criterion?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff responding "Yes" for the Intellectual Merit criterion or the Broader Impacts criterion. Along with overall results, we show responses by position type and by directorate or office. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type or by directorate or office; *p*-values from these tests are reported.

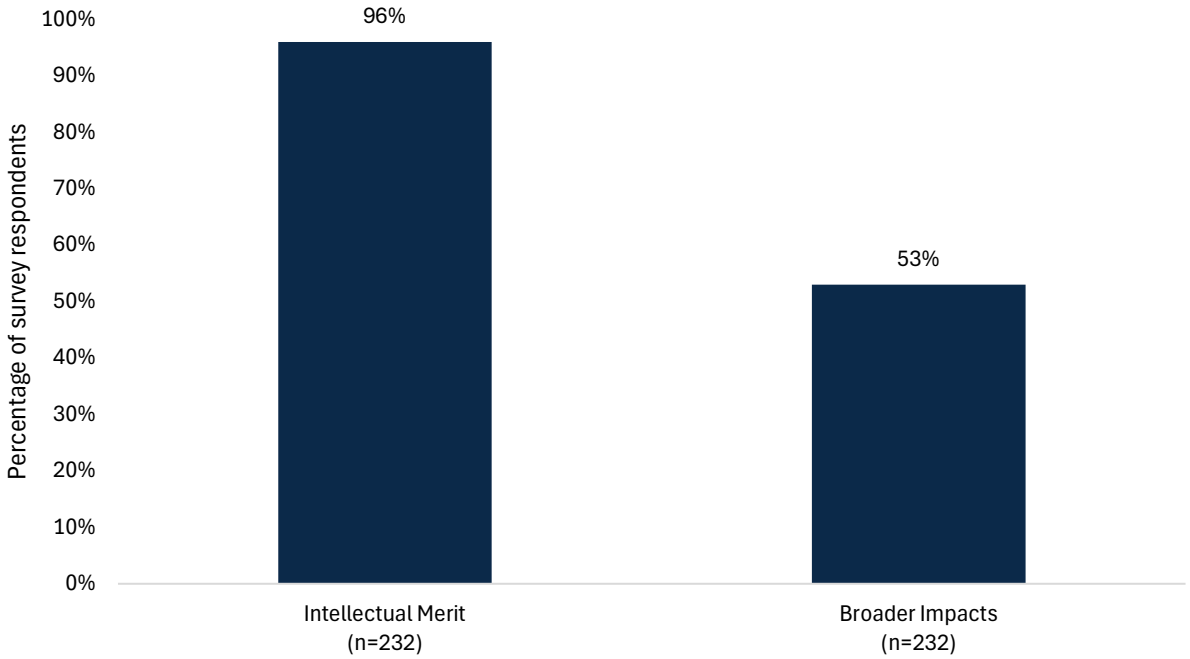
* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 1.5. Percentage of NSF staff who indicated that it is generally clear how to assess a proposal's merit against the Intellectual Merit criterion or the Broader Impacts criterion



Source: Merit Review Examination Staff Survey, question 11: "Do you think it is generally clear how a proposal's merit is to be assessed against the Intellectual Merit criterion and the Broader Impacts criterion?"

Note: In this exhibit, we show the percentage of NSF staff responding "Yes" for the Intellectual Merit criterion or the Broader Impacts criterion. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included.

2. How do PIs, reviewers, and NSF leaders and staff use the merit review criteria?

Exhibit 2.1. Percentage of NSF staff who indicated that the following factors are somewhat or very important for reviewers to assign a high rating to a proposal

Item	Percentage
Merit review principle: All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.	100%
Merit review principle: NSF projects, in the aggregate, should contribute more broadly to achieving societal goals.	74%
Merit review principle: Meaningful assessment and evaluation of NSF funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of Broader Impacts and the resources provided to implement projects.	71%
Intellectual Merit: The potential to advance knowledge.	99%
Broader Impacts: The potential to benefit society and contribute to the achievement of specific, desired societal outcomes.	76%
Merit review element: What is the potential for the proposed activity to advance knowledge and understanding within its own field or across different fields (Intellectual Merit)?	98%
Merit review element: What is the potential for the proposed activity to benefit society or advance desired societal outcomes (Broader Impacts)?	77%
Merit review element: To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?	93%
Merit review element: Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?	96%
Merit review element: How well qualified is the individual, team, or organization to conduct the proposed activities?	96%
Merit review element: Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?	79%
Other important factor for receiving a high rating (Please specify)	19%
Responses (N)	205

Source: Merit Review Examination Staff Survey, question 5: "In your experience, how do reviewers rank the importance of the following when assigning a high rating on a proposal?"

Note: In this exhibit, we show the percentage of NSF staff indicating that the listed factors are "Somewhat important" or "Very important" for reviewers to assign a high rating to a proposal. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Because substantially fewer respondents indicated a response to the "Other important factor for receiving a high rating" question stem, the percentage shown for that item represents the number of respondents who selected "Somewhat important" or "Very important" divided by the total number of respondents to the survey.

Exhibit 2.2. Percentage of NSF staff who indicated that the following factors are somewhat or very important for reviewers to assign a high rating to a proposal (by directorate or office)

Item	BIO	CISE	ENG	GEO	MPS	SBE	EDU	TIP	OD/ OIA	OISE	Directorate not reported	*p*-value
Merit review principle: All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.	100%	100%	100%	96%	100%	100%	100%	100%	100%	D	D	0.774
Merit review principle: NSF projects, in the aggregate, should contribute more broadly to achieving societal goals.	73%	77%	76%	56%	76%	92%	83%	63%	60%	D	D	0.264
Merit review principle: Meaningful assessment and evaluation of NSF funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of Broader Impacts and the resources provided to implement projects.	80%	75%	81%	56%	64%	92%	71%	63%	40%	D	D	0.280
Intellectual Merit: The potential to advance knowledge.	100%	100%	100%	100%	97%	100%	96%	100%	100%	D	D	0.860
Broader Impacts: The potential to benefit society and contribute to the achievement of specific, desired societal outcomes.	83%	69%	71%	56%	70%	100%	96%	63%	60%	D	D	0.010**
Merit review element: What is the potential for the proposed activity to advance knowledge and understanding within its own field or across different fields (Intellectual Merit)?	100%	100%	100%	93%	94%	100%	100%	88%	100%	D	D	0.347
Merit review element: What is the potential for the proposed activity to benefit society or advance desired societal outcomes (Broader Impacts)?	90%	69%	71%	59%	70%	100%	88%	50%	100%	D	D	0.012*
Merit review element: To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?	90%	92%	95%	89%	91%	100%	92%	100%	80%	D	D	0.919
Merit review element: Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?	100%	100%	95%	89%	97%	100%	92%	100%	100%	D	D	0.645
Merit review element: How well qualified is the individual, team, or organization to conduct the proposed activities?	93%	92%	100%	93%	94%	100%	96%	100%	100%	D	D	0.960

Item	BIO	CISE	ENG	GEO	MPS	SBE	EDU	TIP	OD/ OIA	OISE	Directorate not reported	*p*-value
Merit review element: Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?	77%	85%	90%	81%	64%	100%	63%	75%	75%	D	D	0.061
Other important factor for receiving a high rating (Please specify)	23%	23%	14%	15%	18%	38%	4%	25%	60%	D	D	0.193
Responses (N)	31	13	21	27	34	13	24	8	5	D	D	

Source: Merit Review Examination Staff Survey, question 5: "In your experience, how do reviewers rank the importance of the following when assigning a high rating on a proposal?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show, by directorate or office, the percentage of NSF staff indicating that the listed factors are "Somewhat important" or "Very important" for reviewers to assign a high rating on a proposal. Because substantially fewer respondents indicated a response to the "Other important factor for receiving a high rating" question stem, the percentage shown for that item represents the number of respondents who selected "Somewhat important" or "Very important" divided by the total number of respondents to the survey. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. Chi-squared tests were used to assess whether response patterns for each item differ by directorate or office; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 2.3. Percentage of NSF staff who indicated that there are situations in which the importance of the Intellectual Merit criterion and the Broader Impacts criterion are weighted unevenly (overall, by position type, and by directorate or office)

Respondent characteristic	Percentage	Responses (N)
Overall		
All respondents	91%	231
Position type		
Program Officer	91%	193
Deputy Division Director	88%	16
Division Director	100%	7
Deputy Assistant Director and Deputy Office Head	100%	5
Other NSF staff position	90%	10
p-value	0.836	
Directorate or office		
BIO	88%	34
CISE	93%	15
ENG	88%	24
GEO	100%	31
MPS	95%	38
SBE	93%	15
EDU	85%	26
TIP	100%	8
OD/OIA	100%	5
OISE	D	D
Not reported	D	D
p-value	0.417	

Source: Merit Review Examination Staff Survey, question 7: "In your experience, are there situations in which the importance of the Intellectual Merit criterion and the Broader Impacts criterion are weighted unevenly?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff indicating "Yes," there are situations in which the Intellectual Merit criterion and the Broader Impacts criterion are weighted unevenly. Along with overall results, we show responses by position type and by directorate or office. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns differ by position type or by directorate or office; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 2.4. Percentage of NSF staff who indicated that the Intellectual Merit criterion and the Broader Impacts criterion are weighted unevenly with the following frequencies: never, rarely, sometimes, or always (overall, by position type, and by directorate or office)

Respondent characteristic	Never	Rarely	Sometimes	Always	Responses (N)
Overall					
All respondents	0%	8%	66%	27%	209
Position type					
Program Officer	0%	7%	65%	28%	174
Deputy Division Director	0%	14%	71%	14%	14
Division Director	0%	14%	29%	57%	7
Deputy Assistant Director and Deputy Office Head	0%	0%	100%	0%	5
Other NSF staff position	0%	11%	78%	11%	9
p-value				0.263	
Directorate or office					
BIO	0%	14%	62%	24%	29
CISE	0%	0%	64%	36%	14
ENG	0%	0%	62%	38%	21
GEO	0%	6%	74%	19%	31
MPS	0%	3%	53%	44%	36
SBE	0%	0%	79%	21%	14
EDU	0%	14%	73%	14%	22
TIP	0%	0%	38%	63%	8
OD/OIA	0%	0%	80%	20%	5
OISE	D	D	D	D	D
Not reported	D	D	D	D	D
p-value				0.007**	

Source: Merit Review Examination Staff Survey, question 8: "In your experience, how frequently are the merit review criteria weighted unevenly?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff indicating the frequency with which they think the Intellectual Merit criterion and the Broader Impacts criterion are weighted unevenly. Along with overall results, we show responses by position type and by directorate or office. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns differ by position type or by directorate and office; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 2.5. Percentage of NSF staff who indicated that they typically place more weight on the Intellectual Merit criterion, place more weight on the Broader Impacts criterion, or place equal weight on both criteria in their use of the merit review criteria (overall, by position type, and by directorate or office)

Respondent characteristic	More weight on Intellectual Merit	Both equally weighted	More weight on Broader Impacts	Responses (N)
Overall				
All respondents	61%	34%	5%	204
Position type				
Program Officer	64%	31%	5%	169
Deputy Division Director	57%	43%	0%	14
Division Director	43%	43%	14%	7
Deputy Assistant Director and Deputy Office Head	20%	80%	0%	5
Other NSF staff position	44%	44%	11%	9
p-value			0.276	
Directorate or office				
BIO	61%	36%	4%	28
CISE	71%	29%	0%	14
ENG	67%	33%	0%	21
GEO	53%	43%	3%	30
MPS	85%	15%	0%	34
SBE	71%	29%	0%	14
EDU	55%	32%	14%	22
TIP	50%	38%	13%	8
OD/OIA	60%	40%	0%	5
OISE	D	D	D	D
Not reported	D	D	D	D
p-value			0.066	

Source: Merit Review Examination Staff Survey, question 9: "What weight do you typically place on the importance of the Intellectual Merit criterion compared to the Broader Impacts criterion in your use of the merit review criteria?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff indicating the level of weight they typically place on the importance of the Intellectual Merit criterion compared to the Broader Impacts criterion in their use of the merit review criteria. Along with overall results, we show responses by position type and by directorate or office. "More weight on Intellectual Merit" includes the following response options: "Much more weight on Intellectual Merit," "More weight on Intellectual Merit," and "Somewhat more weight on Intellectual Merit." "More weight on Broader Impacts" includes these response options: "Somewhat more weight on Broader Impacts," "More weight on Broader Impacts," and "Much more weight on Broader Impacts." NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns differ by position type or by directorate or office; *p*-values from these tests are reported.

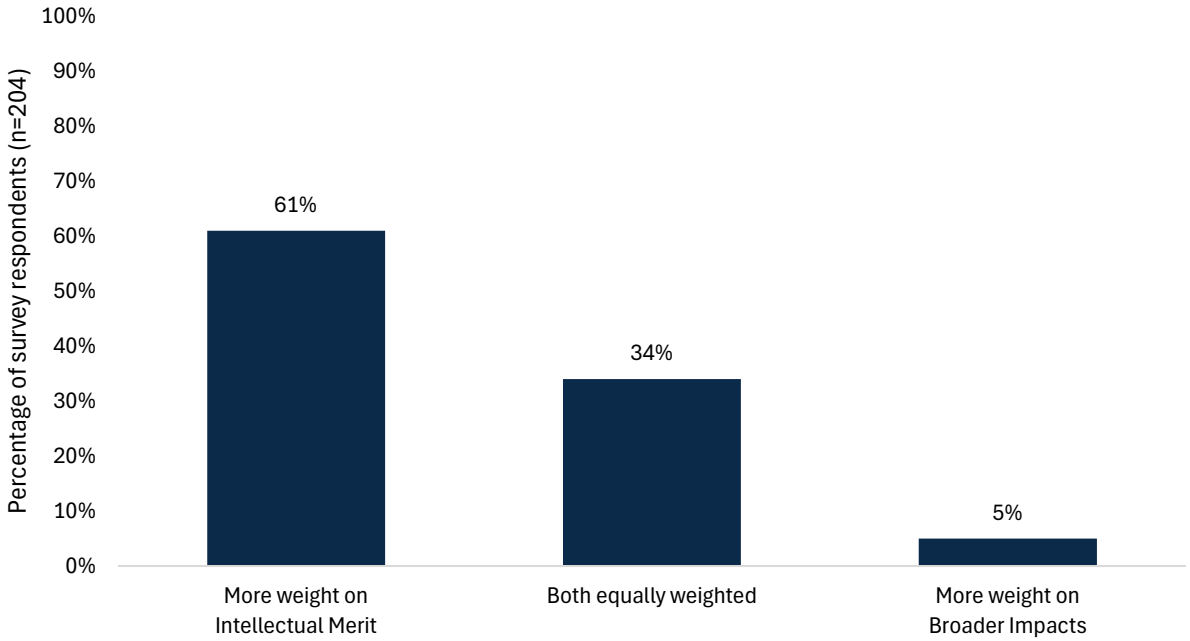
* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 2.6. Percentage of NSF staff who indicated that they typically place more weight on the Intellectual Merit criterion, place more weight on the Broader Impacts criterion, or place equal weight on both criteria in their use of the merit review criteria



Source: Merit Review Examination Staff Survey, question 9: "What weight do you typically place on the importance of the Intellectual Merit criterion compared to the Broader Impacts criterion in your use of the merit review criteria?"

Note: In this exhibit, we show the percentage of NSF staff indicating the level of weight they typically place on the importance of the Intellectual Merit criterion compared to the Broader Impacts criterion in their use of the merit review criteria. "More weight on Intellectual Merit" includes the following response options: "Much more weight on Intellectual Merit," "More weight on Intellectual Merit," and "Somewhat more weight on Intellectual Merit." "More weight on Broader Impacts" includes these response options: "Somewhat more weight on Broader Impacts," "More weight on Broader Impacts," and "Much more weight on Broader Impacts." NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included.

Exhibit 2.7. Percentage of NSF staff who indicated that they do or do not find it challenging to assess a proposal's merit against each criterion (overall, by position type, and by directorate or office)

Respondent characteristic	Intellectual Merit criterion		Broader Impacts criterion		Responses (N)
	Challenging	Not challenging	Challenging	Not challenging	
Overall					
All respondents	18%	82%	39%	61%	233
Position type					
Program Officer	17%	83%	38%	62%	193
Deputy Division Director	18%	82%	47%	53%	17
Division Director	29%	71%	43%	57%	7
Deputy Assistant Director and Deputy Office Head	40%	60%	40%	60%	5
Other NSF staff position	18%	82%	36%	64%	11
p-value		0.685		0.966	
Directorate or office					
BIO	14%	86%	20%	80%	35
CISE	33%	67%	53%	47%	15
ENG	24%	76%	44%	56%	25
GEO	13%	87%	45%	55%	31
MPS	16%	84%	50%	50%	38
SBE	20%	80%	27%	73%	15
EDU	15%	85%	35%	65%	26
TIP	25%	75%	63%	38%	8
OD/OIA	0%	100%	20%	80%	5
OISE	D	D	D	D	D
Not reported	D	D	D	D	D
p-value		0.824		0.186	

Source: Merit Review Examination Staff Survey, question 16: "How challenging do you generally find it to assess a proposal's merit against each criteria?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff indicating that it is challenging or not challenging to assess a proposal's merit against the Intellectual Merit criterion or the Broader Impacts criterion. Along with overall results, we show responses by position type and by directorate or office. "Challenging" includes the following response options: "Somewhat challenging" and "Very challenging." "Not challenging" includes the following response options: "Not at all challenging" and "A little challenging." NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type or by directorate or office; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 2.8. Percentage of NSF staff who indicated that they have had recurrent challenges assessing a proposal's merit against each criterion (overall, by position type, and by directorate or office)

Respondent characteristic	Percentage	Responses (N)
Overall		
All respondents	28%	232
Position type		
Program Officer	29%	192
Deputy Division Director	6%	17
Division Director	14%	7
Deputy Assistant Director and Deputy Office Head	40%	5
Other NSF staff position	45%	11
p-value	0.142	
Directorate or office		
BIO	31%	35
CISE	21%	14
ENG	32%	25
GEO	39%	31
MPS	21%	38
SBE	20%	15
EDU	27%	26
TIP	25%	8
OD/OIA	0%	5
OISE	D	D
Not reported	D	D
p-value	0.579	

Source: Merit Review Examination Staff Survey, question 17: "Have you had any recurrent challenges assessing a proposal's merit against each criterion?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff selecting "Yes," indicating they have had recurrent challenges assessing a proposal's merit against each criterion. Along with overall results, we show responses by position type and by directorate or office. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type or by directorate or office; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 2.9. Percentage of NSF staff who indicated that there is something about the scientific field(s) of their directorate, division, or program that makes it challenging to apply the Intellectual Merit criterion or the Broader Impacts criterion (overall, by position type, and by directorate or office)

Respondent characteristic	Intellectual Merit	Broader Impacts	Responses (N)
Overall			
All respondents	24%	28%	233
Position type			
Program Officer	23%	26%	193
Deputy Division Director	35%	29%	17
Division Director	14%	43%	7
Deputy Assistant Director and Deputy Office Head	20%	20%	5
Other NSF staff position	18%	45%	11
p-value	0.763	0.562	
Directorate or office			
BIO	20%	15%	35
CISE	20%	47%	15
ENG	16%	24%	25
GEO	10%	29%	31
MPS	13%	29%	38
SBE	20%	20%	15
EDU	46%	31%	26
TIP	75%	50%	8
OD/OIA	40%	40%	5
OISE	D	D	D
Not reported	D	D	D
p-value	0.001**	0.518	

Source: Merit Review Examination Staff Survey, question 19: "In your opinion, is there anything about the scientific field(s) of your directorate, division, or program that makes it challenging to apply the Intellectual Merit criterion?"; Merit Review Examination Staff Survey, question 21: "In your opinion, is there anything about the scientific field(s) of your directorate, division, or program that makes it challenging to apply the Broader Impacts criterion?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff selecting "Yes," indicating they believe something about the scientific field(s) of their directorate, division, or program makes it challenging to apply the Intellectual Merit criterion or Broader Impacts criterion. Along with overall results, we show responses by position type and by directorate or office. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type or by directorate or office; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 2.10. Number of RFI responses contributing to themes on how NSF could improve implementation of the merit review process

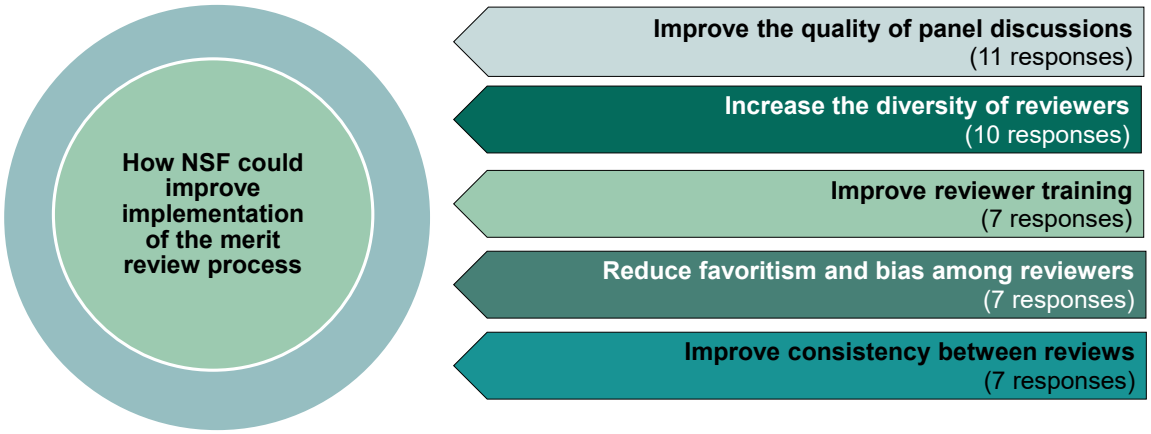
Theme	Number of responses
Improve the quality of panel discussions	11
Increase the diversity of reviewers	10
Improve reviewer training	7
Reduce favoritism and bias among reviewers	7
Improve consistency between reviews	7

Source: Merit Review Examination RFI, question 4, "MRX is interested in the experiences and perspectives of those who have reviewed proposals submitted to NSF. We invite you to share your insights and describe any opportunities you believe would improve implementation of the Merit Review criteria, policy, or processes based on your experience reviewing NSF proposals."

Note: In this exhibit, we show themes from the RFI responses on how NSF could improve implementation of the merit review process and the number of responses contributing to each theme.

RFI = Request for Information.

Exhibit 2.11. Themes from RFI responses on how NSF could improve implementation of the merit review process



Source: Merit Review Examination RFI, question 4, "MRX is interested in the experiences and perspectives of those who have reviewed proposals submitted to NSF. We invite you to share your insights and describe any opportunities you believe would improve implementation of the Merit Review criteria, policy, or processes based on your experience reviewing NSF proposals."

Note: In this exhibit, we show themes from the RFI responses on how NSF could improve implementation of the merit review process and the number of responses contributing to each theme.

RFI = Request for Information.

4. How do PIs, reviewers, and NSF leaders and staff understand and interpret the merit review principles and elements?

Exhibit 4.1. Percentage of NSF staff who indicated a low, moderate, or high level of understanding of the merit review principles for themselves and for Program Officers, NSF leadership (Division Directors, Deputy Division Directors, and Section Heads), reviewers, and principal investigators (overall, by position type, and by directorate or office)

Respondent characteristic	Themselves			Program Officers			Division Directors, Deputy Division Directors, and Section Heads			Reviewers			Principal investigators			Responses (N)
	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	
Overall																
All respondents	0%	5%	95%	1%	9%	91%	6%	8%	86%	8%	36%	55%	13%	41%	46%	226
Position type																
Program Officer	0%	4%	96%	1%	9%	91%	6%	7%	87%	8%	35%	57%	12%	41%	48%	188
Deputy Division Director	0%	6%	94%	6%	0%	94%	6%	0%	94%	19%	19%	63%	19%	38%	44%	16
Division Director	0%	29%	71%	0%	29%	71%	0%	43%	57%	14%	71%	14%	29%	57%	14%	7
Deputy Assistant Director and Deputy Office Head	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	40%	60%	0%	60%	40%	5
Other NSF staff position	0%	10%	90%	0%	10%	90%	10%	20%	70%	10%	50%	40%	20%	40%	40%	10
p-value			0.071			0.202			0.036*			0.261			0.690	
Directorate or office																
BIO	0%	0%	100%	0%	0%	100%	3%	6%	91%	0%	43%	57%	3%	51%	46%	35
CISE	0%	7%	93%	7%	0%	93%	7%	0%	93%	20%	20%	60%	21%	21%	57%	15
ENG	0%	4%	96%	0%	12%	88%	0%	8%	92%	4%	40%	56%	12%	36%	52%	25
GEO	0%	7%	93%	0%	7%	93%	0%	14%	86%	10%	34%	55%	14%	31%	55%	30
MPS	0%	3%	97%	0%	5%	95%	8%	0%	92%	8%	29%	63%	11%	34%	55%	38
SBE	0%	0%	100%	0%	0%	100%	0%	0%	100%	7%	27%	67%	7%	47%	47%	15

Respondent characteristic	Themselves			Program Officers			Division Directors, Deputy Division Directors, and Section Heads			Reviewers			Principal investigators			Responses (N)
	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	
EDU	0%	12%	88%	4%	24%	72%	17%	21%	63%	15%	35%	50%	23%	38%	38%	26
TIP	0%	13%	88%	0%	25%	75%	0%	29%	71%	25%	13%	63%	25%	50%	25%	8
OD/OIA	0%	20%	80%	0%	20%	80%	20%	20%	60%	20%	80%	0%	40%	60%	0%	5
OISE	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
Not reported	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
p-value			0.550			0.089			0.041*			0.195			0.225	

Source: Merit Review Examination Staff Survey, question 31: "How would you rate the overall level of understanding of the merit review principles by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff indicating each level of understanding of the merit review principles for themselves (you) and for Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators. Along with overall results, we show responses by position type, and by directorate or office. "Low" includes the response options "Low level" and "Very low level." "Moderate" includes the response option "Moderate level." "High" includes the response options "High level" and "Very high level." The response option "No basis to rate" was omitted from the table; percentages represent respondents who provided a rating. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests are used to assess whether response patterns for each item differ by position type or by directorate or office; *p*-values from these tests are reported.

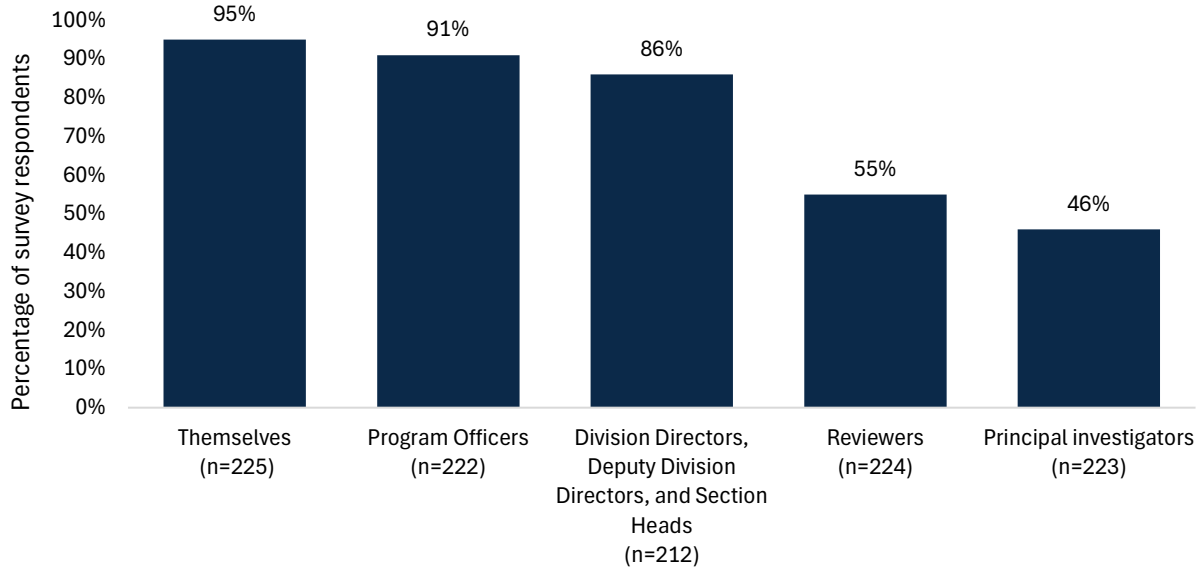
* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 4.2. Percentage of NSF staff who indicated a high or very high level of understanding of the merit review principles for themselves and for Program Officers, NSF leadership (Division Directors, Deputy Division Directors, and Section Heads), reviewers, and principal investigators



Source: Merit Review Examination Staff Survey, question 31: "How would you rate the overall level of understanding of the merit review principles by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators?"

Note: In this exhibit, we show the percentage of NSF staff indicating a high or very high level of understanding of the merit review principles for themselves and for Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators. The response option "No basis to rate" was omitted from the figure; percentages represent respondents who provided a rating. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included.

Exhibit 4.3. Percentage of NSF staff who indicated a low, moderate, or high level of understanding of the merit review elements for themselves and for Program Officers, NSF leadership (Division Directors, Deputy Division Directors, Section Heads), reviewers, and principal investigators (overall, by position type, and by directorate or office)

Respondent characteristic	Themselves			Program Officers			Division Directors, Deputy Division Directors, and Section Heads			Reviewers			Principal investigators			Responses (N)
	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	
Overall																
All respondents	<1%	4%	96%	1%	6%	93%	3%	10%	86%	7%	28%	65%	11%	38%	51%	225
Position type																
Program Officer	0%	4%	96%	1%	7%	92%	3%	11%	86%	7%	29%	64%	11%	39%	51%	187
Deputy Division Director	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	19%	81%	6%	38%	56%	16
Division Director	14%	0%	86%	14%	0%	86%	14%	14%	71%	29%	29%	43%	29%	43%	29%	7
Deputy Assistant Director and Deputy Office Head	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	40%	60%	0%	60%	40%	5
Other NSF staff position	0%	10%	90%	0%	10%	90%	10%	20%	70%	10%	20%	70%	20%	10%	70%	10
p-value			<.001**			0.032*			0.370			0.400			0.446	
Directorate or office																
BIO	0%	0%	100%	0%	0%	100%	3%	6%	91%	9%	26%	66%	11%	34%	54%	35
CISE	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	14%	86%	0%	43%	57%	15
ENG	0%	4%	96%	0%	4%	96%	0%	8%	92%	0%	32%	68%	4%	40%	56%	25
GEO	3%	3%	93%	3%	3%	93%	4%	18%	79%	14%	17%	69%	10%	31%	59%	30
MPS	0%	5%	95%	0%	8%	92%	6%	6%	89%	5%	26%	68%	11%	29%	61%	38
SBE	0%	7%	93%	0%	7%	93%	0%	7%	93%	7%	27%	67%	7%	47%	47%	15
EDU	0%	8%	92%	4%	19%	77%	4%	29%	67%	12%	31%	58%	23%	35%	42%	26
TIP	0%	13%	88%	0%	13%	88%	0%	14%	86%	13%	25%	63%	25%	50%	25%	8

Respondent characteristic	Themselves			Program Officers			Division Directors, Deputy Division Directors, and Section Heads			Reviewers			Principal investigators			Responses (N)
	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	
OD/OIA	0%	0%	100%	0%	0%	100%	0%	20%	80%	20%	40%	40%	20%	60%	20%	5
OISE	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
Not reported	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
p-value			0.917			0.526			0.381			0.721			0.637	

Source: Merit Review Examination Staff Survey, question 32: "How would you rate the overall level of understanding of the merit review elements by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff indicating a low, moderate, or high level of understanding of the merit review elements for themselves (you) and for Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators. Along with overall results, we show responses by position type and by directorate or office. "Low" includes the response options "Low level" and "Very low level." "Moderate" includes the response option "Moderate level." "High" includes the response options "High level" and "Very high level." The response option "No basis to rate" was omitted from the table; percentages represent respondents who provided a rating. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type or by directorate or office; *p*-values from these tests are reported.

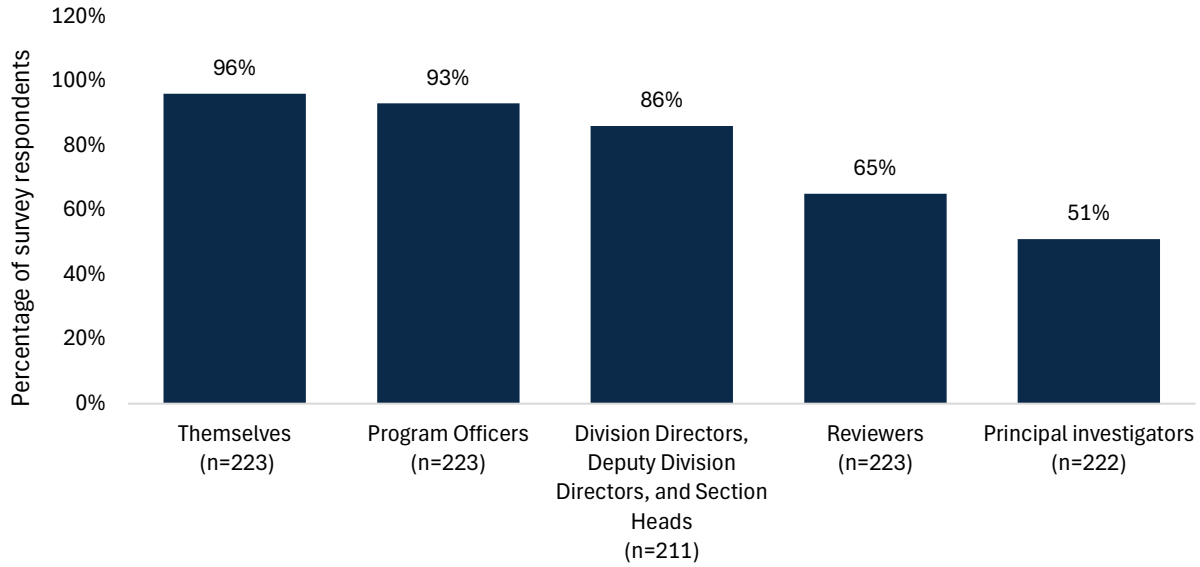
* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 4.4. Percentage of NSF staff who indicated a high or very high level of understanding of the merit review elements for themselves and for Program Officers, NSF leadership (Division Directors, Deputy Division Directors, Section Heads), reviewers, and principal investigators



Source: Merit Review Examination Staff Survey, question 32: "How would you rate the overall level of understanding of the merit review elements by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators?"

Note: In this exhibit, we show the percentage of NSF staff indicating a "high" or "very high" level of understanding of the merit review elements for themselves and for Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators. The response option "No basis to rate" was omitted from the figure; percentages represent respondents who provided a rating. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included.

5. What additional merit review criteria might NSF consider using in the merit review process to better achieve its mission to invest in research that advances scientific knowledge and benefits society?

Exhibit 5.1. Percentage of NSF staff who indicated that there are other important factors for evaluating proposals that the merit review criteria do not capture (overall, by position type, and by directorate or office)

Respondent characteristic	Percentage	Responses (N)
Overall		
All respondents	49%	231
Position type		
Program Officer	46%	191
Deputy Division Director	59%	17
Division Director	100%	7
Deputy Assistant Director and Deputy Office Head	40%	5
Other NSF staff position	64%	11
p-value	0.041*	
Directorate or office		
BIO	41%	34
CISE	67%	15
ENG	36%	25
GEO	55%	31
MPS	39%	38
SBE	40%	15
EDU	65%	26
TIP	75%	8
OD/OIA	80%	5
OISE	D	D
Not reported	D	D
p-value	0.144	

Source: Merit Review Examination Staff Survey, question 24: "Based on your experience, are there any factors important for evaluating proposals that are not captured by the two merit review criteria?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff indicating "Yes," important factors for evaluating proposals are not captured by the two merit review criteria. Along with overall results, we show responses by position type and by directorate or office. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type or by directorate and office; *p*-values from these tests are reported.

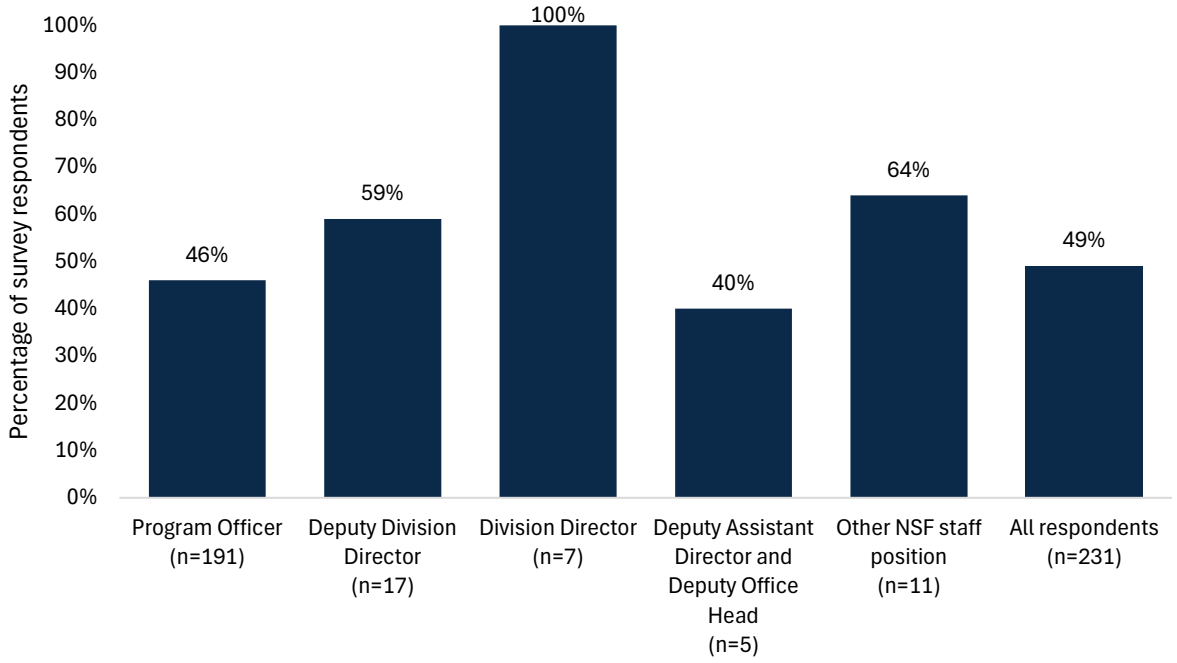
* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 5.2. Percentage of NSF staff who indicated that there are other important factors for evaluating proposals that the merit review criteria do not capture (overall and by position type)



Source: Merit Review Examination Staff Survey, question 24: "Based on your experience, are there any factors important for evaluating proposals that are not captured by the two merit review criteria?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"

Note: In this exhibit, we show the percentage of NSF staff indicating "Yes," important factors for evaluating proposals are not captured by the two merit review criteria. Along with overall results, we show responses by position type. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included.

6. How is reviewer and staff expertise in the merit review criteria assessed?

Exhibit 6.1. Percentage of NSF staff who indicated that they or their unit are somewhat or very effective at identifying which reviewers have the necessary expertise in the merit review criteria (overall, by position type, and by directorate or office)

Respondent characteristic	Percentage	Responses (N)
Overall		
All respondents	98%	228
Position type		
Program Officer	97%	190
Deputy Division Director	100%	17
Division Director	100%	7
Deputy Assistant Director and Deputy Office Head	100%	5
Other NSF staff position	100%	9
p-value	0.906	
Directorate or office		
BIO	100%	34
CISE	100%	15
ENG	92%	25
GEO	100%	31
MPS	97%	37
SBE	100%	14
EDU	100%	26
TIP	100%	8
OD/OIA	100%	5
OISE	D	D
Not reported	D	D
p-value	0.491	

Source: Merit Review Examination Staff Survey, question 26: "How effective do you think [you are/your unit is] at identifying which reviewers have the necessary expertise of the merit review criteria?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff indicating that they or their unit are "Somewhat effective" or "Very effective" at identifying which reviewers have the necessary expertise in the merit review criteria. Along with overall results, we show responses by position type and by directorate or office. For Program Officers, the question fill was "you are;" for all other NSF staff, the question fill was "your unit." NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type or by directorate or office; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 6.2. Percentage of NSF staff who indicated a low, moderate, or high level of understanding of the merit review criteria for themselves and for Program Officers, NSF leadership (Division Directors, Deputy Division Directors, Section Heads), reviewers, and principal investigators (overall, by position type, and by directorate or office)

Respondent characteristic	Themselves			Program Officers			Division Directors, Deputy Division Directors, and Section Heads			Reviewers			Principal Investigators			Responses (N)
	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	
Overall																
All respondents	0%	2%	98%	<1%	6%	93%	3%	9%	88%	3%	33%	63%	6%	45%	49%	230
Position type																
Program Officer	0%	1%	99%	1%	6%	94%	4%	8%	88%	4%	32%	65%	5%	45%	50%	191
Deputy Division Director	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	47%	53%	0%	59%	41%	17
Division Director	0%	29%	71%	0%	29%	71%	0%	29%	71%	0%	57%	43%	14%	57%	29%	7
Deputy Assistant Director and Deputy Office Head	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	40%	60%	0%	60%	40%	5
Other NSF staff position	0%	10%	90%	0%	10%	90%	0%	30%	70%	10%	20%	70%	20%	20%	60%	10
p-value			<.001**			0.433			0.148			0.639			0.301	
Directorate or office																
BIO	0%	0%	100%	0%	0%	100%	3%	3%	94%	0%	23%	77%	3%	34%	63%	35
CISE	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	27%	73%	0%	47%	53%	15
ENG	0%	0%	100%	0%	4%	96%	0%	4%	96%	4%	28%	68%	4%	48%	48%	25
GEO	0%	3%	97%	0%	7%	93%	0%	17%	83%	7%	43%	50%	7%	37%	57%	31
MPS	0%	0%	100%	0%	3%	97%	5%	3%	92%	3%	26%	71%	5%	42%	53%	38
SBE	0%	0%	100%	0%	7%	93%	0%	0%	100%	0%	27%	73%	0%	57%	43%	15
EDU	0%	8%	92%	4%	23%	73%	8%	32%	60%	8%	27%	65%	8%	46%	46%	26
TIP	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	50%	50%	13%	63%	25%	8
OD/OIA	0%	0%	100%	0%	25%	75%	20%	20%	60%	20%	40%	40%	20%	60%	20%	5

Respondent characteristic	Themselves			Program Officers			Division Directors, Deputy Division Directors, and Section Heads			Reviewers			Principal Investigators			Responses (N)
	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	Low	Moderate	High	
OISE	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
Not reported	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D	D
p-value			0.564			0.101			0.009**			0.258			0.722	

Source: Merit Review Examination Staff Survey, question 30: "How would you rate the overall level of understanding of the merit review criteria by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff indicating a low, moderate, or high level of understanding of the merit review criteria for themselves (you), Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators. Along with overall results, we show responses by position and by directorate or office. "Low" includes the response options "Low level" and "Very low level." "Moderate" includes the response option "Moderate level." "High" includes the response options "High level" and "Very high level." The response option "No basis to rate" was omitted from the table; percentages represent respondents who provided a rating. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type or by directorate or office; *p*-values from these tests are reported.

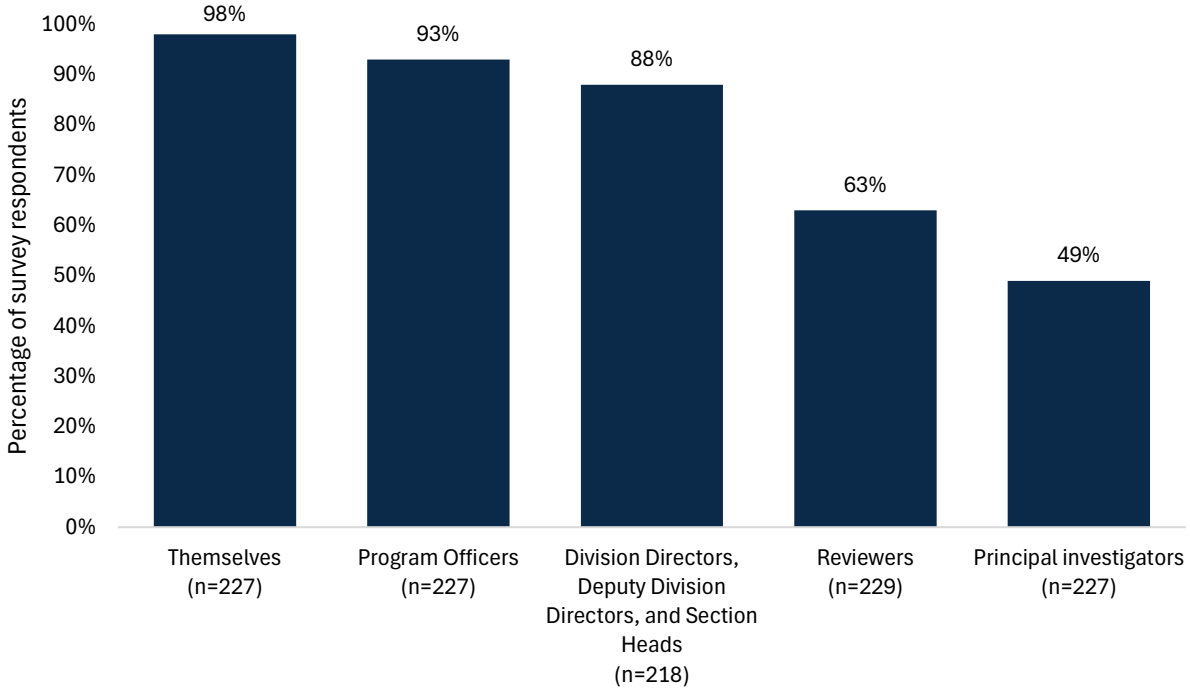
* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 6.3. Percentage of NSF staff who indicated a high or very high level of understanding of the merit review criteria for themselves and for Program Officers, NSF leadership (Division Directors, Deputy Division Directors, Section Heads), reviewers, and principal investigators



Source: Merit Review Examination Staff Survey, question 30: "How would you rate the overall level of understanding of the merit review criteria by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators?"

Note: In this exhibit, we show the percentage of NSF staff indicating a "high" or" very high" level of understanding of the merit review criteria for themselves (you), Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators. The response option "No basis to rate" was omitted from the figure; percentages represent respondents who provided a rating. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included.

9. How do reviewer assessments of each merit review criterion factor into NSF Program Officers' award recommendations?

Exhibit 9.1. Percentage of NSF Program Officers who indicated that, for their division, reviewer assessments of the Intellectual Merit criterion or the Broader Impacts criterion factor into funding recommendations somewhat or to a great extent (overall and by directorate or office)

Respondent characteristic	Intellectual Merit	Broader Impacts	Responses (N)
Overall			
All Program Officer respondents	99%	95%	192
Directorate or office			
BIO	100%	100%	31
CISE	100%	92%	13
ENG	100%	95%	21
GEO	100%	92%	26
MPS	97%	94%	33
SBE	100%	91%	11
EDU	100%	100%	23
TIP	100%	88%	8
OD/OIA	D	D	D
OISE	D	D	D
Not reported	D	D	D
p-value	0.868	0.848	

Source: Merit Review Examination Staff Survey, question 28: "How much do reviewers' assessments of each merit review criterion factor into [funding recommendations/portfolio management] within your division?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF Program Officers indicating that reviewer assessments of each merit review criterion factor into their funding recommendations "Somewhat" or "To a great extent." Along with overall results, we show responses by directorate or office. For Program Officers, the question fill was "funding recommendations." NSF Program Officers who have had responsibility for merit review in the last three years are included. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by directorate or office; *p*-values from these tests are reported.

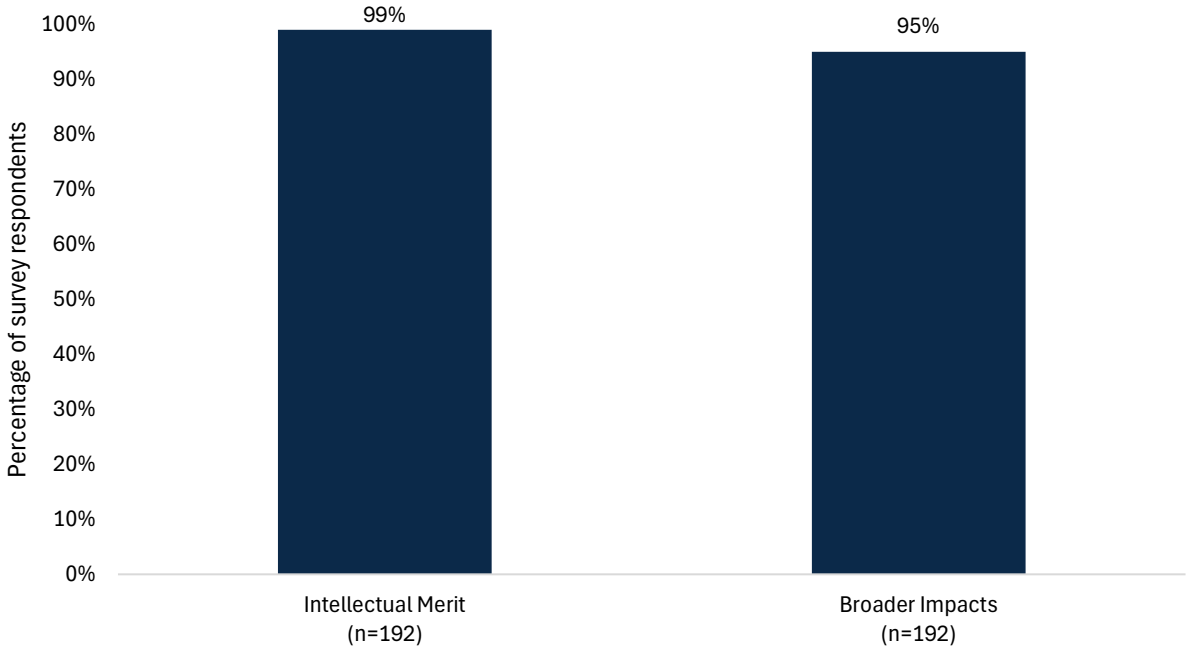
* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 9.2. Percentage of NSF Program Officers who indicated that, for their division, reviewer assessments of the Intellectual Merit criterion or the Broader Impacts criterion factor into funding recommendations somewhat or to a great extent



Source: Merit Review Examination Staff Survey, question 28: "How much do reviewers' assessments of each merit review criterion factor into [funding recommendations/portfolio management] within your division?"

Note: In this exhibit, we show the percentage of NSF Program Officers indicating that reviewer assessments of each merit review criterion factor into their funding recommendations "Somewhat" or "To a great extent." For Program Officers, the question fill was "funding recommendations." NSF Program Officers who have had responsibility for merit review in the last three years are included. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item.

10. How do reviewers' assessments of each merit review criterion factor into NSF division directors' award decisions for individual proposals? At the portfolio level?

Exhibit 10.1. Percentage of Division Directors, Deputy Division Directors, Deputy Assistant Directors, and Deputy Office Heads who indicated that, for their division, reviewers' assessments of Intellectual Merit criterion and or Broader Impacts criterion are factored into their portfolio management somewhat or to a great extent (overall and by position type)

Respondent characteristic	Intellectual Merit	Broader Impacts	Responses (N)
Overall			
All Division Director, Deputy Division Director, Deputy Assistant Director, and Deputy Office Head respondents	93%	93%	29
Position type			
Deputy Division Director	88%	88%	17
Division Director	100%	100%	7
Deputy Assistant Director and Deputy Office Head	100%	100%	5
p-value	0.469	0.469	

Source: Merit Review Examination Staff Survey, question 28: "How much do reviewers' assessments of each merit review criterion factor into [funding recommendations/portfolio management] within your division?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"

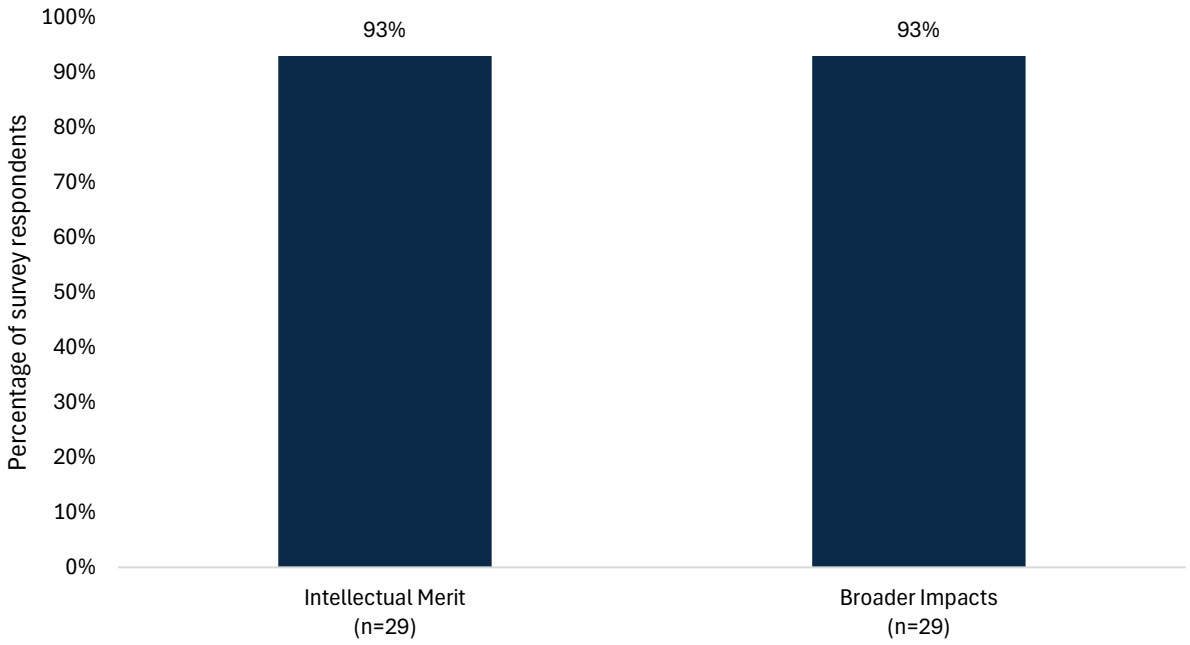
Note: In this exhibit, we show the percentage of NSF Division Directors, Deputy Division Directors, Deputy Assistant Directors, and Deputy Office Heads indicating that reviewer assessments of each merit review criterion factor into their portfolio management "Somewhat" or "To a great extent." Along with overall results, we show responses by position type. For these NSF staff, the question fill was "portfolio management." NSF Division Directors, Deputy Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years are included. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

Exhibit 10.2. Percentage of Division Directors, Deputy Division Directors, Deputy Assistant Directors, and Deputy Office Heads who indicated that, for their division, reviewers' assessments of Intellectual Merit criterion and or Broader Impacts criterion are factored into their portfolio management somewhat or to a great extent



Source: Merit Review Examination Staff Survey, question 28: "How much do reviewers' assessments of each merit review criterion factor into [funding recommendations/portfolio management] within your division?"

Note: In this exhibit, we show the percentage of NSF Division Directors, Deputy Division Directors, Deputy Assistant Directors, and Deputy Office Heads indicating that reviewer assessments of each merit review criterion factor into their portfolio management "Somewhat" or "To a great extent." For these NSF staff, the question fill was "portfolio management." NSF Division Directors, Deputy Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years are included. Intergovernmental Personnel Act (IPA) staff are not included.

12. To what extent do NSF constituencies perceive the merit review policy and processes to be unfairly biased? How do these perceptions differ by constituency?

Exhibit 12.1. Percentage of NSF staff who indicated that they agree or strongly agree that individuals submitting proposals are treated fairly and that the merit review criteria support a fair and accurate assessment of a proposal's merit (overall and by position type, directorate or office, gender, race, ethnicity, and disability)

Respondent characteristic	Individuals submitting proposals are treated fairly		The merit review criteria support a fair and accurate assessment of a proposal's merit	
	Percentage	Responses (N)	Percentage	Responses (N)
Overall				
All respondents	87%	225	85%	225
Position type				
Program Officer	87%	187	85%	187
Deputy Division Director	100%	16	88%	16
Division Director	86%	7	86%	7
Deputy Assistant Director and Deputy Office Head	100%	5	100%	5
Other NSF staff position	60%	10	70%	10
p-value		0.046*		0.608
Directorate or office				
BIO	97%	34	94%	35
CISE	100%	15	100%	15
ENG	88%	25	80%	25
GEO	90%	31	93%	30
MPS	95%	38	92%	38
SBE	93%	15	93%	15
EDU	81%	26	81%	26
TIP	50%	8	38%	8
OD/OIA	20%	5	40%	5
OISE	D	D	D	D
Not reported	D	D	D	D
p-value		<.001**		<.001**
Gender				
Female only	83%	92	83%	92
Male only	93%	91	89%	91
Other category and not reported	83%	42	81%	42
p-value		0.067		0.352

Respondent characteristic	Individuals submitting proposals are treated fairly		The merit review criteria support a fair and accurate assessment of a proposal's merit	
	Percentage	Responses (N)	Percentage	Responses (N)
Race				
American Indian or Alaska Native	NA	0	NA	0
Asian	88%	16	88%	16
Black or African American	60%	15	73%	15
Native Hawaiian or Other Pacific Islander	NA	0	NA	0
White	90%	136	88%	136
Other or multiple racial categories selected	83%	6	67%	6
Not reported	87%	52	83%	52
p-value		0.024*		0.390
Ethnicity				
Hispanic or Latino	100%	13	100%	13
Not Hispanic or Latino	87%	165	85%	165
Not reported	85%	47	81%	47
p-value		0.346		0.233
Disability				
Identifies as having a disability	79%	14	79%	14
Does not identify as having a disability	88%	40	85%	40
Not reported	88%	171	85%	171
p-value		0.615		0.791

Source: Merit Review Examination Staff Survey, question 33: "Based on your experience with the merit review process, to what extent do you agree or disagree with the following statements? 'Individuals submitting proposals are treated fairly,' and 'The merit review criteria support a fair and accurate assessment of a proposal's merit'"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"; Merit Review Examination Staff Survey, question 47: "How do you currently describe yourself?"; Merit Review Examination Staff Survey, question 48: "Are you Hispanic or Latino?"; Merit Review Examination Staff Survey, question 49: "What is your racial background?"; Merit Review Examination Staff Survey, question 50: "Do you identify as having a disability?"

Note: In this exhibit, we show the percentage of NSF staff indicating that they "Agree" or "Strongly agree" that individuals submitting proposals are treated fairly and that the merit review criteria support a fair and accurate assessment of a proposal's merit. Along with overall results, we show responses by position type, directorate or office, gender, race, ethnicity, and disability. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type, directorate or office, gender, race, ethnicity, or disability; *p*-values from these tests are reported.

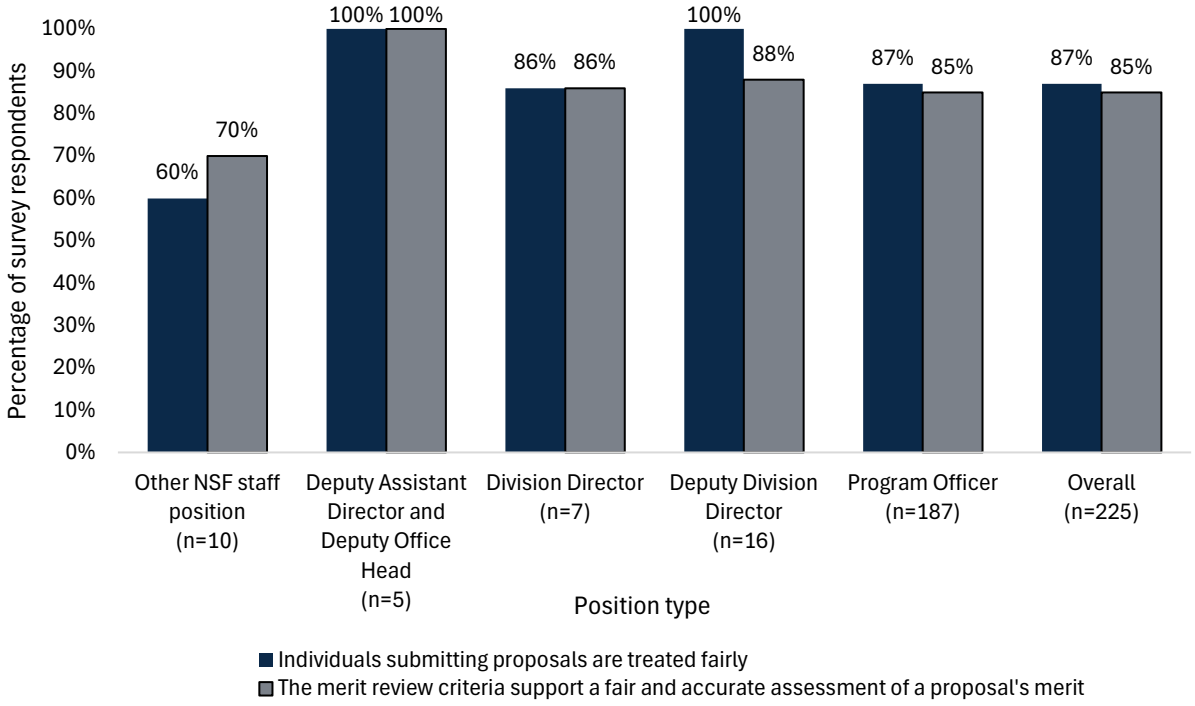
* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS= Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 12.2. Percentage of NSF staff who indicated that they agree or strongly agree that individuals submitting proposals are treated fairly and that the merit review criteria support a fair and accurate assessment of a proposal's merit, by position type



Source: Merit Review Examination Staff Survey, question 33: "Based on your experience with the merit review process, to what extent do you agree or disagree with the following statements? 'Individuals submitting proposals are treated fairly,' and 'The merit review criteria support a fair and accurate assessment of a proposal's merit."; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"

Note: In this exhibit, we show the percentage of NSF staff indicating that they "Agree" or "Strongly agree" that individuals submitting proposals are treated fairly and that the merit review criteria support a fair and accurate assessment of a proposal's merit. Along with overall results, we show responses by position type. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included.

Exhibit 12.3. Percentage of NSF staff who indicated that the merit review process decreases, neither increases nor decreases, or increases the diversity of ideas, principal investigators, or institutions *represented in proposals submitted for review* (overall and by position type, directorate or office, gender, race, ethnicity, and disability)

Respondent characteristic	Diversity of ideas			Diversity of principal investigators			Diversity of institutions			Responses (N)
	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	
Overall										
All respondents	26%	33%	41%	33%	40%	26%	41%	36%	24%	220
Position type										
Program Officer	26%	31%	42%	33%	40%	28%	41%	35%	24%	182
Deputy Division Director	19%	38%	44%	38%	31%	31%	38%	31%	31%	16
Division Director	57%	29%	14%	57%	43%	0%	71%	29%	0%	7
Deputy Assistant Director and Deputy Office Head	0%	60%	40%	0%	60%	40%	0%	60%	40%	5
Other NSF staff position	30%	40%	30%	40%	50%	10%	40%	50%	10%	10
p-value			0.493			0.471			0.363	
Directorate or office										
BIO	18%	38%	44%	30%	30%	39%	42%	24%	33%	34
CISE	20%	20%	60%	27%	40%	33%	40%	27%	33%	15
ENG	28%	28%	44%	32%	56%	12%	40%	48%	12%	25
GEO	23%	39%	39%	42%	42%	16%	45%	42%	13%	31
MPS	25%	28%	47%	19%	42%	39%	28%	36%	36%	36
SBE	36%	14%	50%	29%	36%	36%	29%	36%	36%	14
EDU	40%	28%	32%	40%	36%	24%	48%	32%	20%	25
TIP	38%	25%	38%	50%	13%	38%	63%	13%	25%	8
OD/OIA	60%	20%	20%	80%	20%	0%	80%	20%	0%	5

Respondent characteristic	Diversity of ideas			Diversity of principal investigators			Diversity of institutions			Responses (N)
	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	
OISE	D	D	D	D	D	D	D	D	D	D
Not reported	D	D	D	D	D	D	D	D	D	D
p-value			0.423			0.192			0.304	
Gender										
Female only	30%	30%	40%	38%	39%	22%	42%	38%	20%	90
Male only	20%	34%	46%	25%	43%	33%	36%	36%	28%	89
Other category and not reported	32%	37%	32%	41%	37%	22%	49%	29%	22%	41
p-value			0.402			0.213			0.553	
Race										
American Indian or Alaska Native	NA	NA	NA	NA	NA	NA	NA	NA	NA	0
Asian	13%	38%	50%	38%	19%	44%	44%	13%	44%	16
Black or African American	40%	20%	40%	60%	40%	0%	60%	33%	7%	15
Native Hawaiian or Other Pacific Islander	NA	NA	NA	NA	NA	NA	NA	NA	NA	0
White	26%	32%	41%	30%	44%	26%	39%	40%	21%	133
Other or multiple racial categories selected	33%	17%	50%	50%	17%	33%	50%	17%	33%	6
Not reported	26%	38%	36%	30%	40%	30%	38%	34%	28%	50
p-value			0.774			0.094			0.196	
Ethnicity										
Hispanic or Latino	8%	50%	42%	17%	50%	33%	17%	50%	33%	12
Not Hispanic or Latino	27%	29%	44%	35%	39%	26%	43%	35%	22%	162
Not reported	28%	41%	30%	33%	41%	26%	39%	35%	26%	46
p-value			0.202			0.794			0.500	

Respondent characteristic	Diversity of ideas			Diversity of principal investigators			Diversity of institutions			Responses (N)
	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	
Disability										
Identifies as having a disability	29%	43%	29%	43%	29%	29%	43%	36%	21%	14
Does not identify as having a disability	26%	42%	32%	26%	47%	26%	34%	42%	24%	38
Not reported	26%	30%	44%	34%	40%	26%	42%	34%	24%	168
p-value			0.452			0.740			0.901	

Source: Merit Review Examination Staff Survey, question 36: "How much does the merit review process increase or decrease the diversity of ideas, principal investigators, and institutions represented in proposals submitted for review?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"; Merit Review Examination Staff Survey, question 47: "How do you currently describe yourself?"; Merit Review Examination Staff Survey, question 48: "Are you Hispanic or Latino?; Merit Review Examination Staff Survey, question 49: "What is your racial background?"; Merit Review Examination Staff Survey, question 50: "Do you identify as having a disability?"

Note: In this exhibit, we show the percentage of NSF staff who indicated whether the merit review process increases, decreases, or neither increases nor decreases the diversity of ideas, principal investigators, or institutions represented in proposals submitted for review. Along with overall results, we show responses by position type, directorate or office, gender, race, ethnicity, and disability. "Increases" includes the response options "Slightly increases," "Somewhat increases," and "Greatly increases." "Neither decreases nor increases" includes the response option "Neither decreases nor increases." "Decreases" includes the response options "Greatly decreases," "Somewhat decreases," and "Slightly decreases." NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type, directorate or office, gender, race, ethnicity, or disability; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 12.4. Percentage of NSF staff who indicated that the merit review process decreases, neither increases nor decreases or increases the diversity of ideas, principal investigators, or institutions *represented in awarded projects* (overall and by position type, directorate or office, gender, race, ethnicity, and disability)

Respondent characteristic	Diversity of ideas			Diversity of principal investigators			Diversity of institutions			Responses (N)
	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	
Overall										
All respondents	22%	25%	52%	26%	25%	49%	31%	25%	45%	219
Position type										
Program Officer	22%	24%	54%	24%	24%	52%	28%	25%	48%	181
Deputy Division Director	19%	19%	63%	31%	19%	50%	38%	19%	44%	16
Division Director	57%	29%	14%	43%	57%	0%	71%	29%	0%	7
Deputy Assistant Director and Deputy Office Head	20%	20%	60%	0%	40%	60%	20%	20%	60%	5
Other NSF staff position	20%	50%	30%	40%	30%	30%	50%	30%	20%	10
p-value			0.257			0.171			0.175	
Directorate or office										
BIO	12%	18%	70%	9%	12%	79%	9%	21%	71%	34
CISE	20%	13%	67%	33%	13%	53%	40%	7%	53%	15
ENG	16%	20%	64%	24%	32%	44%	32%	32%	36%	25
GEO	23%	30%	47%	37%	20%	43%	37%	27%	37%	30
MPS	14%	28%	58%	11%	31%	58%	17%	22%	61%	36
SBE	21%	36%	43%	36%	21%	43%	36%	21%	43%	14
EDU	44%	20%	36%	40%	32%	28%	52%	28%	20%	25
TIP	50%	25%	25%	38%	38%	25%	50%	25%	25%	8
OD/OIA	40%	20%	40%	40%	0%	60%	60%	20%	20%	5

Respondent characteristic	Diversity of ideas			Diversity of principal investigators			Diversity of institutions			Responses (N)
	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	
OISE	D	D	D	D	D	D	D	D	D	D
Not reported	D	D	D	D	D	D	D	D	D	D
p-value			0.204			0.030*			0.028*	
Gender										
Female only	26%	25%	49%	30%	24%	46%	33%	25%	43%	89
Male only	18%	22%	60%	19%	28%	53%	28%	26%	46%	89
Other category and not reported	24%	32%	44%	29%	22%	49%	32%	22%	46%	41
p-value			0.407			0.487			0.960	
Race										
American Indian or Alaska Native	NA	NA	NA	NA	NA	NA	NA	NA	NA	0
Asian	6%	44%	50%	19%	31%	50%	31%	25%	44%	16
Black or African American	20%	27%	53%	40%	20%	40%	47%	27%	27%	15
Native Hawaiian or Other Pacific Islander	NA	NA	NA	NA	NA	NA	NA	NA	NA	0
White	25%	20%	55%	25%	25%	51%	29%	24%	47%	134
Other or multiple racial categories selected	33%	33%	33%	17%	33%	50%	17%	33%	50%	6
Not reported	21%	31%	48%	27%	25%	48%	31%	25%	44%	48
p-value			0.463			0.952			0.914	
Ethnicity										
Hispanic or Latino	8%	42%	50%	8%	33%	58%	17%	25%	58%	12
Not Hispanic or Latino	23%	21%	56%	26%	25%	50%	31%	25%	44%	163
Not reported	23%	36%	41%	30%	25%	45%	34%	23%	43%	44
p-value			0.127			0.672			0.817	

Respondent characteristic	Diversity of ideas			Diversity of principal investigators			Diversity of institutions			Responses (N)
	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	Decreases	Neither decreases nor increases	Increases	
Disability										
Identifies as having a disability	29%	36%	36%	36%	29%	36%	36%	36%	29%	14
Does not identify as having a disability	22%	35%	43%	22%	32%	46%	22%	32%	46%	37
Not reported	22%	22%	56%	26%	23%	51%	32%	22%	46%	168
p-value			0.330			0.624			0.383	

Source: Merit Review Examination Staff Survey, question 37: "How much does the merit review process increase or decrease the diversity of ideas, principal investigators, and institutions represented in awarded projects?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"; Merit Review Examination Staff Survey, question 47: "How do you currently describe yourself?"; Merit Review Examination Staff Survey, question 48: "Are you Hispanic or Latino?"; Merit Review Examination Staff Survey, question 49: "What is your racial background?"; Merit Review Examination Staff Survey, question 50: "Do you identify as having a disability?"

Note: In this exhibit, we show the percentage of NSF staff indicating whether the merit review process increases, decreases, or neither increases nor decreases the diversity of ideas, principal investigators, or institutions represented in awarded projects. Along with overall results, we show responses by position type, directorate or office, gender, race, ethnicity, and disability. "Increases" includes the response options "Slightly increases," "Somewhat increases," and "Greatly increases." "Neither decreases nor increases" includes the response option "Neither decreases nor increases." "Decreases" includes the response options "Greatly decreases," "Somewhat decreases," and "Slightly decreases." NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type, directorate or office, gender, race, ethnicity, or disability; *p*-values from these tests are reported.

* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

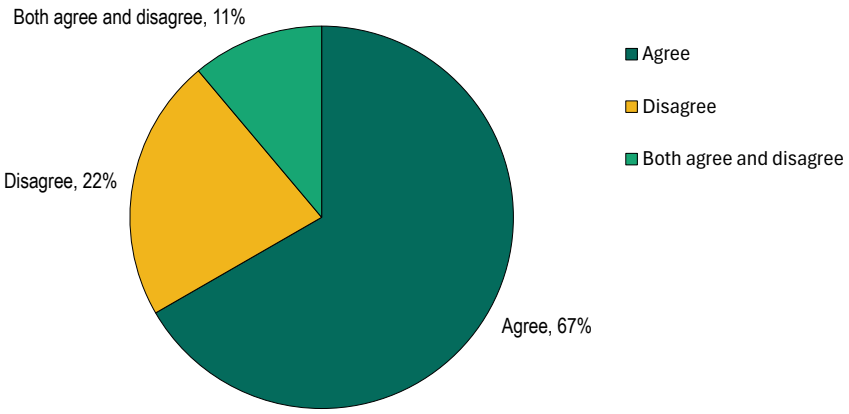
Exhibit 12.5. Number and percentage of interviewed NSF leaders who reported that they agree, disagree, or both agree and disagree that the merit review policies support a fair and accurate assessment of a proposal's merit

Response	Number	Percentage
Agree	6	67%
Disagree	2	22%
Both agree and disagree	1	11%
Responses (N)	9	100%

Source: NSF leadership interviews. Nine NSF leaders were asked, "Do you believe that the merit review policies support a fair and accurate assessment of a proposal's merit? Why or why not?"

Note: In this exhibit, we show the percentage of NSF leaders who indicated that they agree, disagree, or both agree and disagree that the merit review policies support a fair and accurate assessment of a proposal's merit.

Exhibit 12.6. Percentage of interviewed NSF leaders who reported that they agree, disagree, or both agree and disagree that the merit review policies support a fair and accurate assessment of a proposal's merit



Source: NSF leadership interviews. Nine NSF leaders were asked, "Do you believe that the merit review policies support a fair and accurate assessment of a proposal's merit? Why or why not?"

Note: In this exhibit, we show the percentage of NSF leaders who indicated that they agree, disagree, or both agree and disagree that the merit review policies support a fair and accurate assessment of a proposal's merit.

13. What aspects of the merit review policy and processes are perceived to be biased?

Exhibit 13.1. Percentage of NSF staff who indicated that they disagree, are undecided, or agree that NSF staff and NSF reviewers are sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process (overall and by position type, directorate or office, gender, race, ethnicity, and disability)

Respondent characteristic	NSF staff			NSF reviewers			Responses (N)
	Disagree	Undecided	Agree	Disagree	Undecided	Agree	
Overall							
All respondents	22%	19%	59%	25%	20%	56%	224
Position type							
Program Officer	21%	19%	59%	23%	21%	56%	187
Deputy Division Director	27%	7%	67%	27%	7%	67%	15
Division Director	29%	29%	43%	43%	14%	43%	7
Deputy Assistant Director and Deputy Office Head	0%	20%	80%	20%	0%	80%	5
Other NSF staff position	30%	30%	40%	40%	30%	30%	10
p-value			0.719			0.485	
Directorate or office							
BIO	27%	15%	58%	18%	29%	53%	34
CISE	7%	36%	57%	13%	33%	53%	15
ENG	17%	8%	75%	25%	8%	67%	24
GEO	23%	26%	52%	42%	23%	35%	31
MPS	8%	16%	76%	13%	16%	71%	38
SBE	40%	20%	40%	27%	20%	53%	15
EDU	19%	23%	58%	27%	15%	58%	26
TIP	50%	25%	25%	38%	38%	25%	8
OD/OIA	40%	40%	20%	60%	0%	40%	5
OISE	D	D	D	D	D	D	D
Not reported	D	D	D	D	D	D	D
p-value			0.167			0.182	
Gender							
Female only	26%	19%	55%	23%	26%	51%	92
Male only	14%	18%	68%	26%	12%	62%	90
Other category and not reported	27%	24%	49%	26%	21%	52%	42
p-value			0.169			0.211	

Respondent characteristic	NSF staff			NSF reviewers			Responses (N)
	Disagree	Undecided	Agree	Disagree	Undecided	Agree	
Race							
American Indian or Alaska Native	NA	NA	NA	NA	NA	NA	0
Asian	0%	25%	75%	25%	25%	50%	16
Black or African American	40%	20%	40%	47%	27%	27%	15
Native Hawaiian or Other Pacific Islander	NA	NA	NA	NA	NA	NA	0
White	21%	19%	61%	21%	20%	59%	136
Other or multiple racial categories selected	33%	17%	50%	33%	17%	50%	6
Not reported	24%	20%	56%	25%	16%	59%	51
p-value			0.382			0.480	
Ethnicity							
Hispanic or Latino	23%	31%	46%	15%	8%	77%	13
Not Hispanic or Latino	20%	18%	62%	25%	22%	53%	165
Not reported	27%	20%	53%	26%	15%	59%	46
p-value			0.647			0.465	
Disability							
Identifies as having a disability	36%	14%	50%	21%	21%	57%	14
Does not identify as having a disability	26%	21%	54%	21%	21%	59%	39
Not reported	20%	20%	61%	26%	19%	55%	171
p-value			0.625			0.968	

Source: Merit Review Examination Staff Survey, question 29: "NSF staff is sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process."; Merit Review Examination Staff Survey, question 38: "NSF reviewers are sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process."; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"; Merit Review Examination Staff Survey, question 47: "How do you currently describe yourself?"; Merit Review Examination Staff Survey, question 48: "Are you Hispanic or Latino?"; Merit Review Examination Staff Survey, question 49: "What is your racial background?"; Merit Review Examination Staff Survey, question 50: "Do you identify as having a disability?"

Note: In this exhibit, we show the percentage of staff indicating whether they agree or disagree that NSF reviewers or NSF staff are sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process. Along with overall results, we show responses by position type, directorate or office, gender, race, ethnicity, and disability. "Disagree" includes the response options "Strongly disagree" and "Disagree." "Undecided" includes the response option "Undecided." "Agree" includes the response options "Agree" and "Strongly agree." NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type, directorate or office, gender, race, ethnicity, or disability; *p*-values from these tests are reported.

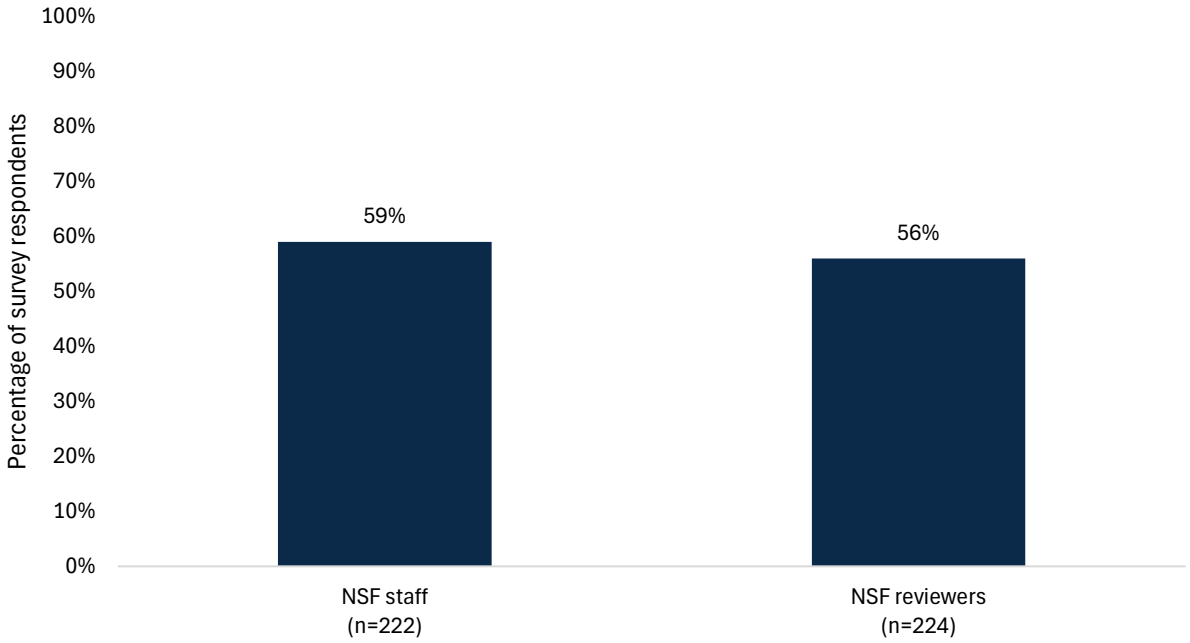
* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 13.2. Percentage of NSF staff who indicated that they agree or strongly agree that NSF staff and NSF reviewers are sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process



Source: Merit Review Examination Staff Survey, question 29: "NSF staff is sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process."; Merit Review Examination Staff Survey, question 38: "NSF reviewers are sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process."

Note: In this exhibit, we show the percentage of staff indicating they "agree" or "strongly agree" that NSF reviewers or NSF staff are sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included.

15. To what extent do the reported outputs and outcomes of funded research align with NSF's mission?

Exhibit 15.1. Percentage of NSF staff who indicated that the data NSF collects about awarded projects are somewhat or very effective at helping NSF assess whether funded projects support NSF's mission to advance scientific knowledge and benefit society (overall and by position type, directorate or office, gender, race, ethnicity, and disability)

Respondent characteristic	Advance scientific knowledge	Benefit society	Responses (N)
Overall			
All respondents	61%	39%	223
Position type			
Program Officer	60%	36%	186
Deputy Division Director	88%	69%	16
Division Director	57%	14%	7
Deputy Assistant Director and Deputy Office Head	D	D	D
Other NSF staff position	40%	40%	10
p-value	0.131	0.032*	
Directorate or office			
BIO	59%	44%	34
CISE	73%	43%	15
ENG	72%	44%	25
GEO	52%	16%	31
MPS	57%	46%	37
SBE	67%	53%	15
EDU	62%	38%	26
TIP	38%	0%	8
OD/OIA	60%	60%	5
OISE	D	D	D
Not reported	D	D	D
p-value	0.759	0.069	

Source: Merit Review Examination Staff Survey, question 39: "How effective do you think these data are for helping NSF to assess whether funded projects support NSF's mission to advance scientific knowledge?"; Merit Review Examination Staff Survey, question 40: "How effective do you think these data are for helping NSF to assess whether funded projects support NSF's mission to benefit society?"; Merit Review Examination Staff Survey, question 1: "Which of the following positions do you currently hold with NSF?"; Merit Review Examination Staff Survey, question 44: "Which directorate or office is your position in?"

Note: In this exhibit, we show the percentage of NSF staff who think the data NSF collects about awarded projects are "Very effective" or "Somewhat effective" at helping NSF assess whether funded projects support NSF's mission to advance scientific knowledge and benefit society. Along with overall results, we show responses by position type and by directorate or office. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included. The number of responses may vary by item. Chi-squared tests were used to assess whether response patterns for each item differ by position type or by directorate or office; *p*-values from these tests are reported.

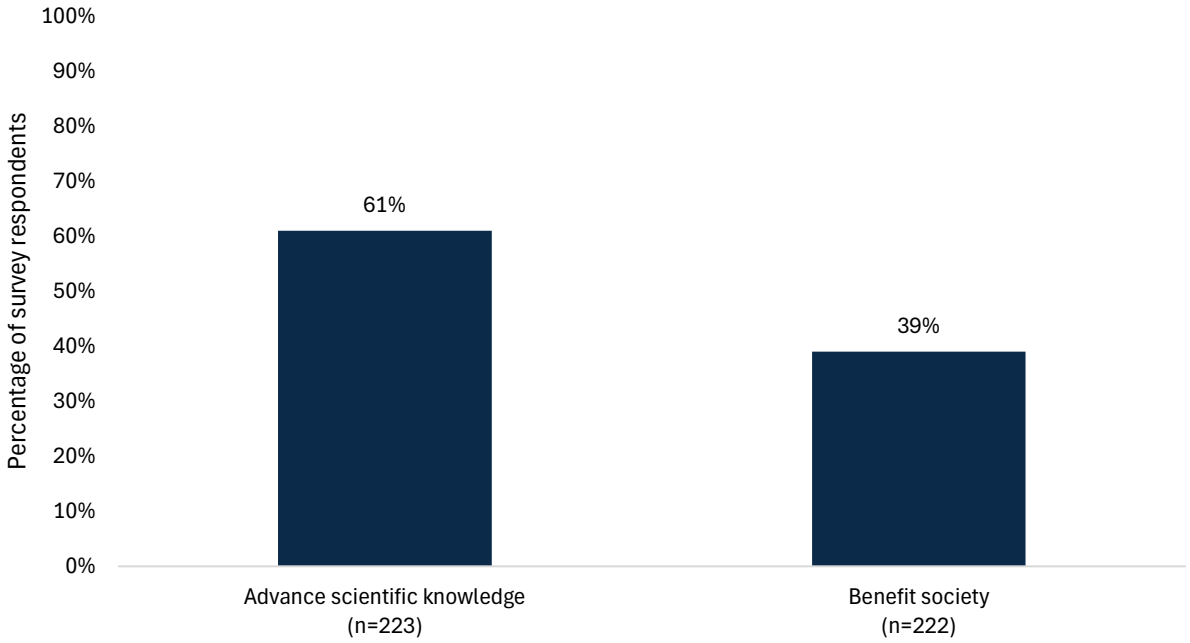
* Significantly different from zero at the .05 level, one-tailed test.

** Significantly different from zero at the .01 level, one-tailed test.

D signals that the estimate was suppressed to avoid disclosure of confidential, sensitive, or otherwise protected information.

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; SBE = Social, Behavioral and Economic Sciences; EDU = Science, Technology, Engineering, and Mathematics (STEM) Education; TIP = Technology, Innovation and Partnerships; OD/OIA = Office of Integrative Activities; OISE = Office of International Science and Engineering.

Exhibit 15.2. Percentage of NSF staff who indicated that the data NSF collects about awarded projects are somewhat or very effective at helping NSF assess whether funded projects support NSF's mission to advance scientific knowledge and benefit society



Source: Merit Review Examination Staff Survey, question 39: "How effective do you think these data are for helping NSF to assess whether funded projects support NSF's mission to advance scientific knowledge?"; Merit Review Examination Staff Survey, question 40: "How effective do you think these data are for helping NSF to assess whether funded projects support NSF's mission to benefit society?"

Note: In this exhibit, we show the percentage of NSF staff who think the data NSF collects about awarded projects are "Very effective" or "Somewhat effective" at helping NSF assess whether funded projects support NSF's mission to advance scientific knowledge and benefit society. NSF staff include Program Officers, Deputy Division Directors, Division Directors, Deputy Assistant Directors, and Deputy Office Heads who have had responsibility for merit review in the last three years. Intergovernmental Personnel Act (IPA) staff are not included.

Appendix B.

Data Collection Items by Guiding Question

Exhibit 16. Items from each data source, by guiding question

Guiding question and relevant items from each data source

1. How do PIs, reviewers, and NSF leaders and staff interpret the merit review criteria?

NSF Staff Survey

- Q4: According to your understanding of the merit review policy, how important are the following factors for receiving a high rating on a proposal? (See Appendix C for all question prompts)
- Q11: Do you think it is generally clear how a proposal's merit is to be assessed against the Intellectual Merit criterion and the Broader Impacts criterion?
- Q12: What is unclear about how a proposal's merit is to be assessed against the Intellectual Merit criterion?
- Q14: What is unclear about how a proposal's merit is to be assessed against the Broader Impacts criterion?
- Q23: What, if any, are the types of questions you get from [reviewers and principal investigators/program officers] related to the merit review criteria?

NSF Leadership Interviews

- Q2: What do you think NSF is trying to accomplish with the Intellectual Merit and Broader Impacts criteria?
- Q3a: Do you consider one [the Intellectual Merit or Broader Impacts criterion] to be more important than the other, and why or why not?
- Q5: What, if anything, about the scientific field of your directorate, division, or program makes it challenging to interpret and apply the merit review criteria?

RFI

- Q1: The MRX is interested in identifying opportunities to improve NSF's current Merit Review criteria, policy, and processes. Importantly, this includes documenting and understanding any areas of misunderstanding, gaps, or lack of clarity regarding (a) the three Merit Review Principles which are the foundations of the Merit Review Process, (b) the two statutory Merit Review Criteria which are used to evaluate all proposals to NSF, and (c) the five Merit Review Elements NSF uses to assess each criterion; for instance: Are the Principles, Criteria, and Elements clear? Could they be improved upon?

2. How do PIs, reviewers, and NSF leaders and staff use the merit review criteria?

NSF Staff Survey

- Q5: In your experience, how do reviewers rank the importance of the following when assigning a high rating on a proposal? (See Appendix C for all question prompts)
- Q6: The NSF Proposal & Award Policies & Procedures Guide (PAPPG) notes that the two merit review criteria (Intellectual Merit and Broader Impacts) are to be given full consideration during the review and decision-making processes. Each criterion is necessary but neither, by itself, is sufficient. In practice, how is full consideration achieved to produce a single [recommendation/decision] for an award?
- Q7: In your experience, are there situations in which the importance of the Intellectual Merit criterion and the Broader Impacts criterion are weighted unevenly?
- Q8: In your experience, how frequently are the merit review criteria weighted unevenly?
- Q9: What weight do you typically place on the importance of the Intellectual Merit criterion compared to the Broader Impacts criterion in your use of the merit review criteria?
- Q16: How challenging do you generally find it to assess a proposal's merit against each criterion?
- Q17: Have you had any recurrent challenges assessing a proposal's merit against each criterion?
- Q18: What recurrent challenges have you had assessing a proposal's merit against each criterion?
- Q19: In your opinion, is there anything about the scientific field(s) of your directorate, division, or program that makes it challenging to apply the Intellectual Merit criterion?
- Q20: What, if anything, about the scientific field(s) of your directorate, division, or program makes it challenging to apply the Intellectual Merit criterion?

- Q21: In your opinion, is there anything about the scientific field(s) of your directorate, division, or program that makes it challenging to apply the Broader Impacts criterion?
- Q22: What, if anything, about the scientific field(s) of your directorate, division, or program makes it challenging to apply the Broader Impacts criterion?

NSF Leadership Interviews

- Q3: In practice, how is full consideration achieved in your leadership of the [directorate's/office's] portfolio?
- Q5: What, if anything, about the scientific field of your directorate, division, or program makes it challenging to interpret and apply the merit review criteria?

RFI

- Q2: NSF strives to conduct a fair, competitive, transparent Merit Review process for the selection of projects. To accomplish this, NSF relies on a process that considers both the technical aspects of a proposed project and its potential to contribute more broadly to advancing NSF's mission using the statutory Intellectual Merit and Broader Impacts Merit Review Criteria. MRX invites suggestions on the implementation of the Merit Review Criteria. We especially invite feedback that would (a) clarify how the Merit Review Criteria can be used in preparing and reviewing proposals, (b) ensure proposals, reviews, and funding decisions demonstrate full consideration of both criteria while maintaining openness to the full spectrum of potential activities under each, and (c) better recognize and support potentially transformative and high-risk/high-reward activities.
- Q4: MRX is interested in the experiences and perspectives of those who have reviewed proposals submitted to NSF. We invite you to share your insights and describe any opportunities you believe would improve implementation of the Merit Review criteria, policy, or processes based on your experience reviewing NSF proposals.

3. What published and ad hoc guidance does NSF offer to PIs and reviewers on interpreting and using the merit review criteria in preparing or evaluating proposals?

NSF Staff Survey

- Q13: In your opinion, what could be done to make it clearer how a proposal's merit is to be assessed against the Intellectual Merit criterion?
- Q15: In your opinion, what could be done to make it clearer how a proposal's merit is to be assessed against the Broader Impacts criterion?

NSF Leadership Interviews

- Q4: What instructions or training do you provide division directors, program officers, and reviewers to help them interpret and apply the merit review criteria?
- Q4a: If division directors have questions about the merit review criteria or are unclear about how to apply them, how would they go about asking for clarification? What about program officers? What about reviewers?

RFI

- Q2: NSF strives to conduct a fair, competitive, transparent Merit Review process for the selection of projects. To accomplish this, NSF relies on a process that considers both the technical aspects of a proposed project and its potential to contribute more broadly to advancing NSF's mission using the statutory Intellectual Merit and Broader Impacts Merit Review Criteria. MRX invites suggestions on the implementation of the Merit Review Criteria. We especially invite feedback that would (a) clarify how the Merit Review Criteria can be used in preparing and reviewing proposals, (b) ensure proposals, reviews, and funding decisions demonstrate full consideration of both criteria while maintaining openness to the full spectrum of potential activities under each, and (c) better recognize and support potentially transformative and high-risk/high-reward activities.

4. How do PIs, reviewers, and NSF leaders and staff understand and interpret the merit review principles and elements?

NSF Staff Survey

- Q31: How would you rate the overall level of understanding of the merit review principles by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators? (See Appendix C for all question prompts)
- Q32: How would you rate the overall level of understanding of the merit review elements by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators? (See Appendix C for all question prompts)

NSF Leadership Interviews

- Q6: Thinking specifically about the review of proposals submitted to or managed by your [directorate/office], how are the merit review principles and elements used to identify which projects to fund?
- Q6a: Specifically, how are the principles and elements used by program officers to make funding recommendations in your directorate/office?
- Q6b: How are principles and elements used by division directors in their portfolio management?
- Q6c: In your experience, are these uses similar to or different from how the principles and elements are used in other directorates and offices?
- Q7: How confident are you that the principles and elements are used consistently within your directorate? Can you tell me about where you think there are areas of consistency and inconsistency?

RFI

- Q1: The MRX is interested in identifying opportunities to improve NSF's current Merit Review criteria, policy, and processes. Importantly, this includes documenting and understanding any areas of misunderstanding, gaps, or lack of clarity regarding (a) the three Merit Review Principles which are the foundations of the Merit Review Process, (b) the two statutory Merit Review Criteria which are used to evaluate all proposals to NSF, and (c) the five Merit Review Elements NSF uses to assess each criterion; for instance: Are the Principles, Criteria, and Elements clear? Could they be improved upon?

5. What additional merit review criteria might NSF consider using in the merit review process to better achieve its mission to invest in research that advances scientific knowledge and benefits society?

NSF Staff Survey

- Q24: Based on your experience, are there any factors important for evaluating proposals that are not captured by the two merit review criteria?
- Q25: Please describe the factor(s) and why you think it would be important for reviewing proposals. What could this capture that is currently missed by the merit review criteria?

NSF Leadership Interviews

- Q9: In your opinion, are there any other important factors for evaluating proposals that are not captured by the merit review criteria?
- Q9a: Tell me a bit more about this factor and why you think it would be important for reviewing proposals. What could this capture that is missed by the current merit review criteria?

RFI

- N/A

6. How is reviewer and staff expertise in the merit review criteria assessed?

NSF Staff Survey

- Q26: How effective do you think [you are/your unit is] at identifying which [NSF program officers/reviewers] have the necessary expertise of the merit review criteria?
- Q30: How would you rate the overall level of understanding of the merit review criteria by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators?

NSF Leadership Interviews

- Q10/10a: Who determines which NSF staff have the necessary expertise of the merit review criteria to evaluate proposals? How do they make that determination?
- Q11/11a: Who determines which reviewers have the necessary expertise of the merit review criteria to evaluate proposals? How do they make that determination?

RFI

- N/A

7. How does reviewer and staff expertise on the merit review criteria factor into reviews, recommendations, and decisions?

NSF Staff Survey

- N/A

NSF Leadership Interviews

- Q12/12a: In your [directorate/office], how do you evaluate whether program officers are reviewing proposals and making recommendations that are aligned to the merit review criteria? What about reviewers?
- Q13: How do you evaluate whether division directors are making decisions about whether to award or decline proposals in a way that is aligned to the merit review criteria?

RFI

- N/A

8. How do reviewers balance the Intellectual Merit and Broader Impacts criteria in assessing proposals?

NSF Staff Survey

- N/A

NSF Leadership Interviews

- N/A

RFI

- Q4: MRX is interested in the experiences and perspectives of those who have reviewed proposals submitted to NSF. We invite you to share your insights and describe any opportunities you believe would improve implementation of the Merit Review criteria, policy, or processes based on your experience reviewing NSF proposals.

9. How do reviewers' assessments of each merit review criterion factor into NSF program officers' award recommendations?

NSF Staff Survey

- Q28: How much do reviewers' assessments of each merit review criterion factor into [funding recommendations/portfolio management] within your division? (See Appendix C for all question prompts)

NSF Leadership Interviews

- N/A

RFI

- N/A

10. How do reviewers' assessments of each merit review criterion factor into NSF division directors' award decisions for individual proposals? At the portfolio level?

NSF Staff Survey

- Q28: How much do reviewers' assessments of each merit review criterion factor into [funding recommendations/portfolio management] within your division? (See Appendix C for all question prompts)

NSF Leadership Interviews

- Q14: How are reviewers' assessments of each merit review criterion factored into portfolio management within your [directorate/office]?

RFI

- N/A

11. To what extent does use of reviewers' assessments of each merit review criterion vary within and across NSF directorates?

NSF Staff Survey

- This question is addressed in our analysis by comparing responses to staff survey Q28 across directorates.

NSF Leadership Interviews

- N/A

RFI

- N/A

12. To what extent do NSF constituencies perceive the merit review policy and processes to be unfairly biased? How do these perceptions differ by constituency?

NSF Staff Survey

- Q33: Based on your experience with the merit review process, to what extent do you agree or disagree with the following statements? (See Appendix C for all question prompts)
- Q36: How much does the merit review process increase or decrease the diversity of ideas, principal investigators, and institutions represented in *proposals submitted for review?* (See Appendix C for all question prompts)
- Q37: How much does the merit review process increase or decrease the diversity of ideas, principal investigators, and institutions represented in *awarded projects?* (See Appendix C for all question prompts)

NSF Leadership Interviews

- Q15: Do you believe that the merit review policies support a fair and accurate assessment of a proposal's merit? Why or why not?
- Q16: Are there any aspects of the merit review process that you think introduce unfair bias into the evaluation of proposals? If so, how?
- Q16a: What, if anything, does your [directorate/office] implement to mitigate potential biases?
- Q17: Do you think the merit review process increases or decreases the diversity of proposals submitted for review? How so? (See Appendix C for all question prompts)
- Q18: Do you think the merit review process increases or decreases the diversity of awarded projects? How so? (See Appendix C for all question prompts)

RFI

- N/A

13. What aspects of the merit review policy and processes are perceived to be biased?

NSF Staff Survey

- Q34: Are there any aspects of the merit review *process* that you think introduce unfair bias into the evaluation of proposals? If so, how?
- Q35: Are there any aspects of the merit review *policies* that you think introduce unfair bias into the evaluation of proposals? If so, how?
- Q29 & Q38: NSF [reviewers/staff] are sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process. (See Appendix C for all question prompts)

NSF Leadership Interviews

- Q16: Are there any aspects of the merit review process that you think introduce unfair bias into the evaluation of proposals? If so, how?

RFI

- N/A

14. What barriers to participation do NSF constituencies perceive in the merit review policy and processes?

NSF Staff Survey

- N/A

NSF Leadership Interviews

- N/A

RFI

- Q3: MRX is interested in the experiences and perspectives of those who have considered submitting and/or submitted proposals in the past. We invite you to share your insights and describe any opportunities you believe would improve implementation of the Merit Review criteria, policy, and processes based on your experience as a proposer or investigator. This includes any experiences that may have encouraged or dissuaded you from submitting proposals to NSF. We are especially interested in learning (a) how NSF guidance (e.g., as provided in the NSF PAPPG, program solicitations, or other funding opportunity announcements), may have played a part in your decision(s) whether to submit proposals, and (b) how NSF might best support investigators interested in submitting a proposal to NSF.

15. To what extent do the reported outputs and outcomes of funded research align with NSF's mission?

NSF Staff Survey

- Q39 & Q40: How effective do you think these data are for helping NSF to assess whether funded projects support NSF's mission to [*advance scientific knowledge/benefit society]*? (See Appendix C for all question prompts)

NSF Leadership Interviews

- Q19: What types of evidence does your [directorate/office] collect and examine to assess whether the projects that received funding helped support NSF's mission to advance scientific knowledge and benefit society? This might include evidence from third-party evaluations, program monitoring systems, or ETAP [NSF Education and Training Application].

RFI

- N/A

16. What guidance does NSF offer to PIs on documenting outcomes that address the merit review criteria in annual and final reports?

NSF Staff Survey

- N/A

NSF Leadership Interviews

- N/A

RFI

- N/A

17. How might NSF better measure the outcomes of funded research with respect to each merit review criterion?

NSF Staff Survey

- Q41: What, if any, additional information do you think NSF should be collecting to monitor progress toward *advancing scientific knowledge* that is not being collected?
- Q42: What, if any, additional information do you think NSF should be collecting to monitor its progress toward *benefitting society* that is not being collected?

NSF Leadership Interviews

- Q20: Is there additional information that NSF could be collecting to monitor its progress to advance scientific knowledge and benefit society?
- Q20a: Is there anything about the *quality* of the data that NSF is collecting that might make it difficult to use to monitor its progress?

RFI

- Q5: MRX is interested in exploring how NSF could better support awardees in demonstrating and documenting outcomes of their awards in advancing knowledge (Intellectual Merit) and benefiting society and contributing to the achievement of desired broader or societal outcomes (Broader Impacts). We invite you to share your insights on how NSF might better support awardees in demonstrating and documenting outcomes of their awards without unnecessarily increasing awardees' administrative burden of reporting.

Appendix C.

Data Collection Instruments

NSF Staff Survey

PROGRAMMER BOX
RESPONDENT TYPES:
PO = Program Officer
DDD = Deputy Division Director
DD = Division Director
DAD = Deputy Assistant Director
DH = Deputy Office Head
Sample size will be approximately 600

BACKGROUND INFORMATION
All respondents If Q1 = 99 or Missing then respondent type is PO.

The first set of questions asks about your role at NSF and if you have any responsibility for merit review of proposals.

1. Which of the following positions do you currently hold with NSF?

Select one only

- ❍ Program Officer or functionally similar role 1
- ❍ Deputy Division Director 2
- ❍ Division Director 3
- ❍ Deputy Assistant Director 4
- ❍ Deputy Office Head 5
- ❍ Other (please specify) 99

Specify [ ] (STRING 500)

BACKGROUND INFORMATION
All respondents

2. What type of appointment do you currently hold with NSF?

Select one only

- ❍ Federal personnel (Permanent, Temporary, VSEE) 1
- ❍ Intergovernmental Personnel Act (IPA) assignment..... 2 Go to END

SOFT CHECK: IF Q2=2; **You selected that you are an Intergovernmental Personnel Act (IPA) assignment (also known as rotator). If that is correct, please select Next.**

EXPERIENCE WITH EVALUATING PROPOSALS FOR MERIT REVIEW CRITERIA
IF Q.2 = 1 OR M

3. When was the last time you had responsibility for implementation and/or oversight of proposal merit review?

Select one only

- ❍ I currently have relevant responsibilities 1
- ❍ 1 year ago..... 2
- ❍ 2-3 years ago..... 3
- ❍ More than 3 years ago..... 4 Go to END
- ❍ I never had relevant responsibilities ..... 5 Go to END

SOFT CHECK: IF Q3= 4; **You selected that it has been more than 3 years since you had relevant responsibilities for merit review proposals. If that is correct, please select Next.** SOFT CHECK: IF Q3=5; **You selected that you never had relevant responsibilities. If that is correct, please select Next.**

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
If Q3 = 1, 2,3 OR M.

For the purposes of this survey, please only think about your responsibility for merit review of proposals in the last 3 years.

The next set of questions will ask for your interpretation and use of the merit review criteria, principles, and elements.

4. According to your understanding of the merit review policy, how important are the following factors for receiving a high rating on a proposal?

Select one per row

	Not at all important	A little important	Somewhat important	Very important
a. **Merit review principle**: All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.	1 ❍	2 ❍	3 ❍	4 ❍
b. **Merit review principle**: NSF projects, in the aggregate, should contribute more broadly to achieving societal goals.	1 ❍	2 ❍	3 ❍	4 ❍
c. **Merit review principle**: Meaningful assessment and evaluation of NSF funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of broader impacts and the resources provided to implement projects.	1 ❍	2 ❍	3 ❍	4 ❍
d. **Intellectual Merit criterion**: The potential to advance knowledge.	1 ❍	2 ❍	3 ❍	4 ❍
e. **Broader Impacts criterion**: The potential to benefit society and contribute to the achievement of specific, desired societal outcomes.	1 ❍	2 ❍	3 ❍	4 ❍
f. **Merit review element**: What is the potential for the proposed activity to advance knowledge and understanding within its own field or across different fields (intellectual merit)?	1 ❍	2 ❍	3 ❍	4 ❍
g. **Merit review element**: What is the potential for the proposed activity to benefit society or advance desired societal outcomes (broader impacts)?	1 ❍	2 ❍	3 ❍	4 ❍
h. **Merit review element**: To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?	1 ❍	2 ❍	3 ❍	4 ❍

	Not at all important	A little important	Somewhat important	Very important
i. **Merit review element**: Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?	1 ❍	2 ❍	3 ❍	4 ❍
j. **Merit review element**: How well qualified is the individual, team, or organization to conduct the proposed activities?	1 ❍	2 ❍	3 ❍	4 ❍
k. **Merit review element**: Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?	1 ❍	2 ❍	3 ❍	4 ❍
l. Other important factor for receiving a high rating (Please specify) [(STRING 500)]	1 ❍	2 ❍	3 ❍	4 ❍

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
If respondent = PO and Q3 = 1, 2 or 3.

5. In your experience, how do *reviewers* rank the importance of the following when assigning a high rating on a proposal?

Select one per row

	Not at all important	A little important	Somewhat important	Very important
a. **Merit review principle**: All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.	1 ❍	2 ❍	3 ❍	4 ❍
b. **Merit review principle**: NSF projects, in the aggregate, should contribute more broadly to achieving societal goals.	1 ❍	2 ❍	3 ❍	4 ❍
c. **Merit review principle**: Meaningful assessment and evaluation of NSF funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of broader impacts and the resources provided to implement projects.	1 ❍	2 ❍	3 ❍	4 ❍
d. **Intellectual Merit**: The potential to advance knowledge.	1 ❍	2 ❍	3 ❍	4 ❍

		Not at all important	A little important	Somewhat important	Very important
e.	**Broader Impacts**: The potential to benefit society and contribute to the achievement of specific, desired societal outcomes.	1 ❍	2 ❍	3 ❍	4 ❍
f.	**Merit review element**: What is the potential for the proposed activity to advance knowledge and understanding within its own field or across different fields (Intellectual Merit)?	1 ❍	2 ❍	3 ❍	4 ❍
g.	**Merit review element**: What is the potential for the proposed activity to benefit society or advance desired societal outcomes (Broader Impacts)?	1 ❍	2 ❍	3 ❍	4 ❍
h.	**Merit review element**: To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?	1 ❍	2 ❍	3 ❍	4 ❍
i.	**Merit review element**: Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?	1 ❍	2 ❍	3 ❍	4 ❍
j.	**Merit review element**: How well qualified is the individual, team, or organization to conduct the proposed activities?	1 ❍	2 ❍	3 ❍	4 ❍
k.	**Merit review element**: Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?	1 ❍	2 ❍	3 ❍	4 ❍
l.	Other important factor for receiving a high rating (Please specify) (STRING 500)	1 ❍	2 ❍	3 ❍	4 ❍

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
All respondents If PO fill "recommendation" If DD, DDD, DAD, or DH fill "decision"

6. The NSF Proposal & Award Policies & Procedures Guide (PAPPG) notes that the two merit review criteria (Intellectual Merit and Broader Impacts) are to be given full consideration during the review and decision-making processes. Each criterion is necessary but neither, by itself, is sufficient. In practice, how is full consideration achieved to produce a single [recommendation/decision] for an award??

(RANGE 4,000)

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
All respondents If Q7 = 1 go to Q8 and Q9 If Q7 = 0 go to Q11.

7. In your experience, are there situations in which the importance of the Intellectual Merit criterion and the Broader Impacts criterion are weighted unevenly?

- ❍ Yes 1
- ❍ No ..... 0 Go to Q11

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
If Q7 = 1

8. In your experience, how frequently are the merit review criteria weighted unevenly?

Select one only

- ❍ Never 1
- ❍ Rarely 2
- ❍ Sometimes 3
- ❍ Always 4

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
If Q7 = 1

9. What weight do you typically place on the importance of the Intellectual Merit criterion compared to the Broader Impacts criterion in your use of the merit review criteria?

Select one only

- ❍ Much more weight on Intellectual Merit 1
- ❍ More weight on Intellectual Merit 2
- ❍ Somewhat more weight on Intellectual Merit 3
- ❍ Both equally 4
- ❍ Somewhat more weight on Broader Impacts 5
- ❍ More weight on Broader Impacts 6
- ❍ Much more weight on Broader Impacts 7

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
All respondents If Q11.a AND Q11.b = 1 go to Q16. If Q11.a = 0 go to Q12 and Q13. If Q11.b = 0 go to Q14 and Q15.

11. Do you think it is generally clear how a proposal's merit is to be assessed against the Intellectual Merit criterion and the Broader Impacts criterion?

Select one per row

	Yes	No	
a. Intellectual Merit criterion	1 ❍	0 ❍	If 0 go to Q12
b. Broader Impacts criterion	1 ❍	0 ❍	If 0 go to Q14

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
If Q11.a = 0

12. What is unclear about how a proposal's merit is to be assessed against the *Intellectual Merit* criterion?

(STRING 4,000)

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
If Q11.a = 0

13. In your opinion, what could be done to make it clearer how a proposal's merit is to be assessed against the *Intellectual Merit* criterion?

(STRING 4,000)

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
If Q11.b = 0

14. What is unclear about how a proposal's merit is to be assessed against the *Broader Impacts* criterion?

(STRING 4,000)

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
If Q11.b = 0

15. In your opinion, what could be done to make it clearer how a proposal's merit is to be assessed against the *Broader Impacts* criterion?

(STRING 4,000)

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
All respondents

16. How challenging do you generally find it to assess a proposal's merit against each criteria?

Select one per row

	Not at all challenging	A little challenging	Somewhat challenging	Very challenging
a. Intellectual Merit criterion	1 ❍	2 ❍	3 ❍	4 ❍
b. Broader Impacts criterion	1 ❍	2 ❍	3 ❍	4 ❍

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
All respondents If Q17 = 0 go to Q19

17. Have you had any recurrent challenges assessing a proposal's merit against each criterion?

- ❍ Yes..... 1
- ❍ No ..... 0 Go to Q19

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
If Q17 = 1

18. What recurrent challenges have you had assessing a proposal's merit against each criterion?

(STRING 4,000)

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
All respondents If Q19 = 0 go to Q21

19. In your opinion, is there anything about the scientific field(s) of your directorate, division, or program that makes it challenging to apply the *Intellectual Merit* criterion?

- ❍ Yes..... 1
- ❍ No ..... 0 Go to Q21

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
If Q19 = 1

20. What, if anything, about the scientific field(s) of your directorate, division, or program makes it challenging to apply the *Intellectual Merit* criterion?

(STRING 4,000)

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
All respondents If Q21 = 0 got to Q23.

21. In your opinion, is there anything about the scientific field(s) of your directorate, division, or program that makes it challenging to apply the *Broader Impacts* criterion?

- ❍ Yes..... 1
- ❍ No ..... 0 Go to Q23

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
If Q21 = 1

22. What, if anything, about the scientific field(s) of your directorate, division, or program makes it challenging to apply the *Broader Impacts* criterion?

(STRING 4,000)

INTERPRETATION AND UTILIZATION OF MERIT REVIEW CRITERIA BY PIS, REVIEWERS, AND NSF PROGRAM STAFF
All respondents If PO fill “reviewers and principal investigators” If DD, DDD, DAD, or DH fill “program officers”

23. What, if any, are the types of questions you get from [reviewers and principal investigators/program officers] related to the merit review criteria?

(STRING 4,000)

PERCEPTIONS AMONG PIS, REVIEWERS, AND NSF PROGRAM STAFF ON THE VALUE OF ADDITIONAL MERIT REVIEW CRITERIA
All respondents If Q24 = 0 go to Q26

24. Based on your experience, are there any factors important for evaluating proposals that are not captured by the two merit review criteria?

- ❍ Yes..... 1
- ❍ No ..... 0 Go to Q26

PERCEPTIONS AMONG PIS, REVIEWERS, AND NSF PROGRAM STAFF ON THE VALUE OF ADDITIONAL MERIT REVIEW CRITERIA
If Q24 = 1

25. Please describe the factor(s) and why you think it would be important for reviewing proposals. What could this capture that is currently missed by the merit review criteria?

(STRING 4,000)

ASSESSING AND ACCOUNTING FOR REVIEWER AND NSF STAFF EXPERTISE ABOUT MERIT REVIEW CRITERIA
Respondent = DD, DDD, DAD, or DH fill "your unit" Respondent = PO fill "you are"

26. How effective do you think [you are/your unit is] at identifying which reviewers have the necessary expertise of the merit review criteria?

Select one only

- ❍ Not at all effective 1
- ❍ A little effective..... 2
- ❍ Somewhat effective 3
- ❍ Very effective 4

HOW REVIEWERS USE THE MERIT REVIEW CRITERIA AND HOW REVIEWS ARE USED BY NSF PROGRAM OFFICERS AND DIVISION DIRECTORS
All respondents If respondent = PO fill "funding recommendations" Respondent = DD, DDD, DAD, or DH fill "portfolio management"

28. How much do reviewers' assessments of each merit review criterion factor into [funding recommendations/portfolio management] within your division?

Select one per row

	Not at all	A little	Somewhat	To a great extent
a. Intellectual Merit criterion	1 ❍	2 ❍	3 ❍	4 ❍
b. Broader Impacts criterion	1 ❍	2 ❍	3 ❍	4 ❍

ASSESSING AND ACCOUNTING FOR REVIEWER AND NSF STAFF EXPERTISE ABOUT MERIT REVIEW CRITERIA
All respondents

30. How would you rate the overall level of understanding of the merit review *criteria* by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators?

Select one per row

	Very low level	Low level	Moderate level	High level	Very high level	No basis to rate
a. You	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
b. Program Officers	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
c. Division Directors, Deputy Division Directors, and Section Heads	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
d. Reviewers	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
e. Principal investigators	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍

The next questions will ask about how merit review principles and elements are applied in your directorate.

PROGRAMMER BOX

THE CLIENT WOULD LIKE THE DEFINITION BELOW TO BE INCLUDED WITH THE QUESTIONS. WE SHOULD DO A HOVER DEFINTION FOR THE WORDS HIGLIGHTED IN YELLOW IN EACH QUESTION. IF IT'S EASIER THEY COULD ALSO GO TO A SEPARATE PAGE IF THEY CLICK ON THE HIGHLIGHTED WORD IN THE QUESTIONS.

The definitions for principles and elements are from the NSF Proposal & Award Policies & Procedures Guide (PAPPG) Chapter III: NSF Proposal Processing and Review.

Principles

- **All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.**
- **NSF projects, in the aggregate, should contribute more broadly to achieving societal goals. These broader impacts may be accomplished through the research itself, through activities that are directly related to specific research projects, or through activities that are supported by, but are complementary to, the project. The project activities may be based on previously established and/or innovative methods and approaches, but in either case must be well justified.**

- **Meaningful assessment and evaluation of NSF funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of broader impacts and the resources provided to implement projects. If the size of the activity is limited, evaluation of that activity in isolation is not likely to be meaningful. Thus, assessing the effectiveness of these activities may best be done at a higher, more aggregated, level than the individual project.**

Elements

- **What is the potential for the proposed activity to:**
  - **Advance knowledge and understanding within its own field or across different fields (Intellectual Merit)**
  - **Benefit society or advance desired societal outcomes (Broader Impacts)?**
- **To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?**
- **Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?**
- **How well qualified is the individual, team, or organization to conduct the proposed activities?**
- **Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?**

APPLICATION OF THE MERIT REVIEW PRINCIPLES AND ELEMENTS
All respondents

31. How would you rate the overall level of understanding of the merit review *principles* by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators?

Select one per row

	Very low level	Low level	Moderate level	High level	Very high level	No basis to rate
a. You	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
b. Program Officers	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
c. Division Directors, Deputy Division Directors, and Section Heads	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
d. Reviewers	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
e. Principal investigators	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍

APPLICATION OF THE MERIT REVIEW PRINCIPLES AND ELEMENTS
All respondents

32. How would you rate the overall level of understanding of the merit review *elements* by you, Program Officers, Division Directors, Deputy Division Directors, Section Heads, reviewers, and principal investigators?

Select one per row

	Very low level	Low level	Moderate level	High level	Very high level	No basis to rate
a. You	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
b. Program Officers	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
c. Division Directors, Deputy Division Directors, and Section Heads	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
d. Reviewers	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍
e. Principal investigators	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍

PERCEPTIONS OF UNFAIR BIAS AND BARRIERS TO PARTICIPATION WITHIN MERIT REVIEW POLICIES AND PROCESS
All respondents

The next question is about your perceptions of fairness in the merit review criteria and process.

33. Based on your experience with the merit review process, to what extent do you agree or disagree with the following statements?

Select one per row

	Strongly disagree	Disagree	Neither disagree or agree	Agree	Strongly agree
a. Individuals submitting proposals are treated fairly.	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍
b. The merit review criteria support a fair and accurate assessment of a proposal's merit.	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍

PERCEPTIONS OF UNFAIR BIAS AND BARRIERS TO PARTICIPATION WITHIN MERIT REVIEW POLICIES AND PROCESS
All respondents

The next questions are about perceptions of unfair bias in the merit review policies and process.

Bias can be introduced by differential treatment, when individuals are treated unequally because of their background, or disparate impact, when individuals are treated equally according to a given set of criteria, but the criteria favor members of particular groups (Pager and Shepherd 2008).

34. Are there any aspects of the merit review *process* that you think introduce unfair bias into the evaluation of proposals? If so, how?

(STRING 4,000)

PERCEPTIONS OF UNFAIR BIAS AND BARRIERS TO PARTICIPATION WITHIN MERIT REVIEW POLICIES AND PROCESS
All respondents

35. Are there any aspects of the merit review *policies* that you think introduce unfair bias into the evaluation of proposals? If so, how?

(STRING 4,000)

PERCEPTIONS OF UNFAIR BIAS AND BARRIERS TO PARTICIPATION WITHIN MERIT REVIEW POLICIES AND PROCESS
All respondents

The Proposal & Award Policies & Procedures Guide notes that merit review criteria and process are designed to inform NSF investments "in a robust and diverse portfolio of projects that creates new knowledge and enables breakthroughs in understanding across all areas of science and engineering research and education".

The next questions are about how the merit review process supports or hinders achieving a diverse portfolio of projects that creates new knowledge. Diversity can include diversity in the ideas being submitted and/or the characteristics of principal investigators or institutions submitting the proposal.

36. How much does the merit review process increase or decrease the diversity of ideas, principal investigators, and institutions represented in *proposals submitted for review?*

Select one per row

	Greatly decreases	Somewhat decreases	Slightly decreases	Neither decreases nor increases	Slightly increases	Somewhat increases	Greatly increases
a. Diversity of ideas	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍	7 ❍
b. Diversity of principal investigators	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍	7 ❍
c. Diversity of Institutions	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍	7 ❍

PERCEPTIONS OF UNFAIR BIAS AND BARRIERS TO PARTICIPATION WITHIN MERIT REVIEW POLICIES AND PROCESS
All respondents

37. How much does the merit review process increase or decrease the diversity of ideas, principal investigators, and institutions represented in *awarded projects?*

Select one per row

	Greatly decreases	Somewhat decreases	Slightly decreases	Neither decreases nor increases	Slightly increases	Somewhat increases	Greatly increases
a. Diversity of ideas	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍	7 ❍
b. Diversity of principal investigators	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍	7 ❍
c. Diversity of Institutions	1 ❍	2 ❍	3 ❍	4 ❍	5 ❍	6 ❍	7 ❍

PERCEPTIONS OF UNFAIR BIAS AND BARRIERS TO PARTICIPATION WITHIN MERIT REVIEW POLICIES AND PROCESS
All respondents

How much do you agree or disagree with the following statement?

38. *NSF reviewers* are sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process.

Select one only

- ❍ Strongly disagree 1
- ❍ Disagree 2
- ❍ Undecided 3
- ❍ Agree 4
- ❍ Strongly agree 5

PERCEPTIONS OF UNFAIR BIAS AND BARRIERS TO PARTICIPATION WITHIN MERIT REVIEW POLICIES AND PROCESS
All respondents

How much do you agree or disagree with the following statement?

29. *NSF staff* is sufficiently diverse in terms of individual and institutional characteristics to achieve the goals of the merit review process.

Select one only

- ❍ Strongly disagree 1
- ❍ Disagree 2
- ❍ Undecided 3
- ❍ Agree 4
- ❍ Strongly agree 5

HOW NSF CURRENTLY DEMONSTRATES, BOTH QUANTITATIVELY AND QUALITATIVELY, THE OUTPUTS AND OUTCOMES OF NSF-FUNDED RESEARCH TO SUPPORT ITS STATUTORY MISSION
All respondents

The next questions are about the types of data NSF collects about *awarded projects* to evaluate how well the merit review process is helping NSF to meet its goals to advance scientific knowledge and benefit society.

Think about the outputs and outcomes of research that NSF asks principal investigators to report on in their annual and final reports, such as project accomplishments, products, and impact. Also consider evidence from third party evaluations, program monitoring systems, and/or NSF Education and Training Application (ETAP).

39. How effective do you think these data are for helping NSF to assess whether funded projects support NSF's mission to *advance scientific knowledge*?

Select one only

- ❍ Not at all effective 1
- ❍ A little effective..... 2
- ❍ Somewhat effective 3
- ❍ Very effective 4

HOW NSF CURRENTLY DEMONSTRATES, BOTH QUANTITATIVELY AND QUALITATIVELY, THE OUTPUTS AND OUTCOMES OF NSF-FUNDED RESEARCH TO SUPPORT ITS STATUTORY MISSION
All respondents

40. How effective do you think these data are for helping NSF to assess whether funded projects support NSF's mission to *benefit society*?

Select one only

- ❍ Not at all effective 1
- ❍ A little effective..... 2
- ❍ Somewhat effective 3
- ❍ Very effective 4

HOW NSF CURRENTLY DEMONSTRATES, BOTH QUANTITATIVELY AND QUALITATIVELY, THE OUTPUTS AND OUTCOMES OF NSF-FUNDED RESEARCH TO SUPPORT ITS STATUTORY MISSION
All respondents

41. What, if any, additional information do you think NSF should be collecting to monitor their progress toward *advancing scientific knowledge* that is not being collected?

Please include both quantitative and qualitative outputs and outcomes if relevant.

(STRING 4,000)

HOW NSF CURRENTLY DEMONSTRATES, BOTH QUANTITATIVELY AND QUALITATIVELY, THE OUTPUTS AND OUTCOMES OF NSF-FUNDED RESEARCH TO SUPPORT ITS STATUTORY MISSION
All respondents

42. What, if any, additional information do you think NSF should be collecting to monitor their progress toward *benefiting society* that is not being collected?

Please include both quantitative and qualitative outputs and outcomes if relevant.

(STRING 4,000)

OVERALL PERSPECTIVE ON MERIT REVIEW PROCESS AND CRITERIA
All respondents

43. Is there anything else about the merit review process or policies that you think would be helpful for NSB-NSF Commission on Merit Review (MRX) to know?

(STRING 4,000)

BACKGROUND INFORMATION
All respondents

The next questions request demographic information and are for statistical purposes only. Your responses are voluntary.

Please pick the category or categories that you feel best describe yourself. You may also select the option to not specify a category for each question.

44. Which directorate or office is your position in?

Select one only

- ❍ Biological Sciences (BIO) 1
- ❍ Computer and Information Science and Engineering (CISE) 2
- ❍ Engineering (ENG) 3
- ❍ Geosciences (GEO) 4
- ❍ Mathematical and Physical Sciences (MPS) 5
- ❍ Social, Behavioral and Economic Sciences (SBE) 6
- ❍ STEM Education (EDU) 7
- ❍ Technology, Innovation and Partnerships (TIP) 8
- ❍ Integrative Activities (OIA) 9
- ❍ International Science and Engineering (OISE) 10
- ❍ I do not wish to provide this information 0

BACKGROUND INFORMATION
All respondents

45. How long have you held your current NSF position?

Select one only

- ❍ Less than 1 year 1
- ❍ 1 – 3 years 2
- ❍ 3 – 5 years 3
- ❍ 5 – 10 years 4
- ❍ More than 10 years 5
- ❍ I do not wish to provide this information 0

BACKGROUND INFORMATION
All respondents

46. In total, how long have you worked for NSF?

Select one only

- ❍ Less than 1 year 1
- ❍ 1 – 3 years 2
- ❍ 3 – 5 years 3
- ❍ 5 – 10 years 4
- ❍ 10 – 15 years 5
- ❍ More than 15 years 6
- ❍ I do not wish to provide this information 0

DEMOGRAPHIC INFORMATION
All respondents

47. How do you currently describe yourself?

Select all that apply

- ☐ Female 1
- ☐ Male 2
- ☐ Transgender 3
- ☐ I use a different term (please specify) 99

Specify [ ] (STRING 100)

- ❍ I do not wish to provide this information 0

DEMOGRAPHIC INFORMATION
All respondents

48. Are you Hispanic or Latino?

Select one only

- ❍ Yes, I am Hispanic or Latino 1
- ❍ No, I am not Hispanic or Latino 2
- ❍ I do not wish to provide this information 0

DEMOGRAPHIC INFORMATION
All respondents

49. What is your racial background?

Select all that apply

- ☐ American Indian or Alaska Native 1
- ☐ Asian 2
- ☐ Black or African American 3
- ☐ Native Hawaiian or Other Pacific Islander 4
- ☐ White 5
- ☐ Other (please specify) 99

Specify [ ] (STRING 100)

- ❍ I do not wish to provide this information 0

DEMOGRAPHIC INFORMATION
All respondents

50. Do you identify as having a disability?

Select *yes* if any of the following apply to you:

- **Deaf or serious difficulty hearing.**
- **Blind or serious difficulty seeing even when wearing glasses.**
- **Serious difficulty walking or climbing stairs.**
- **Other disability related to a physical, mental, or emotional condition.**

- ❍ Yes 1
- ❍ No 0
- ❍ I do not wish to provide this information 2

NSF Leadership Interview Protocol

1. First, can you describe how you are involved with the merit review process?

Slide: Merit review criteria definitions

- Intellectual Merit: the potential for a proposed project to advance knowledge
- Broader Impacts: the potential for a proposed project to benefit society and contribute to the achievement of specific, desired societal outcomes

2. I would like to hear about how you understand the goals of merit review. What do you think NSF is trying to accomplish with the Intellectual Merit and Broader Impacts criteria?

3. The PAPPG notes that the two merit review criteria are to be given full consideration during the review and decision-making processes. Each criterion is necessary but neither, by itself, is sufficient. In practice, how is full consideration achieved in your leadership of the [directorate's/office's] portfolio?

   a. Do you consider one to be more important than the other, and why or why not?

4. What instructions or training do you provide division directors, program officers, and reviewers to help them interpret and apply the merit review criteria?

   a. We are interested in understanding more about the resources available for individuals who have questions about the merit review criteria. If division directors have questions about the merit review criteria or are unclear about how to apply them, how would they go about asking for clarification?

      i. What about program officers?

      ii. What about reviewers?

5. Tell me a bit about applying the merit review criteria in your [directorate/office] or program specifically. What, if anything, about the scientific field of your directorate, division, or program makes it challenging to interpret and apply the merit review criteria?

Slide: Merit review principles and elements

Principles

- All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.
- NSF projects, in the aggregate, should contribute more broadly to achieving societal goals. These broader impacts may be accomplished through the research itself, through activities that are directly related to specific research projects, or through activities that are supported by, but are complementary to, the project. The project activities may be based on previously established and/or innovative methods and approaches, but in either case must be well justified.
- Meaningful assessment and evaluation of NSF funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of broader impacts and the resources provided to implement projects. If the size of the activity is limited, evaluation of that activity in isolation is not likely to be meaningful. Thus, assessing the effectiveness of these activities may best be done at a higher, more aggregated, level than the individual project.

Elements

- What is the potential for the proposed activity to:
- Advance knowledge and understanding within its own field or across different fields (Intellectual Merit); and Benefit society or advance desired societal outcomes (Broader Impacts)?
- To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?
- Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?
- How well qualified is the individual, team, or organization to conduct the proposed activities?
- Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?

6. Thinking specifically about the review of proposals submitted to or managed by your [directorate/office], how are the merit review principles and elements used to identify which projects to fund?

   a. Specifically, how are the principles and elements used by program officers to make funding recommendations in your directorate?

   b. How are the principles and elements used by division directors in their portfolio management?

   c. In your experience, are these uses similar to or different from how the principles and elements are used in other directorates and offices?

7. How confident are you that the principles and elements are used consistently within your Directorate? Can you tell me about where you think there are areas of consistency and inconsistency?

8. We know that the COV reports are a key source of feedback to support continuous improvement of Merit Review practices. Can you walk me through how you use COV feedback within your [directorate/office] and anything from the most recent COV report that you are currently working on in your current role?

9. The Proposal & Award Policies & Procedures Guide notes that merit review criteria and process are designed to inform NSF investments "in a robust and diverse portfolio of projects that creates new knowledge and enables breakthroughs in understanding across all areas of science and engineering research and education". In your opinion, are there any other important factors for evaluating proposals are not captured by the merit review criteria?

   a. *If YES:* Tell me a bit more about this factor and why you think it would be important for reviewing proposals. What could this capture that is missed by the current merit review criteria?

10. Who determines which NSF staff have the necessary expertise of the merit review criteria to evaluate proposals?

   a. How do they make that determination?

11. Who determines which reviewers have the necessary expertise of the merit review criteria to evaluate proposals?

   a. How do they make that determination?

12. In your [directorate/office], how do you evaluate whether program officers are reviewing proposals and making recommendations that are aligned to the merit review criteria?

   a. What about reviewers?

13. How do you evaluate whether division directors are making decisions about whether to award or decline proposals in a way that is aligned to the merit review criteria?

14. How are reviewers' assessments of each merit review criterion factored into portfolio management within your [directorate/office]?

15. Do you believe that the merit review policies support a fair and accurate assessment of a proposal's merit? Why or why not?

16. Are there any aspects of the merit review process that you think introduce unfair bias into the evaluation of proposals? If so, how?

   a. What, if anything, does your [directorate/office] implement to mitigate potential biases?

17. Do you think the merit review process increases or decreases the diversity of proposals submitted for review? How so?

   a. Probe: Ask for diversity of ideas addressed in proposals, characteristics of PIs and institutions if not mentioned.

18. Do you think the merit review process increases or decreases the diversity of awarded projects? How so?

   a. Probe: Ask for diversity of ideas addressed in projects, characteristics of PIs and institutions if not mentioned.

19. What types of evidence does your [directorate/office] collect and examine to assess whether the projects that received funding helped support NSF's mission to advance scientific knowledge and benefit society? This might include evidence from third-party evaluations, program monitoring systems, or ETAP [NSF Education and Training Application]. *If needed, probe for both quantitative and qualitative outputs and outcomes.*

20. Is there additional information that NSF could be collecting to monitor its progress to advance scientific knowledge and benefit society?

   a. Is there anything about the *quality* of the data that NSF is collecting that might make it difficult to use to monitor its progress?

21. Is there anything else about the merit review process or the two criteria that you think would be helpful for NSB-NSF Commission on Merit Review (MRX) to know?

Request for Information

INFORMATION REQUESTS

Please refer to definitions provided at the end of this letter for terms used in these Information Requests.

1. The MRX is interested in identifying opportunities to improve NSF's current Merit Review criteria, policy, and processes. Importantly, this includes documenting and understanding any areas of misunderstanding, gaps, or lack of clarity regarding (a) the three Merit Review Principles which are the foundations of the Merit Review Process, (b) the two statutory Merit Review Criteria which are used to evaluate all proposals to NSF, and (c) the five Merit Review Elements NSF uses to assess each criterion; for instance: Are the Principles, Criteria, and Elements clear? Could they be improved upon? The MRX welcomes feedback on any or all of these, and particularly on the Broader Impacts Criterion. Chapter 3 of NSF's Proposal & Award Policies and Procedures Guide (PAPPG) defines terms in this Information Request. See https://new.nsf.gov/policies/pappg/24-1/ch-3-proposal-processing-review#a-merit-review-principles-and-criteria-af2
   Individuals responding to this request are encouraged to indicate whether their perspectives are informed by experience(s) preparing and/or reviewing proposals to NSF.

2. NSF strives to conduct a fair, competitive, transparent Merit Review process for the selection of projects. To accomplish this, NSF relies on a process that considers both the technical aspects of a proposed project and its potential to contribute more broadly to advancing NSF's mission using the statutory Intellectual Merit and Broader Impacts Merit Review Criteria. MRX invites suggestions on the implementation of the Merit Review Criteria. We especially invite feedback that would (a) clarify how the Merit Review Criteria can be used in preparing and reviewing proposals, (b) ensure proposals, reviews, and funding decisions demonstrate full consideration of both criteria while maintaining openness to the full spectrum of potential activities under each, and (c) better recognize and support potentially transformative and high-risk/high-reward activities.
   Individuals responding to this request are encouraged to indicate whether their perspectives are informed by experience(s) preparing and/or reviewing proposals to NSF.

3. MRX is interested in the experiences and perspectives of those who have considered submitting and/or submitted proposals in the past. We invite you to share your insights and describe any opportunities you believe would improve implementation of the Merit Review criteria, policy, and processes based on your experience as a proposer or investigator. This includes any experiences that may have encouraged or dissuaded you from submitting proposals to NSF. We are especially interested in learning (a) how NSF guidance (e.g., as provided in the NSF PAPPG, program solicitations, or other funding opportunity announcements), may have played a part in your decision(s) whether to submit proposals, and (b) how NSF might best support investigators interested in submitting a proposal to NSF.
   Individuals responding to this request are encouraged to indicate whether they submitted or decided not to submit a proposal, and whether these experiences occurred within the past five years.

4. MRX is interested in the experiences and perspectives of those who have reviewed proposals submitted to NSF. We invite you to share your insights and describe any opportunities you believe would improve implementation of the Merit Review criteria, policy, or processes based on your experience reviewing NSF proposals.
   Individuals responding to this request are encouraged to indicate whether they served on a panel and/or as ad hoc reviewers, and whether these experiences occurred within the past five years.

5. MRX is interested in exploring how NSF could better support awardees in demonstrating and documenting outcomes of their awards in advancing knowledge (Intellectual Merit) and benefiting society and contributing to the achievement of desired broader or societal outcomes (Broader Impacts). We invite you to share your insights on how NSF might better support awardees in demonstrating and documenting outcomes of their awards without unnecessarily increasing awardees' administrative burden of reporting.
   Individuals responding to this request are encouraged to indicate whether their suggestions are based on experiences as investigators, users of public outcomes reports, or another perspective.

6. MRX welcomes any other comments on or suggestions for improving NSF's current Merit Review criteria, policy, and processes. It also welcomes information about aspects of Merit Review criteria, policy, and processes that are currently working well.

DEFINITIONS FOR TERMS USED IN THIS RFI

Merit Review Policy—

Principles

1. All NSF projects should be of the highest quality and have the potential to advance, if not transform, the frontiers of knowledge.
2. NSF projects, in the aggregate, should contribute more broadly to achieving societal goals. These broader impacts may be accomplished through the research itself, through activities that are directly related to specific research projects, or through activities that are supported by, but are complementary to, the project. The project activities may be based on previously established and/or innovative methods and approaches, but in either case must be well justified.
3. Meaningful assessment and evaluation of NSF funded projects should be based on appropriate metrics, keeping in mind the likely correlation between the effect of broader impacts and the resources provided to implement projects. If the size of the activity is limited, evaluation of that activity in isolation is not likely to be meaningful. Thus, assessing the effectiveness of these activities may best be done at a higher, more aggregated, level than the individual project.

Criteria

Both criteria are to be given full consideration during the review and decision-making processes; each criterion is necessary but neither, by itself, is sufficient. Therefore, proposers must fully address both criteria.

- Intellectual Merit (IM): the potential for a proposed project to advance knowledge.
- Broader Impacts (BI): the potential for a proposed project to benefit society and contribute to the achievement of specific, desired societal outcomes.

Elements

1. What is the potential for the proposed activity to:
    a. Advance knowledge and understanding within its own field or across different fields (Intellectual Merit); and
    b. Benefit society or advance desired societal outcomes (Broader Impacts)?
2. To what extent do the proposed activities suggest and explore creative, original, or potentially transformative concepts?
3. Is the plan for carrying out the proposed activities well-reasoned, well-organized, and based on a sound rationale? Does the plan incorporate a mechanism to assess success?
4. How well qualified is the individual, team, or organization to conduct the proposed activities?

5. Are there adequate resources available to the PI (either at the home organization or through collaborations) to carry out the proposed activities?

This description of NSF's Merit Review policy is from NSF's 2024 Proposal and Award Policies and Procedures Guide (PAPPG), part I, chapter 3. https://new.nsf.gov/policies/pappg/24-1.

Transformative Research

Transformative research is defined as research driven by ideas that have the potential to radically change our understanding of an important existing scientific or engineering concept or leading to the creation of a new paradigm or field of science or engineering. Such research also is characterized by its challenge to current understanding or its pathway to new frontiers. See NSB's statement Enhancing Support of Transformative Research at NSF: https://www.nsf.gov/pubs/2007/nsb0732/nsb0732.pdf.

Broadening Participation

"Broadening participation in STEM" is the comprehensive phrase NSF uses to refer to the Foundation's goal of increasing the representation and diversity of individuals, organizations, and geographic regions that contribute to STEM education, research, and innovation. To broaden participation in STEM, it is necessary to address issues of equity, inclusion, and access in STEM education, training, and careers. Whereas all NSF funding programs might support broadening participation components, some funding programs primarily focus on supporting broadening participation research and projects. Examples can be found on the NSF Broadening Participation in STEM website. See https://new.nsf.gov/funding/initiatives/broadening-participation, and the NSF PAPPG, Introduction, https://new.nsf.gov/policies/pappg/24-1.

Mathematica Inc.

Our employee-owners work nationwide and around the world.

Find us at **mathematica.org** and **edi-global.com**.

National Science Board

NSB-2025-23

Alexandria, Virginia
703-292-7000
www.nsf.gov/nsb