



National Science Foundation
Evaluation and Assessment Capability

Assessment of Stakeholder Experiences With NSF's Merit Review Process

Findings From the 2021 Merit
Review Survey

About the Evaluation and Assessment Capability Section

[The Evaluation and Assessment Capability \(EAC\)](#) Section bolsters NSF efforts to make informed decisions and promote a culture of evidence. Located in the Office of Integrative Activities of the Office of the Director, EAC provides centralized technical support, tools, and resources to conduct evidence-building activities and to build capacity for evidence generation and use across the agency. EAC is led by NSF's Chief Evaluation Officer.

About this report

This report was prepared for EAC under contract number 49100420C0012 and OMB clearance number 3145-0257. The views expressed are those of the authors and should not be attributed to NSF, nor does mention of trade names, commercial products, or organizations imply endorsement of same by the U.S. Government.

Preferred citation

Bowman, M., Bryant, A., Griffiths, R., Hare, A., Huey, L., McCall, J., Scanlon, J., and Wakar, B. (2023). *Assessment of Stakeholder Experiences With NSF's Merit Review Process: Findings From the 2021 Merit Review Survey*. Alexandria, VA: National Science Foundation.

NSF Quality Certification: Level 1

Quality Certifications

Level 1 — The author(s)/contractor(s) are responsible for the quality and conclusions presented in this report

Level 2 — NSF verified that this report underwent quality assurance procedures and contributed to assessing its content

Level 3 — NSF independently reproduced the analysis presented in this report



Assessment of Stakeholder Experiences With NSF's Merit Review Process

Findings From the 2021 Merit Review Survey

Marietta Bowman	Senior Study Director	mbowman@insightpolicyresearch.com
Andrew Bryant	Data Analyst	abryant@insightpolicyresearch.com
Richard Griffiths	Senior Statistician	rgriffiths@insightpolicyresearch.com
Amanda Hare	Associate Director	ahare@insightpolicyresearch.com
Lyndsay Huey	Analyst	lhuey@insightpolicyresearch.com
James McCall	Survey Researcher	jmccall@insightpolicyresearch.com
Jeffrey Scanlon	Data Scientist	jscanlon@insightpolicyresearch.com
Breanna Wakar	Senior Data Scientist	bwakar@insightpolicyresearch.com

Submitted to:

National Science Foundation

2415 Eisenhower Ave.

Alexandria, VA 22314

Submitted by:

RIVA Solutions, Inc.

1676 International Drive, Suite #520

McLean, VA 22102

Contents

Contents.....	2
1 Executive Summary	5
1.1 Study Overview	5
1.2 Summary of Key Findings	5
1.3 Implications.....	7
2 Introduction	8
2.1 Background on the Merit Review Process.....	8
2.2 Four Dimensions of Applicant and Reviewer Experience.....	9
3 Data and Analytic Methods	10
3.1 Data Sources.....	10
3.1.1 Administrative data.....	10
3.1.2 2019 survey data.....	10
3.1.3 2021 survey data.....	11
3.2 Data Management and Analytic Approach	12
3.2.1 Data cleaning and variable construction	12
3.2.2 Nonresponse bias and weighting adjustments	15
3.2.3 Multiple regression: examination of group differences	15
3.3 Limitations	16
3.3.1 Response rate and generalizability	16
3.3.2 Small cell sizes.....	16
3.3.3 Conclusions about differences by race and ethnicity	17
3.3.4 Limits of the regression modeling approach.....	18
3.3.5 Comparison of 2021 and 2019 data	18
4 Assessing the 2021 Merit Review Process	20
4.1 Descriptive Statistics	20
4.2 Satisfaction With the Merit Review Process	22
4.2.1 Overall satisfaction among all respondents.....	22
4.2.2 Applicant satisfaction index scores	27
4.2.3 Reviewer satisfaction index scores	31
4.3 Burden in the Merit Review Process	36
4.3.1 Applicant burden	36
4.3.2 Reviewer burden	39
4.4 Proposal and Review Quality in the Merit Review Process	41
4.4.1 Primary factors that would improve the merit review process	41

4.4.2	Applicant perceptions of quality.....	43
4.4.3	Reviewer perceptions of quality	48
4.5	Fairness in the Merit Review Process	52
4.5.1	Overall perceptions of fairness.....	52
4.5.2	Applicant perceptions of fairness	54
4.5.3	Reviewer perceptions of fairness	56
5	Conclusions.....	58
5.1	Summary of Key Findings	58
5.1.1	Satisfaction with the merit review process	58
5.1.2	Assessment of burden of the merit review process.....	58
5.1.3	Perceptions of quality within the merit review process	58
5.1.4	Perceptions of fairness within the merit review process.....	58
5.1.5	Trend differences compared with 2019 survey	59
5.2	Implications.....	59
5.2.1	Facilitating applicants' exposure to or transparency around the merit review process may increase satisfaction and perceptions of fairness.....	59
5.2.2	Redesigning future survey sampling and recruitment efforts may enhance the conclusions that can be drawn by population subgroups and other relevant process factors	59
5.2.3	Examination or expansion of efforts to support underrepresented groups may be warranted.....	60
5.2.4	Increasing awareness of available accommodations and supports for individuals with disabilities may increase satisfaction and perceptions of fairness	60
5.2.5	Establishment of ongoing assessment methods can help determine how satisfaction with and perceptions of the merit review process change over time	61
5.2.6	Investigate if applicant funding success may explain differences between satisfaction with and perceptions of the merit review process across subgroups	61
	Acknowledgements.....	62
	References.....	62
	Appendix A. Supplemental Information on Data	A-1
A.1	Administrative Data	A-1
A.2	2021 Survey Instrument	A-4
A.3	2021 Survey Data Collection Methods	A-17
	Appendix B. Supplemental Information on Analytic Approach.....	B-1
B.1	2021 Survey Data Cleaning and Variable Construction	B-1
	<i>Addressing Small Cell Sizes</i>	B-1
	<i>Variable Construction</i>	B-2
B.2	Nonresponse Bias Analysis.....	B-3

B.3	Correcting for Nonresponse Bias With Cell Weighting	B-8
B.4	Multiple Regression Equations	B-9
Appendix C.	Supplemental Findings	C-1
C.1	Descriptive Statistics	C-1
C.2	Satisfaction Dimension: Regression Tables	C-12
C.3	Burden: Regression Tables	C-20
C.4	Quality: Regression Tables	C-39
C.5	Fairness: Regression Tables.....	C-59

1 Executive Summary

The National Science Foundation (NSF) receives over 43,000 requests for science and engineering research funding each year (NSF 2022a, NSF 2022b). To ensure the highest quality research is funded, every proposal NSF receives undergoes a rigorous review process to assess its merit—known as the merit review process. NSF contracted with RIVA Solutions, Inc. and Insight Policy Research (Insight), together referred to as Team RIVA, to conduct the 2021 Merit Review Survey of individuals who have submitted proposals to NSF (applicants) and/or reviewed proposals on behalf of NSF (reviewers) via the merit review process.

1.1 Study Overview

This study's primary goals are to (1) understand applicants' and reviewers' overall satisfaction and perceptions of burden, quality, and fairness of the merit review process and (2) examine how these experiences compare with the feedback from the 2019 Merit Review Survey. To achieve these goals, the study team examined research questions across four key dimensions:

- **Satisfaction:** To what extent are individuals satisfied with the merit review process? To what extent has satisfaction with the merit review process changed since 2019?
- **Burden:** What is the burden associated with the merit review process? To what extent has the burden associated with the merit review process changed since 2019?
- **Quality:** What are the perceptions of proposal and review quality? To what extent have perceptions of proposal and review quality changed since 2019?
- **Fairness:** What are the perceptions of fairness of the merit review process? To what extent have perceptions of fairness of the merit review process changed since 2019?

The study team used three data sources to address these research questions. In fall 2021, Insight surveyed more than 17,000 individuals who submitted proposals to NSF and/or reviewed proposals on behalf of NSF during fiscal year (FY) 2019 and FY 2020 to assess the current status of reviewers' and principal investigators' (PIs) perceptions of and experiences with the merit review process. The study team supplemented the 2021 survey data with data from a prior iteration of the survey conducted in 2019 to examine changes in perceptions and experiences over time. The team used NSF administrative data to identify the survey universe and examine the potential for nonresponse bias.

1.2 Summary of Key Findings

In general, assessments of satisfaction and fairness were positive. Over half of respondents were satisfied with the merit review process, and two out of three perceived the process to be fair. About three of four reviewers indicated the majority of proposals they reviewed in recent years have been of high quality. However, less than half of applicants agreed or strongly agreed that the quality of written reviews was high. Applicants reported submitting an average of three proposals between October 1, 2018, and when they completed the survey in fall 2021. They also reported spending an average of 126 hours preparing their most recent proposal. Reviewers reported reviewing an average of 10 proposals since October 1, 2018, and spending an average of 7 hours reading, writing, and submitting their most recent review. (These reported review hours do not include time traveling to or participating in a panel discussion.)

Several patterns emerged for subgroups of respondents across the four research dimensions. Respondents with reviewer experience were more likely to be satisfied with the review process and rated its quality and fairness more highly. Asian respondents were more satisfied with the

review process and rated the quality of written reviews more highly than White, non-Hispanic respondents, though Asian respondents rated the quality of proposals less highly.

Detailed analyses of subgroups for each dimension provided further insights into this variation:

- **Satisfaction** varied by role in the merit review process, institutional affiliation, disability status, and race and ethnicity. Respondents who had only reviewed proposals were more satisfied than respondents who had both reviewed and submitted proposals while those who had only submitted proposals were less satisfied than those who had done both. Respondents from non-R1¹ minority-serving institutions (MSIs) were more satisfied than their peers, and respondents with disabilities were less satisfied than those without a disability. Asian, non-Hispanic respondents were more satisfied with the merit review process than White, non-Hispanic respondents. In contrast, respondents who identified as American Indian, Native Hawaiian, or selected “other” were less satisfied than White, non-Hispanic respondents.
- **Burden** varied by the NSF directorate or office with which respondents were affiliated. Applicants affiliated with the directorates of Computer and Information Science and Engineering, Education and Human Resources,² and Engineering submitted more proposals and spent more hours preparing them than their peers in other directorates. Reviewers affiliated with these same directorates also conducted more reviews but spent fewer hours preparing them relative to reviewers in other directorates.
- **Perceptions of quality** varied by role in the merit review process, NSF directorate or office, gender, and race and ethnicity. Applicants who also reviewed proposals rated the quality of written reviews more highly than those who had not. Reviewers who also submitted proposals rated the quality of proposals lower than those who did not have this experience. Female respondents rated the quality of written reviews and proposals more highly than males, while Asian respondents rated the quality of written reviews higher but proposals lower than respondents in other racial or ethnic groups.
- **Perceptions of fairness** varied by race and role in the merit review process. Hispanic and Black, non-Hispanic respondents were less likely than White, non-Hispanic respondents to think the process was fair. Respondents who had only reviewed proposals were most likely to think the process was fair. Those who had only submitted proposals were least likely to share this view.

Respondent satisfaction, perceptions of fairness with the merit review process, and perceptions of review quality were lower in 2021 compared with 2019. Respondents were also significantly more likely to report declining a request to serve as a reviewer in 2021 than in 2019. This study did not examine changes in perceived quality of written reviews or changes in burden because of changes made to the survey items between 2019 and 2021.

¹ R1 represents doctoral universities with very high research activities, as defined by the Carnegie Classification of Institutions of Higher Education (N.d.).

² Education and Human Resources directorate is now known as STEM Education directorate.

1.3 Implications

Based on these findings, Team RIVA identified six potential implications for future policy and research:

- Survey findings demonstrated consistent differences in the experiences of applicants who had also served as reviewers and those who had not. Most notably, individuals who were applicants only were less satisfied with the merit review process and perceived the process to be less fair. Facilitating applicants' exposure to or transparency around the merit review process may increase satisfaction and perceptions of fairness.
- The response rate for the survey (27 percent) may limit the conclusions the study team could draw for some subpopulations. Subgroup analyses were limited by small sample sizes, particularly for small race and ethnicity groups. Alternative survey designs that increase response rates and sample sizes for subpopulations would help further mitigate bias and improve subgroup analyses and the reliability of the data.
- Patterns of survey findings by institutional affiliation suggested respondents affiliated with non-R1/MSI may be more satisfied with the merit review process, both as applicants and reviewers, than respondents affiliated with the more commonly funded R1/non-MSIs. However, Black applicants and reviewers were less satisfied with the merit review process than their White counterparts. Examination of efforts to support underrepresented groups, such as outreach to non-R1/MSI institutions, may be warranted to understand the effects of these activities and opportunities for their expansion to other underrepresented populations.
- Survey results showed that respondents with disabilities were less satisfied with the merit review process and less likely to perceive the merit review process as fair than their counterparts who did not report having a disability. Increasing awareness of available accommodations and supports for individuals with disabilities may improve satisfaction and perceptions of fairness.
- Survey results showed that respondents were generally less satisfied with the merit review process in 2021 than in 2019. However, the global health pandemic and social unrest around race and politics that occurred over this time made it difficult to determine what additional factors outside NSF's control may have affected this decrease in satisfaction. Establishing ongoing methods to assess how satisfaction with and perceptions of the merit review process change over time would provide more targeted data on how NSF procedures and policies relate to applicant and reviewer experiences.
- This study does not control for overall applicant success rate or the most recent proposal award outcome when examining respondent satisfaction and perceptions of burden, quality, and fairness. However, applicant success could be related to these outcomes to some degree. Additional analyses of the 2021 Merit Review Survey data may benefit from understanding if applicant funding success explains the observed differences in subgroups, such as applicants' and reviewers' perceptions of the merit review process experience.

2 Introduction

Founded by Congress in 1950, the National Science Foundation (NSF) is an independent federal agency charged with supporting research and education in all nonmedical fields of science and engineering, at all levels of education. To achieve its primary goal to “promote the progress of science; [and] to advance the national health, prosperity, and welfare” within the United States (NSF 2022b:1), NSF issues funding to colleges, universities, K–12 school systems, businesses, and other institutions to advance science and engineering research. In FY 2021, NSF received over 43,000 competitive proposals for funding and selected approximately 11,000 for funding. It ultimately awarded \$8.1 billion in funding to support more than 300,000 people at approximately 1,900 entities (NSF 2022a, NSF 2022b). To ensure the highest quality research is funded, every proposal NSF receives undergoes a rigorous review process to assess its merit—known as the merit review process.

In alignment with the strategic goal outlined in the 2018-2022 Strategic Plan to “excel at NSF operations and management” (NSF 2022c:28), NSF contracted with RIVA Solutions, Inc. and Insight Policy Research (referred to as Team RIVA in this report) to conduct the 2021 Merit Review Survey. Conducted biennially since 2015, this survey collects information on the experiences and perceptions of individuals who have submitted proposals to NSF (applicants) and/or reviewed proposals on behalf of NSF (reviewers) via the merit review process.

This study’s primary goals are to (1) understand applicants’ and reviewers’ overall satisfaction and perceptions of quality, fairness, and burden and (2) examine how these experiences compared to the feedback from the 2019 Merit Review Survey (Hare et al. 2020).

2.1 Background on the Merit Review Process

NSF strives to fund the highest quality proposals to advance the interests of science and engineering research in the United States. The merit review process is central to that goal—it ensures funding decisions are based upon fair, unbiased, and thoughtful deliberation of ideas (NSF 2022b). As outlined in NSF’s *Proposal and Award Policies and Procedures Guide*, nearly all proposals NSF receives must undergo merit review and are evaluated against two National Science Board-approved criteria: intellectual merit and broader impacts (NSF 2021a). Intellectual merit refers to the potential of the proposed research to advance knowledge, whereas broader impacts refer to its potential to benefit society more broadly and contribute to the achievement of desired societal outcomes (NSF 2021a). Individual programs may have additional review criteria that must be met, which are specified within the program’s solicitation documentation (NSF 2021a). The merit review process is composed of three phases:

- 1) Proposal preparation and submission:** NSF announces funding opportunities and receives both unsolicited and solicited proposals. Applicants submit their proposals electronically to NSF according to the guidance provided in the *NSF Proposal Award Policies & Procedures Guide* (PAPPG) (NSF 2021a). NSF reviews the proposals for compliance with PAPPG requirements, and if confirmed, the proposals are assigned to the NSF program officer (PO) who operates within the appropriate program.

- 2) Proposal review and processing:** The PO determines which type of review the proposals will undergo (individual ad hoc review, review by a panel of peer reviewers, or a combination of both). The PO then identifies experts in the relevant field and invites them to assess proposals that align with their specific expertise and provide written reviews of each proposal's strengths and weaknesses. Reviewers are identified based on the PO's knowledge of the research area, references listed in the proposal, review of journal articles on related topics, or suggestions from applicants. All proposals are reviewed to assess the broader impacts and intellectual merit of the proposed work and any additional review criteria specified in a solicitation. Following expert reviews, the NSF PO synthesizes written reviews and reflects on panel discussions (if applicable) to issue funding recommendations while balancing these decisions against program's funding portfolios and priorities. Proposals recommended for funding are forwarded to the Division Director or another NSF official for additional review.
- 3) Award processing:** Each proposal recommended for award undergoes an administrative review by NSF's Office of Budget, Finance, and Award Management. If it passes this review, the proposal is awarded.

The NSF merit review process ensures a competitive and transparent evaluation of proposals to identify research endeavors with the most potential to advance knowledge and contribute to achieving societal goals (NSF 2022a; NSF 2022b). The success of the merit review process relies on the assumptions that scientists will continue to submit their ideas for cutting-edge research and experts in their respective fields will continue to provide high-quality reviews of those proposals. Applicants and reviewers may not choose to participate in the merit review process if they become dissatisfied with the process or it becomes too burdensome, the quality of proposals and reviews is not high, or scientists perceive the process to be unfair, jeopardizing merit review. NSF conducts a survey of applicants and reviewers biannually to assess their satisfaction with the process.

2.2 Four Dimensions of Applicant and Reviewer Experience

This report describes the findings from the 2021 Merit Review Survey across four dimensions characterizing the merit review process: satisfaction, burden, quality, and fairness. Analyses evaluate whether experiences within these four dimensions vary by race and ethnicity, gender, disability status, NSF directorate, institution type, early career status, and experience as an applicant and/or a reviewer. Findings from 2021 are compared with those from the prior survey, conducted in 2019. The primary research questions that guide this study follow:

- **Satisfaction:** To what extent are individuals satisfied with the merit review process? To what extent has satisfaction with the merit review process changed since 2019?
- **Burden:** What is the burden associated with the merit review process? To what extent has the burden associated with the merit review process changed since 2019?
- **Quality:** What are the perceptions of proposal and review quality? To what extent have the perceptions of proposal and review quality changed since 2019?
- **Fairness:** What are the perceptions of fairness of the merit review process? To what extent have the perceptions of fairness of the merit review process changed since 2019?

3 Data and Analytic Methods

This study used a repeated cross-sectional design and descriptive analyses to provide an overview of applicant and reviewer perceptions of the 2021 merit review process. Analysis focused on four research dimensions touched on in the 2021 survey questionnaire: satisfaction, burden, quality, and fairness; each dimension included multiple survey items. Analyses were conducted for the survey population overall, by role (applicant or reviewer), and by seven population characteristics that align with NSF's strategic objective to ensure accessibility and inclusivity (NSF, 2020c). These characteristics were race and ethnicity, gender, disability status, primary NSF directorate, institutional affiliation, and early career status. The team also conducted year-over-year analyses to compare responses from the 2021 Merit Review Survey with those from the Merit Review Survey conducted in 2019. Section 3.2 includes full descriptions of these analyses.

3.1 Data Sources

This study drew on three data sources to evaluate applicant and reviewer satisfaction, burden, quality, and fairness: (1) NSF administrative data, (2) the 2019 NSF Merit Review Survey, and (3) the 2021 Merit Review Survey.

3.1.1 Administrative data

The team collected administrative data from NSF data systems to identify the universe for the 2021 survey. The universe was restricted to individuals who had served as applicants or reviewers in FY 2019 or FY 2020. Records were first deduplicated by NSF reviewer and applicant IDs and then deduplicated by generating a set of keys (combinations of descriptive variables) for each record. The team used these keys as points of comparison to determine if two records belonged to the same individual even if they had been assigned different reviewer or applicant IDs. After records deduplication, 66,585 unique individuals in the final universe file were eligible to receive the survey invitation (see appendix A.1 for additional details on administrative data variables). In addition to establishing the universe for the 2021 survey, the administrative data were also used to conduct an analysis of the potential for nonresponse bias, as described further in section 3.2. In March 2023, NSF asked the study team to pull updated administrative data on applicant demographics from NSF's DataLake.³ The study team used these new data to update the administrative race and ethnicity, disability status, and gender variables previously constructed and the nonresponse bias analyses and survey weights. See appendix table A.1.3 for additional details on demographic variables pulled from NSF's DataLake.

3.1.2 2019 survey data

Data from the 2019 Merit Review Survey were used to conduct year-over-year analyses comparing 2019 and 2021 survey results across dimensions. For the 2019 Merit Review Survey, NSF provided 146,129 email addresses for all individuals who reviewed proposals or submitted proposals to NSF between FY 2016 and FY 2018. After removing invalid email addresses, NSF email addresses, and emails associated with duplicate reviewer or principal investigator (PI) observations, the final contact list consisted of 89,012 unique reviewers and PIs. As in the 2021 survey, questions in the 2019 survey focused on perceptions of satisfaction, burden, quality, and fairness. Data were collected in September and October 2019. The overall

³ DataLake is a SQL server containing most of NSF's business data (proposals, panels, awards) and proposal content.

response rate was 30 percent (26,216 respondents). Details on survey questions, methodology, and findings can be found in the *2019 Merit Review Report* (Hare et al. 2020).

3.1.3 2021 survey data

Data from the 2021 Merit Review Survey were used to answer the primary research questions for 2021 and conduct year-over-year comparisons to the 2019 survey data. Although most of the survey items were retained from the 2019 instrument to enable cross-year analyses, the team made minor modifications to the 2021 instrument to clarify survey question wording or drop questions that were not relevant for the planned 2021 survey analyses. Changes were also made to the demographic questions to align with current best practices that include more inclusive response categories.

The first two questions of the survey instrument (see appendix A.2) confirmed the respondents' eligibility by asking if they participated as an applicant or reviewer in the merit review process since October 1, 2018. The survey asked the respondents about their experiences with the merit review process within their reported roles. Applicants were asked about the number of proposals submitted, perceptions of reviews of declined proposals (if applicable), experience submitting proposals, and perceived quality of reviews they received. Reviewers were asked about the number of reviews completed, reasons for declining to participate (if applicable), and perceptions of the quality of the proposals they reviewed. All respondents, regardless of role, were also asked to report their overall satisfaction with the merit review process and their intention to continue to submit or review proposals in the future.

To mitigate a trend toward declining response, the team embedded an experiment within the data collection period to determine if shortening the number of words in the initial invitation and first reminder emails would yield more responses than longer versions used in prior survey years. See appendix A.2 for the results of the experiment and its effects on the study design.

3.1.3.1 2021 survey response

Table 3.1 shows the number of respondents to the survey.

Table 3.1. 2021 Merit Review Survey data analyses inclusions and exclusions

	Count
Number of individuals in the universe (N)	66,585
Number of respondents*	17,330
Number of individuals excluded from response rate calculation**	3,073
Nonrespondents	46,182
Response rate	27%

Note: * Respondents include those who completed or partially completed a survey. A survey was considered partially completed if the questions used to construct the satisfaction indices were answered, even if other questions were unanswered.

** A total of 3,073 individuals were excluded from the response rate calculation for the following reasons: (1) the study team did not have valid email addresses for these individuals prior to survey launch ($N = 18$); (2) the survey email invitation was returned as undeliverable ($N = 1,683$); or (3) the respondent indicated they had not submitted or reviewed a proposal since October 1, 2018 ($N = 1,372$).

Over the course of 12 weeks, 17,330 individuals completed the survey for an overall response rate of 27 percent.⁴ Of these completions, 5,554 (32.1 percent) of respondents identified themselves as applicants, 3,494 (20.2 percent) identified as reviewers, and 8,282 (47.8 percent) participated in the merit review process in both roles. Partial completes (90 respondents) were individuals who answered all survey items deemed critical but did not make it through the entire web survey. (See appendix B for additional information about nonresponse bias analysis and weighting activities.)

3.2 Data Management and Analytic Approach

The study team analyzed responses to the 2021 Merit Review Survey to assess applicant and reviewer experiences with and perceptions of the merit review process in the context of the four research dimensions—satisfaction, burden, quality, and fairness. This section provides an overview of the study team’s three-step approach to analyzing the survey and administrative data, including data cleaning, nonresponse analysis and weighting, and multiple regression analyses. Appendix B provides supplemental information about the study team’s approach.

3.2.1 Data cleaning and variable construction

After closing the 2021 Merit Review Survey, the team cleaned the survey datafile and prepared data for analyses (see appendix B.1 for more details). Data processing procedures to ensure high-quality data included checking ranges to verify that each field contained only allowable codes, checking skip patterns, and checking numeric sums to verify that numeric responses made up the sum of component data items. A final survey disposition status was assigned to each case to identify complete, partially complete,⁵ and ineligible surveys. The team used this information to calculate the response rate (see table 3.1).⁶

Next, the team used categorical survey data on applicant and reviewer characteristics to construct independent variables (predictors in the regression analyses) for 2021 and 2019 (see appendix B.1 for more details regarding variable construction, including decisions to collapse classes and definitions of specific classes). Table 3.2 lists the 2021 independent variables and their respective classes. The 2019 independent variables mirrored those in 2021 with two exceptions: 2019 institutional affiliation had only one academic category while 2021 had multiple determinations (e.g., R1⁷/minority-serving institution [MSI], non-R1/MSI), and disability status was not measured in 2019.

Lastly, the team used administrative data from NSF (see appendix tables A.1.1. and A.1.2) to impute values when data were missing for independent variables. Only a small number of observations were imputed: 45 observations were imputed for gender, 56 for institutional affiliation, 69 for disability status, and 130 for early career status.

Table 3.2. 2021 Merit Review Survey independent variables

Independent variable	Classes
NSF experience	■ Experience as applicant only ■ Experience as reviewer only ■ Experience as both applicant and reviewer

⁴ Calculated using the American Association for Public Opinion Research (2016) response rate 2 formula, which removes cases where the respondent indicated they had not submitted or reviewed a proposal within the survey timeframe.

⁵ Respondent had completed at least the survey items that were used to construct the satisfaction indices.

⁶ Respondent reported they had not submitted a proposal or served as a reviewer since October 1, 2018.

⁷ R1 represents doctoral universities with very high research activities, as defined by the Carnegie Classification of Institutions of Higher Education (N.d.).

Independent variable	Classes
Race and ethnicity ^a	■ Asian, non-Hispanic ■ Black or African American, non-Hispanic ■ Hispanic, any race ■ Multiple races, non-Hispanic ■ White, non-Hispanic ■ Other^b • Native Hawaiian or Other Pacific Islander, non-Hispanic (NHPI) • American Indian or Alaska Native, non-Hispanic (AIAN) • Other race, non-Hispanic (other) ■ Unknown or do not wish to provide
Early career status	■ Early career (less than 10 years since highest degree earned) ■ Not early career
Institutional affiliation	■ Institution of higher education: R1/MSI ■ Institution of higher education: R1/non-MSI ■ Institution of higher education: non-R1/MSI ■ Institution of higher education: non-R1/non-MSI ■ For-profit organization ■ Nonprofit organization ■ Government (local, state, federal, or tribal) ■ Other • Institution of higher education: R1/unknown MSI • Institution of higher education: non-R1/unknown MSI • Federally funded research and development center • Primary or secondary academic institution • Other, non-higher education organization
Disability status	■ Disabled (hearing, mobility, vision, or other disability) ■ Not disabled ■ Do not wish to provide
NSF directorate/office	■ Biological Sciences (BIO) ■ Computer and Information Science and Engineering (CISE) ■ Education and Human Resources (EHR) ■ Engineering (ENG) ■ Geosciences (GEO) ■ Mathematical and Physical Sciences (MPS) ■ Social, Behavioral, and Economic Sciences (SBE) ■ Office of the Director (OD) • Office of Integrative Activities (OIA) • International Science and Engineering (OISE) ■ Other (specify)—open-ended response in Q2 if above response options did not apply
Gender	■ Female ■ Male ■ Other, do not wish to provide, or unknown
Merit Review Survey year ^c	■ 2019 ■ 2021

Note: MSI = minority-serving institution; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.)

^a Variable was constructed using separate race and ethnicity survey questions.

^b Because of small cell sizes for specific race and ethnicity categories, the following race and ethnicity groups were collapsed into “other” for analytic purposes: Native Hawaiian/Pacific Islander (non-Hispanic), Native American/Alaska Native (non-Hispanic), and other race. See section 3.3.3.2 for additional details.

^c This independent variable is only included in the year-over-year analysis models.

The team also used responses to survey items to construct dependent variables (see table 3.3 for a description of all dependent variables organized by research dimension). Most of the dependent variables were constructed as three, four, or five-level Likert scale items, which assessed the degree to which a respondent agreed or disagreed with a particular sentiment about the merit review process (i.e., 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree). A small portion of the dependent variables were continuous (e.g., the number of proposals or reviews submitted to NSF since October 1, 2018, or the number of hours required to write and submit the most recent proposal or review). Others were constructed as nominal (or unordered) variables. The team also used several survey items to construct satisfaction indices, which were treated as continuous dependent variables in their respective regression models. Separate indices were created for applicants and reviewers. See appendix B.1 for more details regarding construction of the satisfaction indices. See appendix table C.1.2 for a detailed list of each dependent variable for applicants and reviewers.

Table 3.3. Dependent variables by research dimension

Research questions by dimension
Assessment of satisfaction with the merit review process
How satisfied are applicants and reviewers with the merit review process? To what extent has satisfaction with the merit review process changed since 2019?
■ Overall satisfaction with the merit review process ■ Applicant satisfaction scores^a ■ Reviewer satisfaction scores^b
Assessment of burden in the merit review process
What is the burden associated with the merit review process for applicants and reviewers? To what extent has the burden associated with the merit review process changed since 2019?
■ Applicants: Number of proposals submitted to NSF ■ Applicants: Time spent preparing the most recent proposal ■ Applicants: Amount of effort to write and complete a proposal compared with other agencies ■ Reviewers: Number of reviews conducted for NSF ■ Reviewers: Amount of time to conduct a review ■ Reviewers: Time of day reviewers typically worked on their reviews ■ Reviewers: Employer's view of participating as a reviewer ■ Reviewers: Decline to serve as a reviewer ■ Reviewers: Extent to which a "lack of time" influenced the decision to decline ■ Reviewers: Extent to which too many NSF review requests influenced the decision to decline ■ Reviewers: Extent to which competing professional pressures influenced the decision to decline ■ Reviewers: Extent to which commitments as a reviewer to other agencies influenced the decision to decline ■ Reviewers: Extent to which an inability to travel to face-to-face panels influenced the decision to decline
Assessment of quality in the merit review process
What are applicants' perceptions of review quality and reviewers' perceptions of proposal quality? To what extent have perceptions of proposal and review quality changed since 2019?
■ Applicants: Extent to which reviews improved understanding of the proposal process ■ Applicants: Extent to which reviews provided useful information for improving future proposals ■ Applicants: Extent to which reviews influenced the decision to submit to another funding agency ■ Applicants: Extent to which reviews discouraged revising and resubmitting proposal to NSF ■ Applicants: Satisfaction with the quality of the information provided by NSF ■ Applicants: Extent to which written reviews are thorough ■ Applicants: Extent to which written reviews are technically sound ■ Applicants: Extent to which written reviews are of high overall quality

Research questions by dimension
■ Applicants: Extent to which panel summaries are of high quality ■ Applicants: Extent to which information provided by NSF regarding the outcome of the competition is of high quality ■ Reviewers: Proposals reviewed have been of high quality ■ Applicants and reviewers: Factors for improvement
Assessment of fairness in the merit review process
What are applicants' and reviewers' perceptions of fairness in the merit review process? To what extent have perceptions of fairness of the merit review process changed since 2019?
■ Applicants: Perceptions that individuals submitting proposals are treated fairly ■ Reviewers: Perceptions that individuals submitting proposals are treated fairly ■ Applicants and reviewers: NSF's merit review process is fair overall

Note:

^a Applicant satisfaction indices included 2019 survey items Q26B, Q28A, Q28B, Q28C, Q28D, Q29A, Q29B, and Q29E and 2021 survey items Q18B, Q20A, Q20B, Q20C, Q20D, Q20E, Q21A, Q21B, and Q21E.

^b Reviewer satisfaction indices included 2019 survey items Q29A, Q29B, and Q29E and 2021 survey items Q21A, Q21B, and Q21E.

Finally, the team reviewed open-ended responses from the comment question at the end of the survey and selected some that were relevant to highlight throughout the report. These quotes are intended to provide additional context but are not representative of the statistical findings reported.

3.2.2 Nonresponse bias and weighting adjustments

To assess the potential for nonresponse bias, the team examined the relationships between key dependent variables (such as perception of fairness and satisfaction with NSF's merit review process) and (1) the probability of responding to the survey; (2) population characteristics; and (3) the stage (early, nonrespondent follow-up, or late) at which a respondent completed the survey. The team also examined the relationships between the probability of responding to the survey and population characteristics. See appendix B.2 for additional details.

To mitigate the potential for nonresponse bias in the estimates, the team assigned a weight to each respondent to ensure respondents were representative of the population on demographic characteristics (directorship, gender, and race and ethnicity) that were related to the probability of response and key dependent variables. See appendix B.3 for additional details.

3.2.3 Multiple regression: examination of group differences

The team employed regression modeling to detect subgroup differences among applicants and reviewers on the key dependent variables (see table 3.3). Models were fit using the seven independent variables, including NSF experience, race and ethnicity, early career status, institutional affiliation, disability status, NSF directorship/office, and gender (see table 3.2), with outcomes under each of the four research dimensions as the dependent variables. The analysis employed linear regression for all continuous dependent variables, logistic regression for all binary dependent variables, ordinal logistic regression for all ordinal dependent variables, and multinomial logistic regression for all nominal dependent variables. See appendix B.4 for more details, including regression model equations.

Following a postestimation⁸ of adjusted probabilities (logistic models) or means (linear models), significance tests determined whether the average marginal effects for each independent

⁸ Postestimation used STATA's margins commands (STATA.com, N.d.)

variable were statistically different from zero (e.g., whether there was a nonzero average difference in mean satisfaction scores between female and male applicants). The calculated adjusted probabilities/means were used to illustrate findings from the regression models in figures in chapter 4. The average marginal effects were used to describe the observed class differences (e.g., differences in the adjusted probabilities/means) for highly significant (regression coefficients with $p < 0.001$) independent variables. For example, according to the 2021 survey results, female reviewers were 11 percentage points more likely than male reviewers to decline an NSF proposal review request. In this example, 11 percentage points is the difference between the adjusted probabilities of declining for females and males (e.g., $0.45 - 0.34 = 0.11$ or 11 percentage points).

Four dependent variables were log transformed to correct for skewness in the distribution. These variables included number of proposals submitted, hours spent developing and submitting a recent proposal, number of proposal reviews, and hours spent conducting a recent review. Statistically significant coefficients ($p < 0.001$) were exponentiated to aid interpretation. The exponentiated value was then subtracted from 1, which resulted in a final value to be interpreted as the percent change (positive or negative) in the dependent variable for a significant class within an independent variable compared with the reference group.

The analysis also examined change in applicant and reviewer experiences since 2019 and whether change was conditioned by the following four population characteristics: race and ethnicity, gender, institution affiliation, and early career status. For these tests, models were fit using six of the seven key population characteristics⁹ as independent variables, an analysis year indicator, and four interaction terms (e.g., the product of the year indicator and each of the four key population characteristics). See appendix B.4 for more details.

3.3 Limitations

This section describes elements of study design and operationalization that should be considered when interpreting the findings.

3.3.1 Response rate and generalizability

The response rate for the survey (27 percent) may limit the generalizability of the results. To the extent that respondents and nonrespondents differ from each other, nonresponse bias may limit the inferences that can be drawn from the study. The study team conducted nonresponse bias analyses to assess the potential impact of nonresponse on survey inferences and developed a nonresponse weighting method to mitigate the nonresponse bias. The analysis results indicated little evidence of bias on demographic characteristics, similar to what was observed in the 2019 data (see appendix B.2). However, to the extent that nonresponse weighting does not account for differences between respondents and nonrespondents, the inferences that can be drawn from the study are limited.

3.3.2 Small cell sizes

Small cell sizes required the team to collapse classes within certain independent variables to allow the regression models to successfully run and ensure stability of the estimated model

⁹ Disability status was not available in the 2019 data and, therefore, was not used in the year-to-year analyses.

coefficients. These adjustments were made uniformly across all models to ensure consistency in analyses.

- **NSF directorate/office:** collapsed Office of Integrative Activities (OIA) and Office of International Science and Engineering (OISE) into Office of the Director (OD)
- **Institution type:** collapsed institutions of higher education of unknown MSI status (both R1 and non-R1 institutions), federally funded research and development centers, and primary or secondary academic institutions into “other”
- **Gender:** collapsed other and do not wish to provide into “unknown”
- **Race and ethnicity:** combined self-reported Native Hawaiian/Pacific Islander, Native American/Alaska Native, and other into “other, NHPI, AIAN”

Small cell sizes also affected analysis of the fairness dimension of the extent to which applicants agreed or disagreed that individuals submitting proposals were treated fairly. Small cell sizes for one institution type class (R1/MSI unknown) created predicted probabilities of zero, which interfered with the calculations. The decision was made to drop institution affiliation type as an independent variable in the model to enable the model to run successfully with the other five independent variables. This only affected the year-to-year analyses; the 2021 analyses ran with all independent variables.

3.3.3 Conclusions about differences by race and ethnicity

The findings presented for race and ethnicity are limited in the conclusions that can be drawn for Asian respondents and those collapsed into the “other, NH/PI, NA/AN” race and ethnicity category, which included Native Hawaiian/Pacific Islander, Native American/Alaska Native, and Hispanic.

3.3.3.1 Findings for Asian respondents

The 2021 Merit Review Survey adopted the approach other federal surveys use to measure race and ethnicity. The survey asked respondents to select from a list of options, including a general “Asian” racial category. Because this category does not take into account the vast intragroup heterogeneity in socioeconomic statuses and racial experiences present across groups of people originating in the Far East, Southeast Asia, or the Indian subcontinent (Office of Management and Budget 1997; Kim 2007; Min 2002; Dhingra 2007), conclusions drawn for the “Asian” category should be interpreted with this important limitation in mind.

3.3.3.2 Findings for Native Hawaiian or other Pacific Islander, American Indian or Alaska Native, and other race and ethnicity groups

As a result of the regression analytic approach this research employed, in combination with low overall response rates for Native Hawaiian/Pacific Islander, Native American/Alaska Native, and other race and ethnicity categories, respondents self-identifying in these groups were collapsed into a single “other, NHPI, AIAN” category. See appendix table C.1.1 for a breakdown of survey responses for the groups included in this category.

Shotton, Lowe, and Waterman (2013) discuss longstanding issues related to how the academic community typically addresses small sample sizes of Native Americans in quantitative research. The authors point out that researchers often do not report or discuss these data or note them as not statistically significant. These practices have similar implications as collapsing different known race and ethnic groups into a single category because they reinforce the invisibility of these populations (Shotton et al. 2013).

While the decision to collapse Native Hawaiian/Pacific Islander, Native American/Alaska Native, and other race and ethnic groups into a single “other, NHPI, AIAN” race and ethnicity category ensures the analysis includes the data from these groups, doing so fails to account for the diversity existing across these groups. Thus, the regression findings for this category cannot be meaningfully interpreted.

3.3.4 Limits of the regression modeling approach

A limitation of the regression model structure is that it does not allow for pairwise comparisons between all classes of an independent variable. Rather, one must choose a reference class for each independent variable included in the analysis. For the majority of independent variables, the team consistently chose the largest class (i.e., male respondents, White respondents, non-early career respondents, those without disabilities, those affiliated with R1/non-MSI institutions, those submitting to or reviewing for the Directorate for Mathematical and Physical Sciences [MPS]). The team deviated from this rule in the analysis of NSF experience. When modeling applicants or reviewers separately, the team set either applicant-only or reviewer-only as the reference class. This approach was motivated by the direction of the research questions, which were primarily concerned with the impact of having experience as both an applicant and reviewer. When modeling a pooled sample of applicants and reviewers, the team defaulted to the established rule of setting the largest class as the reference, which facilitated more intuitive comparisons. The team then compared all other classes within a given independent variable to the reference class. Although this approach is relatively intuitive for variables with few classes (e.g., gender), interpretation can be more challenging for variables with larger numbers of classes (e.g., NSF directorate/office or institution affiliation).

3.3.5 Comparison of 2021 and 2019 data

The regression models comparing 2019 and 2021 data were adjusted in three ways, as outlined below, because of variation in the data available across years. It should also be noted that the COVID-19 health pandemic and deep social and political unrest across the United States occurred between the two survey periods. Additional research is needed to understand the possible effects of these broader events on participation in the merit review process.

3.3.5.1 Changes to survey questions and administrative data

Changes to survey questions and administrative data used in 2019 and 2021 changed the data available to construct institution affiliation type. To adjust for this change, the study team constructed a seven-value variable for both 2019 and 2021 for the year-to-year analyses to ensure a direct mapping could be made. The most significant change to this new variable was the collapsing of all higher education categories into a single “higher education institution” value, which was the reference group for the year-to-year analyses. As a result, mean comparisons cannot be made at the more granular R1 and MSI institution status for year-to-year analyses. The seven-value variables were (1) higher education institution, (2) primary or secondary academic institution, (3) nonprofit organization, (4) for-profit or industry organization, (5) federally funded research and development center, (6) government, and (7) other. See appendix table C.1.1 for a description of the complete 12-category variable used in the standalone analysis of 2021 survey data.

3.3.5.2 Changes in response scale or survey reference data

As a result of the changes in response scale or reference dates for eight survey items between 2019 and 2021, these dependent variables were dropped from the year-to-year analysis. Single-year analyses of these data from the 2021 survey were still conducted.

The response scale for two survey items changed in 2021 from a 4-point to a 5-point Likert scale with the addition of a middle neutral value, resulting in strongly agree, agree, neither agree nor disagree, disagree, and strongly disagree. These two items follow:

- Overall, the majority of proposals I have reviewed in recent years have been of high quality.
- Individuals submitting proposals are treated fairly.

The response scale for six survey items changed from a 5-point to a 6-point Likert scale between 2019 and 2021 with the addition of a middle neutral value, resulting in strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, and not applicable. These six items follow:

- Written reviews are thorough.
- Written reviews are technically sound.
- Written reviews are of high quality.
- The panel summary or summaries are of high quality.
- The information provided regarding the outcomes of the competition is of high quality.
- Individuals submitting proposals are treated fairly.

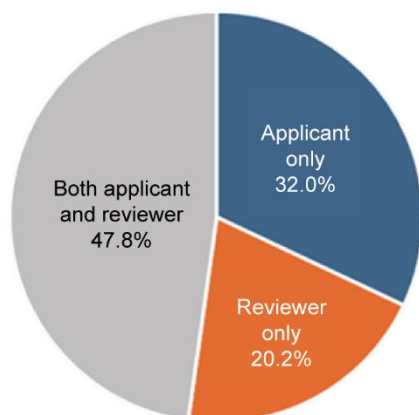
The 2019 Merit Review Survey asked about proposals submitted and reviews conducted between October 2015 (the start of FY 2016) and August 2019 (the time of the survey)—a 47-month span. The 2021 Merit Review Survey asked about October 2018 to September 2021 timeframe—a 36-month span. As a result, respondent reports of the number of proposals submitted to NSF and reviews conducted for NSF are not easily comparable because the reference period lengths differ. Therefore, these two dependent variables were not included in the year-to-year analyses and only reported for 2021.

4 Assessing the 2021 Merit Review Process

4.1 Descriptive Statistics

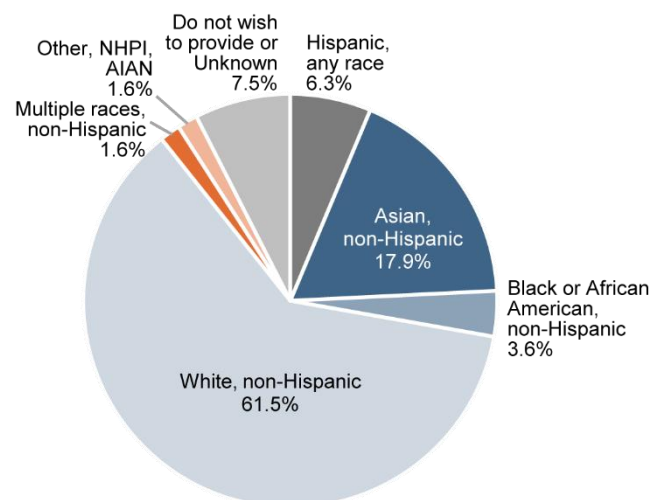
As figures 4.1.1–4.1.7 show, about half of survey respondents had both submitted and reviewed NSF proposals between October 1, 2018, and the time of the survey launch September 9, 2021. The majority identified as White, non-Hispanic, male, nondisabled, and non-early career scientists. The largest proportion of respondents reported their scholarly activities were most closely affiliated with MPS and the fewest with OD.

Figure 4.1.1. Respondent role in the merit review process



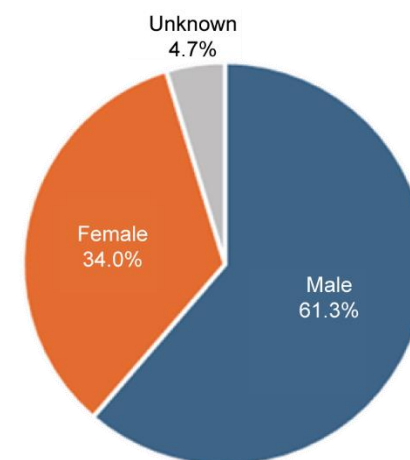
Note: $N = 17,330$
See appendix table C.1.1 for detailed results.
Source: 2021 Merit Review Survey

Figure 4.1.2. Respondent race and ethnicity



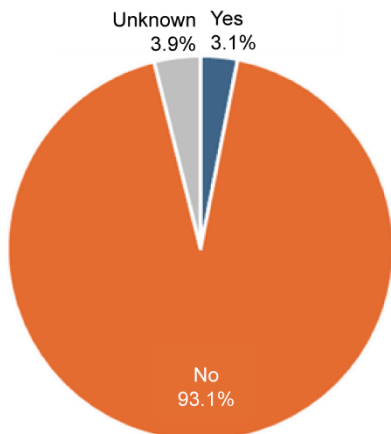
Note: $N = 17,330$
Because of small cell sizes that rendered regression models inestimable, the study team collapsed respondents who identified as American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, or other race and ethnicity into one “other” category. See appendix table C.1.1 for more details on this category and the subgroups that compose it.
Source: 2021 Merit Review Survey

Figure 4.1.3. Respondent gender



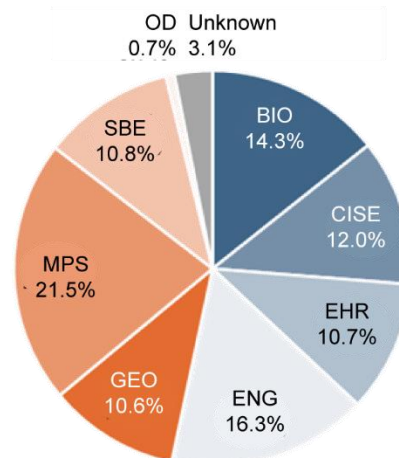
Note: $N = 17,330$
See appendix table C.1.1 for detailed results.
Source: 2021 Merit Review Survey

Figure 4.1.4. Respondent disability status



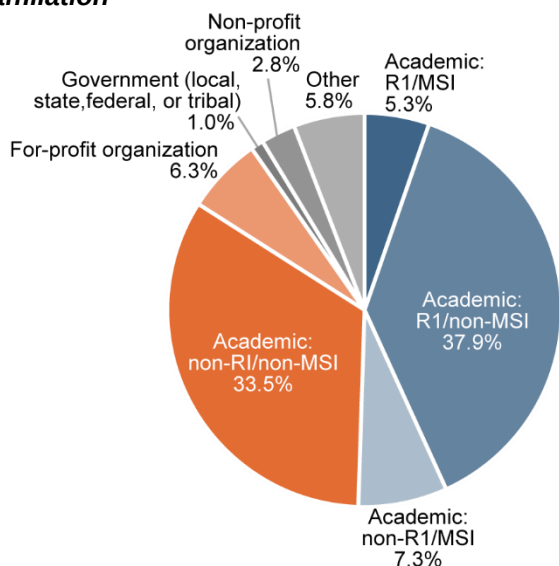
Note: *N* = 17,330
See appendix table C.1.1 for detailed results.
Source: 2021 Merit Review Survey

Figure 4.1.5. Respondent NSF directorate or office affiliation



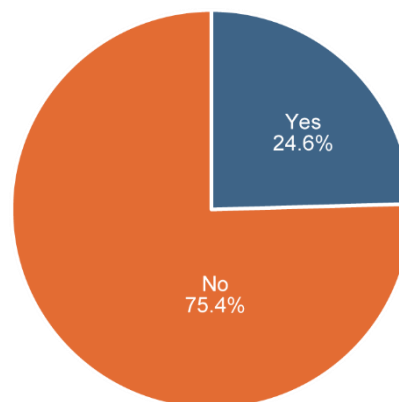
Note: *N* = 17,330
See appendix table C.1.1 for detailed results.
BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; OD = Office of the Director; SBE = Social, Behavioral, and Economic Sciences
Source: 2021 Merit Review Survey

Figure 4.1.6. Respondent institutional affiliation



Note: *N* = 17,330
See appendix table C.1.1 for detailed results.
MSI = minority-serving institution; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); "other" includes primary and secondary institution, federally funded research and development center, academic R1/unknown, and academic non-R1/unknown
Source: 2021 Merit Review Survey

Figure 4.1.7. Respondent early career status



Note: *N* = 17,330
See appendix table C.1.1 for detailed results.
Source: 2021 Merit Review Survey

4.2 Satisfaction With the Merit Review Process

Satisfaction Highlights

- Over half of respondents were satisfied with the merit review process.
- Respondent satisfaction was lower among respondents in 2021 than in 2019.
- Respondents with reviewer experience (regardless of whether they had also been applicants) and those from non-R1/MSIs were generally more satisfied than their peers without reviewer experience and those from R1/non-MSIs.
- By contrast, respondents with only applicant experience and those with disabilities were less satisfied than their counterparts without applicant experience and without disabilities.
- Asian, non-Hispanic respondents were more satisfied with the merit review process than their White, non-Hispanic peers. In contrast, respondents in the “other” race and ethnicity category, which includes American Indian and Native Hawaiian respondents, were less satisfied.

This study examined satisfaction with the merit review process in two ways. First, the team analyzed overall satisfaction across the entire respondent pool using a single survey item. The survey item asked all respondents to which extent they agreed or disagreed with the following statement: “I am satisfied with the NSF merit review process.”

Second, the team constructed and analyzed two satisfaction indices: one measuring applicant satisfaction and one measuring reviewer satisfaction. Unlike the single survey item representing overall satisfaction among all respondents, the two satisfaction indices provided a more robust and multidimensional indicator of satisfaction. Both satisfaction index scores ranged from 0 to 100. See appendix B.1 for additional details on the construction of each index.

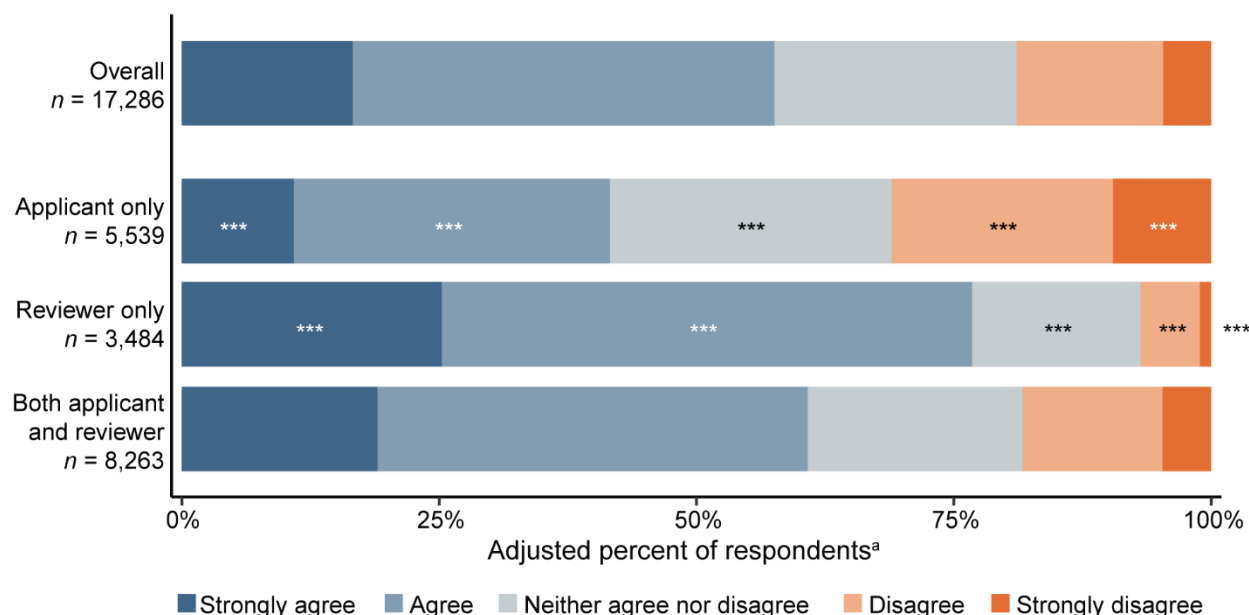
All analyses were modelled controlling for key respondent characteristics (i.e., role in the merit review process, race and ethnicity, gender, disability status, NSF directorate or office, institution affiliation, and early career status). To focus the discussion, the report highlights findings statistically significant at $p < 0.001$. Detailed regression tables present all average marginal effects (AMEs) and p -values in appendix C.

4.2.1 Overall satisfaction among all respondents

Overall, the majority (54 percent) of survey respondents agreed or strongly agreed they were satisfied with NSF’s merit review process (see appendix table C.1.3). However, this represents a lower level of satisfaction since the previous Merit Review Survey. In 2021, survey respondents were 14 percentage points less likely to agree and 7 percentage points less likely to strongly agree they were satisfied compared with survey respondents in 2019 (see appendix table C.2.4.B; Hare et al. 2020).

As figures 4.2.1–4.2.4 illustrate, overall satisfaction with the merit review process depended on whether a respondent had submitted a proposal for funding, institutional affiliation, disability status, and race and ethnicity.

Figure 4.2.1. Applicant and reviewer agreement with being satisfied with the NSF review process, by role in the merit review process



Note: N = 17,286

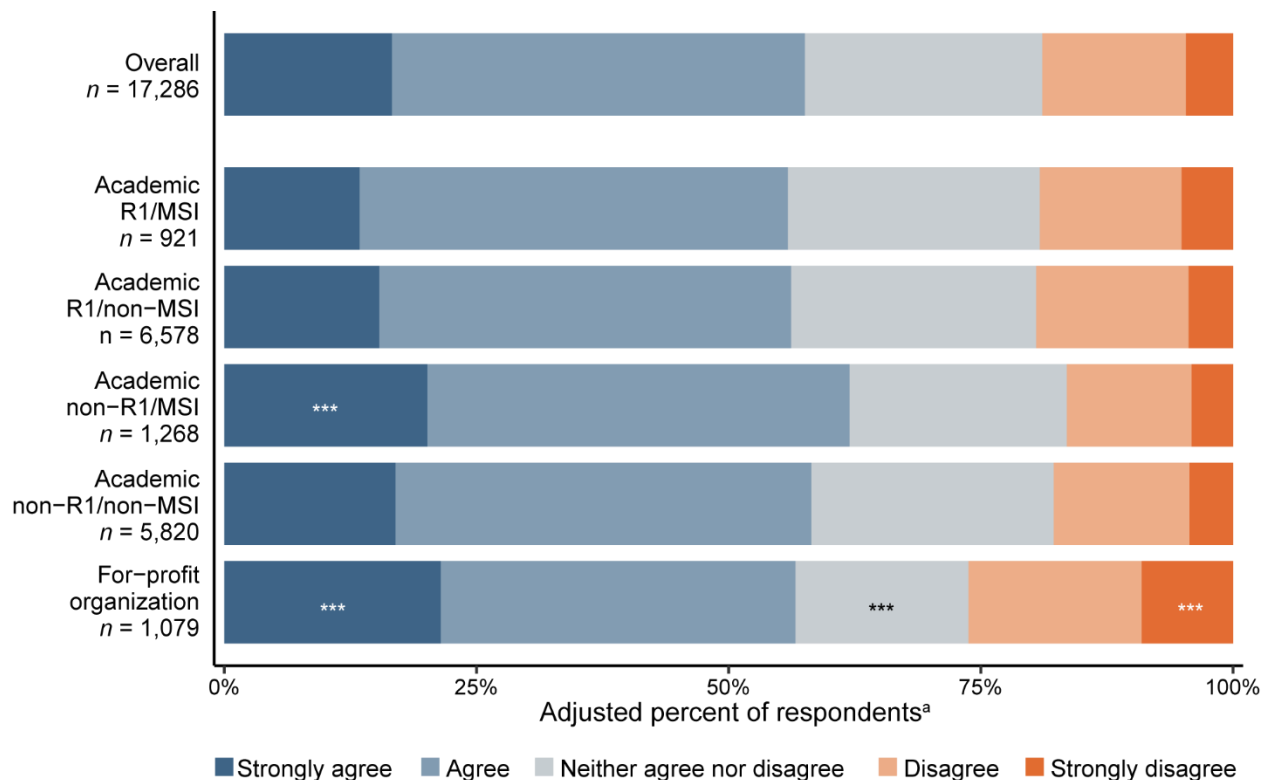
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.1 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (both applicant and reviewer).

Source: 2021 Merit Review Survey

Figure 4.2.1 shows that respondents who had only submitted proposals (applicants) were less likely to strongly agree they were satisfied with the process than those who had also reviewed proposals (an 8 percentage point difference). Conversely, those who had only reviewed proposals were more likely to strongly agree (a 6 percentage point difference) that they were satisfied with the proposal process.

Figure 4.2.2. Applicant and reviewer agreement with being satisfied with the NSF review process, by institutional affiliation



Note: N = 17,286

MSI = minority-serving institution; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.)

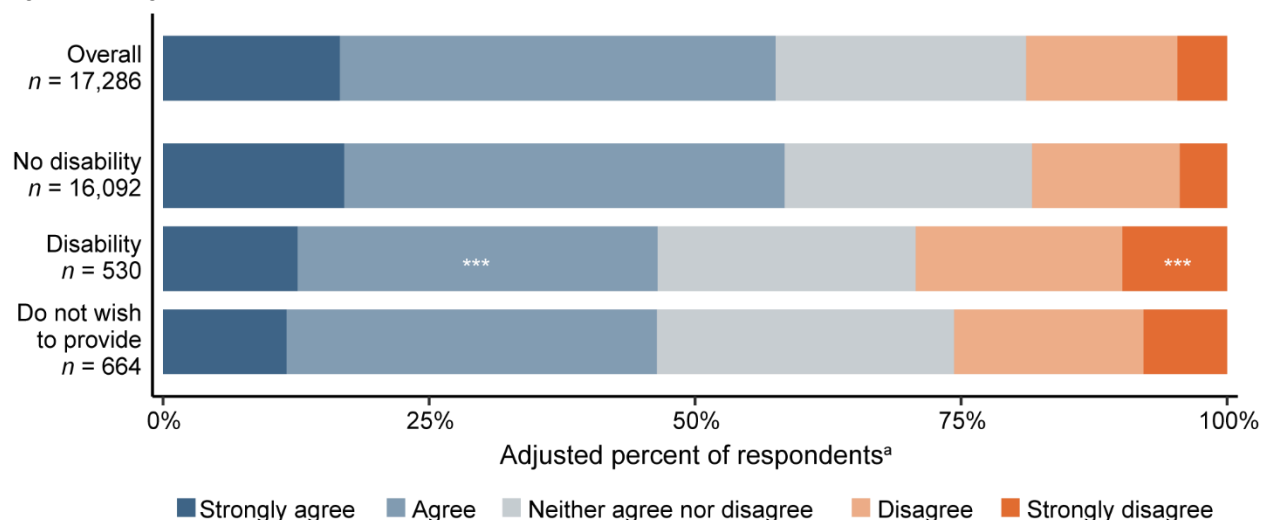
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.1 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (R1/non-MSI).

Source: 2021 Merit Review Survey

As figure 4.2.2 illustrates, respondents affiliated with non-R1/MSIs were more likely to strongly agree they were satisfied with the merit review process than those affiliated with R1/non-MSIs—the most common institution type (a 5 percentage point difference). Respondents affiliated with for-profit organizations tended to have stronger responses in either direction, meaning that they were more likely to strongly agree or strongly disagree they were satisfied than those affiliated with R1/non-MSIs (6 and 5 percentage point differences, respectively).

Figure 4.2.3. Applicant and reviewer agreement with being satisfied with the NSF review process, by disability status



Note: N = 17,286

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.1 for detailed regression results.

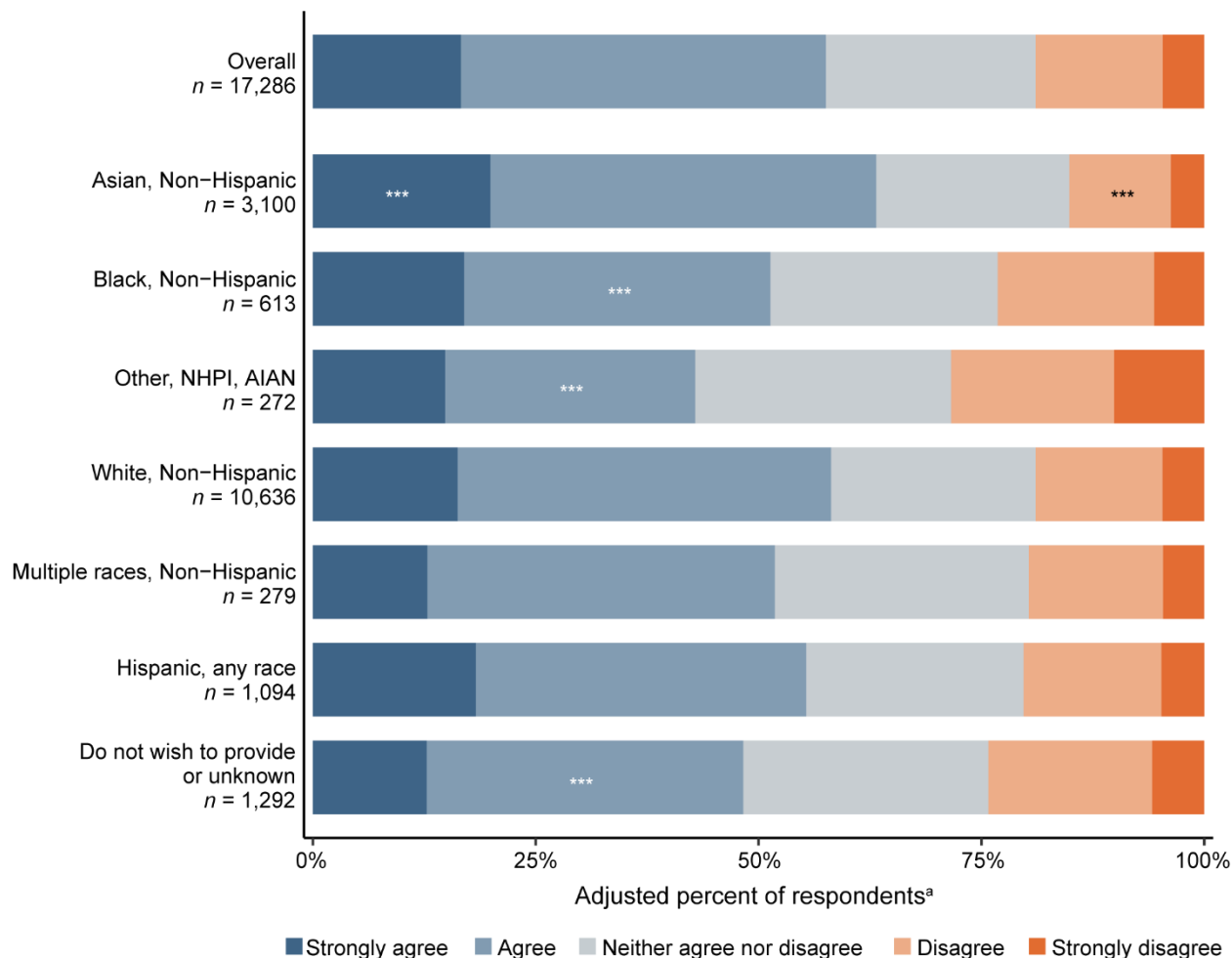
*** indicates significant at $p < 0.001$ relative to the reference group (no disability).

Source: 2021 Merit Review Survey

Figure 4.2.3 shows that respondents who reported having a vision, hearing, mobility, or other serious disability were less likely to agree they were satisfied with the process than those without a disability (an 8 percentage point difference). Although the survey did not directly ask respondents to expand on their reports related to satisfaction and disability, one open-ended comment offered some possible contributing factors:

"The recent changes in the formatting of supporting files are a huge nightmare for neurodivergent people. There is no assistance or special consideration for deadlines or requirements for neurodivergent individuals or persons with disabilities that I know of. This discourages those folks from applying and reduces their chances of success."

Figure 4.2.4. Applicant and reviewer agreement with being satisfied with the NSF review process, by race and ethnicity



Note: N = 17,286

AIAN = American Indian or Alaska Native; NHPI = Native Hawaiian or Other Pacific Islander

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.1 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (White, non-Hispanic).

Source: 2021 Merit Review Survey

Figure 4.2.4 shows that respondents who identified as Asian, non-Hispanic were more likely to strongly agree they were satisfied with the review process than White, non-Hispanic respondents (a 4 percentage point difference). Black, non-Hispanic respondents were less likely to agree they were satisfied than White, non-Hispanic respondents (an 8 percentage point difference). Respondents in the race and ethnicity categories of “do not wish to provide or unknown” and “other, NHPI, AIAN” were less likely to agree they were satisfied with the review process than White, non-Hispanic respondents (a 6 and 14 percentage point difference, respectively).

Applicant and reviewer overall satisfaction also varied by directorate/office and early career status. Responses ranged from 10 to 20 percent of respondents associated with OD and Education and Human Resources (EHR), respectively, reporting they strongly agree they were

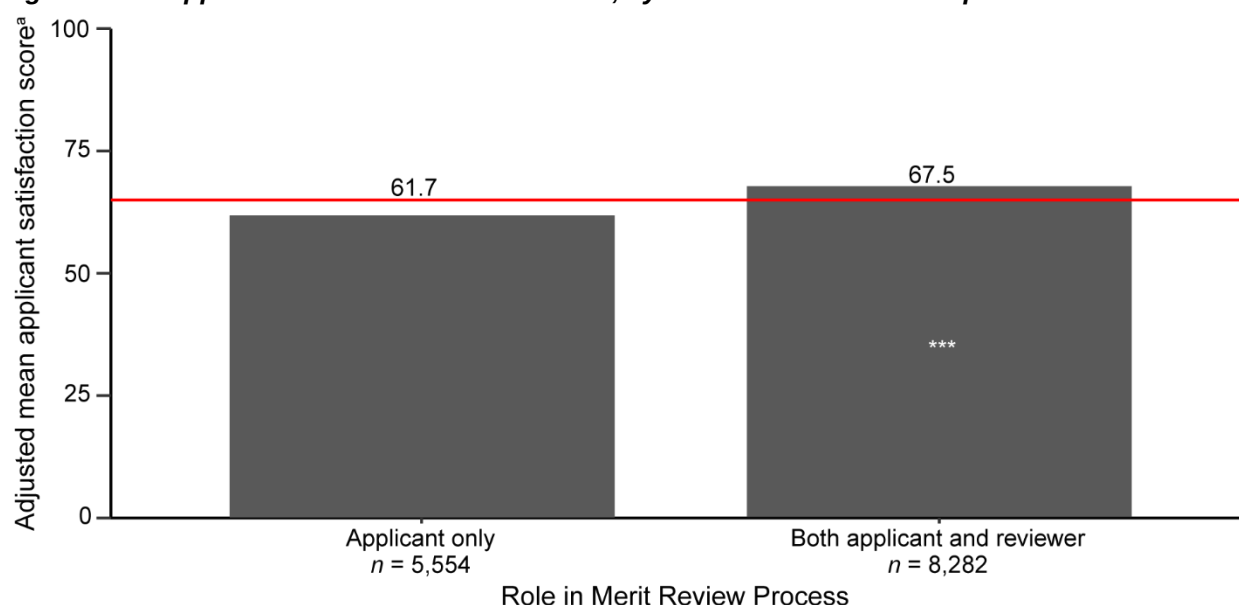
satisfied with the merit review process overall. Early career respondents were more likely to agree they were satisfied with the merit review process than respondents further along in their careers (a 4 percentage point difference). See appendix table C.2.1 for detailed regression results.

4.2.2 Applicant satisfaction index scores

Overall, applicants had an average satisfaction index score of 66, out of a possible range of 0 to 100 (see appendix table C.1.3), which is 5 points lower compared with respondents from the previous Merit Review Survey conducted in 2019 (see appendix table C.2.5; Hare et al. 2020).

As with overall satisfaction, applicant satisfaction scores varied by role, institutional affiliation, disability status, race and ethnicity, and directorate (see figures 4.2.5–4.2.9).

Figure 4.2.5. Applicant satisfaction index scores, by role in the merit review process



Note: $N = 13,836$

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.2 for detailed regression results.

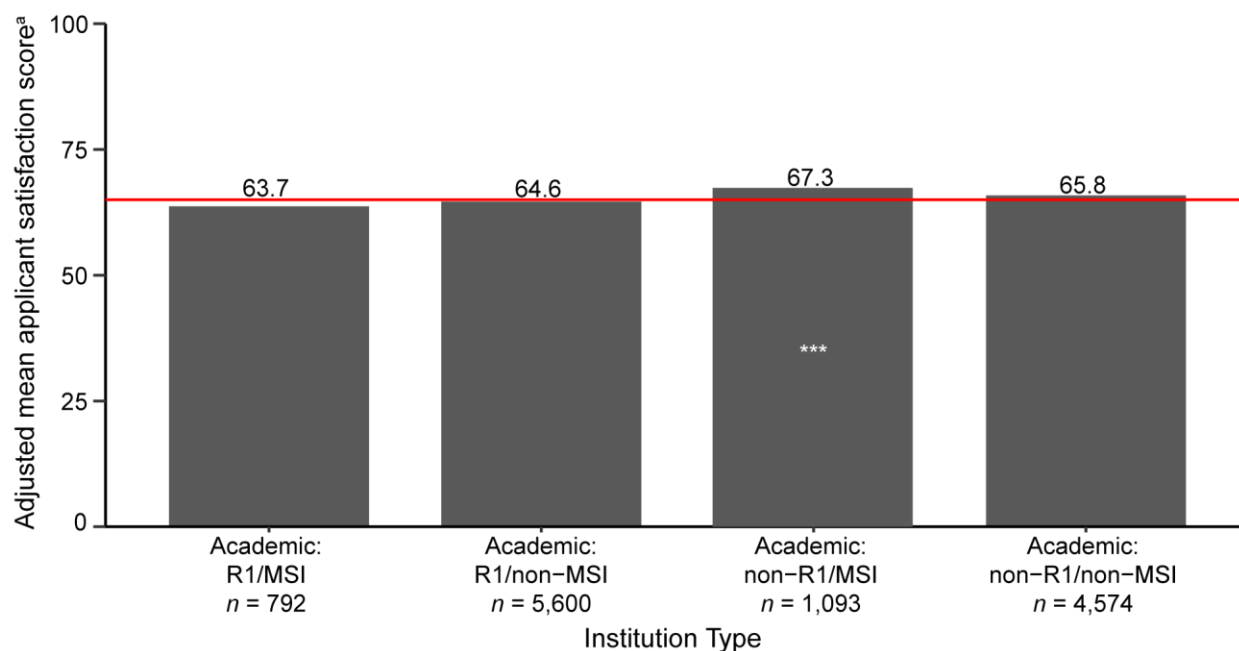
*** indicates significant at $p < 0.001$ relative to the reference group (applicant only).

— indicates the overall adjusted mean applicant satisfaction score.

Source: 2021 Merit Review Survey

As figure 4.2.5 illustrates, applicants who also reviewed proposals scored higher on the satisfaction index than those who had not (a 6-point difference).

Figure 4.2.6. Applicant satisfaction index scores, by institutional affiliation



Note: $N = 13,836$

Figure only shows adjusted means for academic institution categories ($N = 12,059$).

MSI = minority-serving institution; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.)

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.2 for detailed regression results.

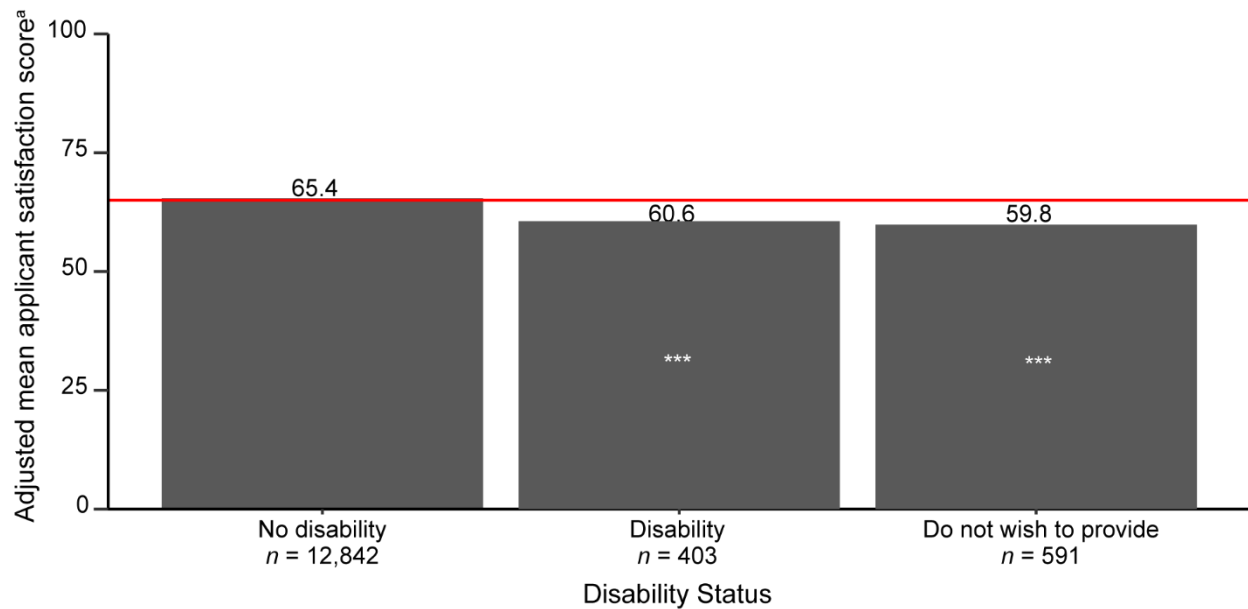
*** indicates significant at $p < 0.001$ relative to the reference group (R1/non-MSI).

— indicates the overall adjusted mean applicant satisfaction score.

Source: 2021 Merit Review Survey

As figure 4.2.6 shows, applicants associated with non-R1/MSIs scored higher on the satisfaction index than those from R1/non-MSI institutions (a 3-point difference).

Figure 4.2.7. Applicant satisfaction index scores, by disability status



Note: $N = 13,836$

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.2 for detailed regression results.

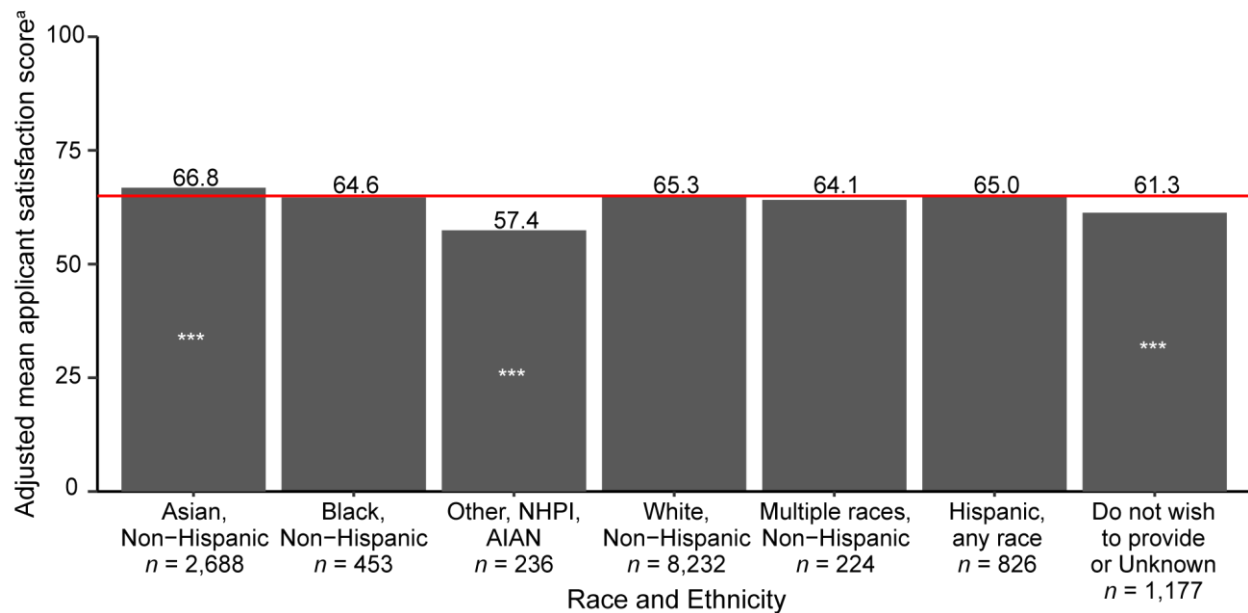
*** indicates significant at $p < 0.001$ relative to the reference group (no disability).

— indicates the overall adjusted mean applicant satisfaction score.

Source: 2021 Merit Review Survey

Figure 4.2.7 shows that applicants with a hearing, mobility, vision, or other disability had lower satisfaction than those without a disability (a 5-point difference).

Figure 4.2.8. Applicant satisfaction index scores, by race and ethnicity



Note: N = 13,836

AIAN = American Indian or Alaska Native; NHPI = Native Hawaiian or Other Pacific Islander

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.2 for detailed regression results.

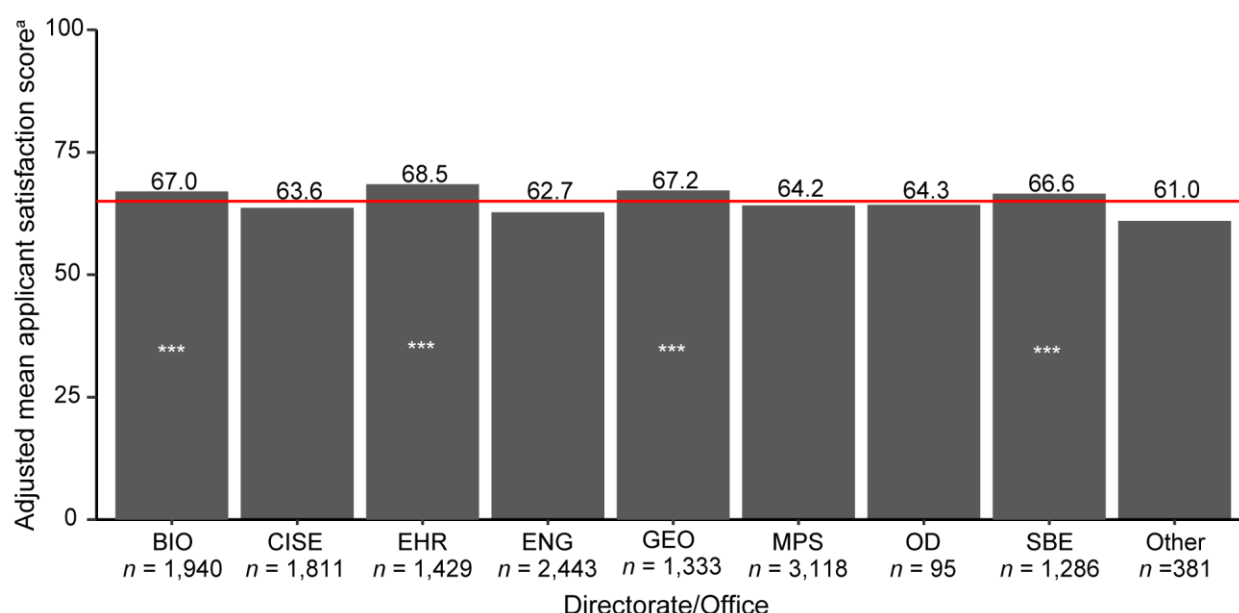
*** indicates significant at $p < 0.001$ relative to the reference group (White, non-Hispanic).

— indicates the overall adjusted mean applicant satisfaction score.

Source: 2021 Merit Review Survey

Figure 4.2.8 shows that applicants who identified as Asian, non-Hispanic were more satisfied than White, non-Hispanic applicants (a 2-point difference). Applicants in the other, NHPI, AIAN race and ethnicity category and those who did not wish to provide their race and ethnicity information were less satisfied than White, non-Hispanic applicants (an 8- and 4-point difference, respectively).

Figure 4.2.9. Applicant satisfaction index scores, by directorate



Note: $N = 13,836$

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; OD = Office of the Director; SBE = Social, Behavioral, and Economic Sciences; Other = directorate/office not listed in the survey

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.2 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (MPS).

— indicates the overall adjusted mean applicant satisfaction score.

Source: 2021 Merit Review Survey

Figure 4.2.9 illustrates that satisfaction scores varied across NSF directorates/offices, ranging from 61 for those associated with “other” to 69 for those associated with EHR.

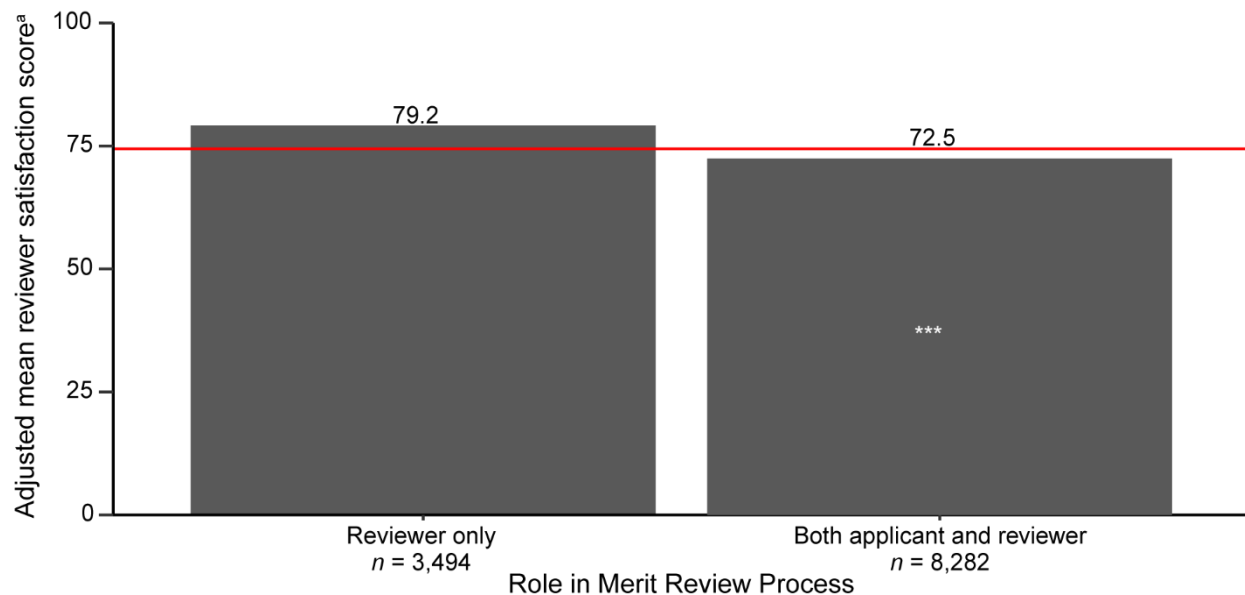
Applicant satisfaction also varied by gender and early career status. Female applicants were more satisfied than their male colleagues (a 2-point difference). Early career applicants were more satisfied than those further along in their careers (a 4-point difference). See appendix table C.2.2 for detailed regression results.

4.2.3 Reviewer satisfaction index scores

Overall, reviewers had an average satisfaction index score of 75 out of a possible range of 0 to 100 (see appendix table C.1.4). Similar to applicants, this score is 5 points lower than the previous Merit Review Survey conducted in 2019 (see appendix table C.2.6; Hare et al. 2020).

Reviewer satisfaction scores varied by role, disability status, race and ethnicity, and directorate (see figures 4.2.10–4.2.13).

Figure 4.2.10. Reviewer satisfaction index scores, by role in the merit review process



Note: $N = 11,776$

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.3 for detailed regression results.

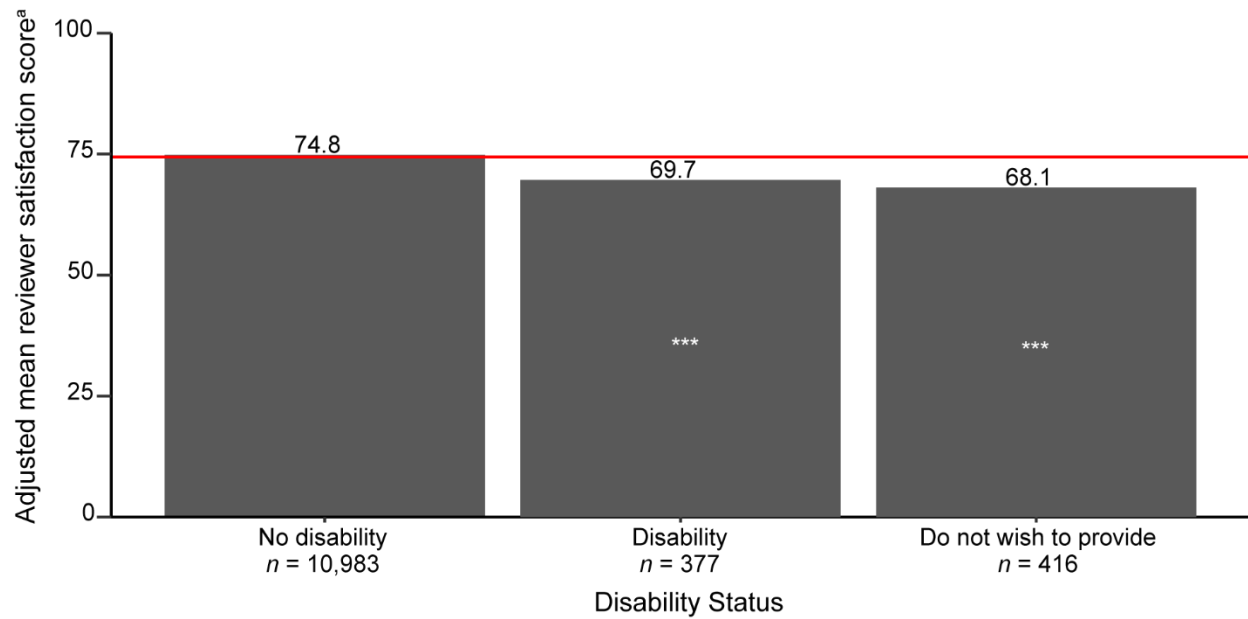
*** indicates significant at $p < 0.001$ relative to the reference group (reviewer only).

— indicates the overall adjusted mean reviewer satisfaction score.

Source: 2021 Merit Review Survey

As figure 4.2.10 shows, reviewers who also submitted NSF proposals scored lower on the satisfaction index than those who had only conducted reviews (a 7-point difference).

Figure 4.2.11. Reviewer satisfaction index scores, by disability status



N = 11,776

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.3 for detailed regression results.

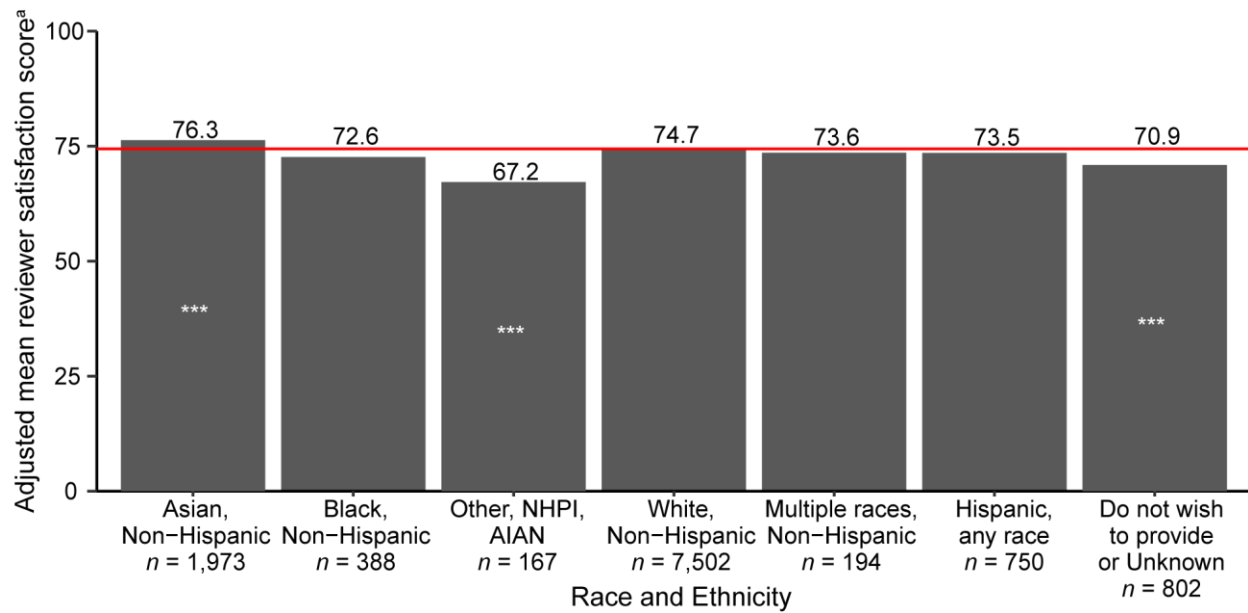
*** indicates significant at $p < 0.001$ relative to the reference group (no disability).

— indicates the overall adjusted mean reviewer satisfaction score.

Source: 2021 Merit Review Survey

As figure 4.2.11 illustrates, reviewers with hearing, mobility, vision, or another disability had lower satisfaction than those without a disability (a 5-point difference).

Figure 4.2.12. Reviewer satisfaction index scores, by race and ethnicity



N = 11,776

AIAN = American Indian or Alaska Native; NHPI = Native Hawaiian or Other Pacific Islander

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.3 for detailed regression results.

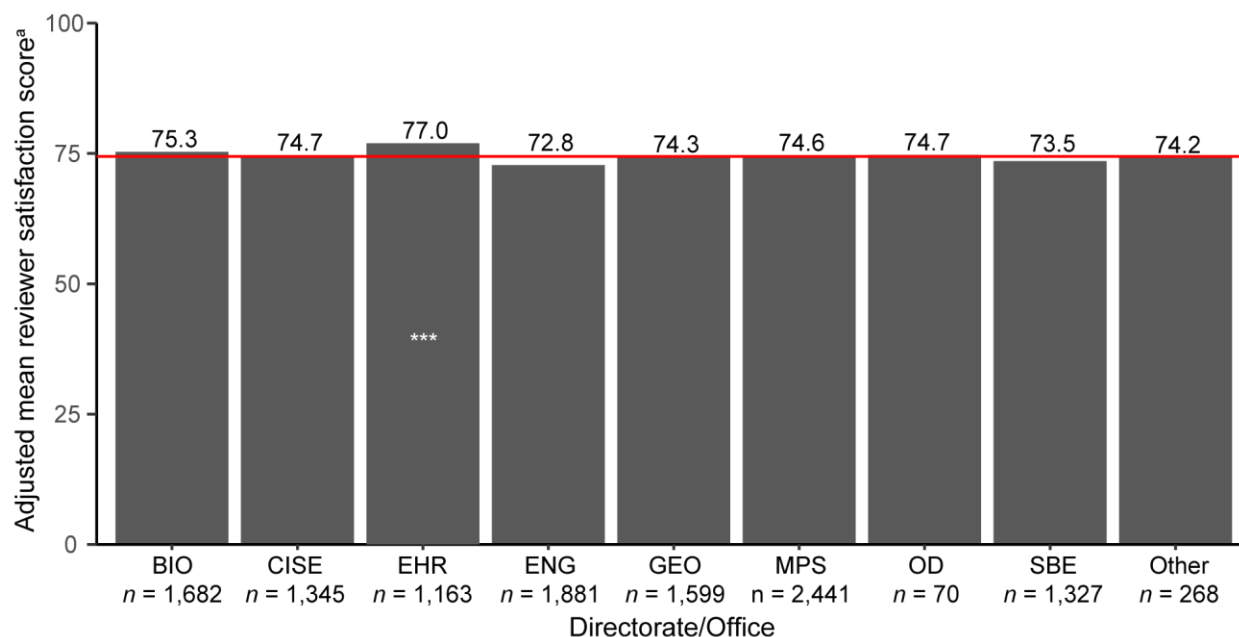
*** indicates significant at $p < 0.001$ relative to the reference group (White, non-Hispanic).

— indicates the overall adjusted mean reviewer satisfaction score.

Source: 2021 Merit Review Survey

Figure 4.2.12 shows that reviewers who identified as Asian, non-Hispanic had higher satisfaction than White, non-Hispanic reviewers (a 2-point difference). Reviewers in the other, NHPI, AIAN race and ethnicity category and those who did not wish to provide their race and ethnicity information had lower satisfaction than White, non-Hispanic reviewers (a 7- and 4-point difference, respectively).

Figure 4.2.13. Reviewer satisfaction index scores, by directorate



Note: $N = 11,776$

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; OD = Office of the Director; SBE = Social, Behavioral, and Economic Sciences; Other = Directorate/office not listed in the survey

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.2.3 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (MPS).

— indicates the overall adjusted mean applicant satisfaction score.

Source: 2021 Merit Review Survey

Figure 4.2.13 shows that reviewer satisfaction scores varied by NSF directorate/office affiliation, ranging from 73 for those associated with ENG to 77 for those associated with EHR. See appendix table C.2.3 for detailed regression results.

4.3 Burden in the Merit Review Process

Burden Highlights

- Applicants submitted an average of three proposals and spent an average of 126 hours preparing their most recent proposal.
- Reviewers reviewed an average of 10 proposals and spent an average of 7 hours reading the proposal and writing and submitting their most recent review.
- Respondents were significantly more likely to report declining a review request in 2021 than in 2019.
- Applicant and reviewer burden varied across NSF directorates. Although applicants affiliated with the directorates of Computer and Information Science and Engineering (CISE), EHR, and ENG submitted large numbers of proposals and spent many hours preparing them relative to applicants in other directorates, they were more likely to perceive NSF's merit review process as requiring less effort than other federal agencies' proposal submission systems. Reviewers affiliated with CISE, EHR, and ENG also conducted large numbers of reviews but spent fewer hours preparing them relative to those in other directorates and were more likely to decline invitations to review.

This study used several survey items to evaluate the burden of submitting proposals to NSF or reviewing proposals on behalf of NSF via the merit review process. The team operationalized applicant burden as the number of proposals submitted, the number of hours required to prepare one's most recent proposal, and the amount of effort required to complete an NSF proposal compared with other agencies. Reviewer burden was operationalized as the number of proposals reviewed, the number of hours required to complete one's most recent review, and whether a respondent had declined to review an NSF proposal.

All analyses are modelled controlling for key respondent characteristics (i.e., role in the merit review process, race and ethnicity, gender, disability status, NSF directorate or office, institution affiliation, and early career status).

4.3.1 Applicant burden

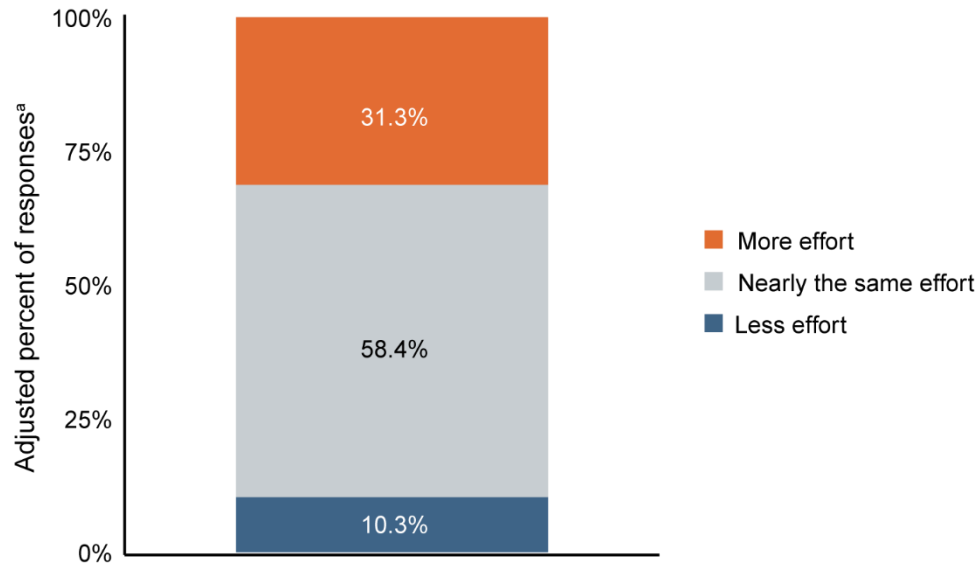
The survey asked applicants to provide the number of proposals they submitted to their primary directorate or office since October 1, 2018, and the number of hours they spent preparing their most recent proposal (including writing, formatting, and submitting).¹⁰ Overall, applicants submitted an average of three proposals and spent an average of 126 hours preparing their most recent proposal (see appendix table C.1.3). In 2019, the plurality of applicants (28 percent) reported they spent between 81 and 120 hours preparing their most recent proposal (Hare et al. 2020); 23 percent of 2021 survey respondents also reported spending between 81 and 120 hours.^{11, 12}

¹⁰ The number of proposals submitted and hours spent preparing a proposal were log transformed, so the findings presented in text are exponentiated regression coefficients.

¹¹ This study could not examine change in the number of proposals submitted over time because the lengths of the two survey reference periods were not comparable (47 months in 2019 versus 36 months in 2021; see section 3.2 for more details).

¹² Analysis of change in the number of hours to complete a most recent proposal is limited to a descriptive comparison because the 2021 and 2019 surveys operationalized this variable differently. The 2021 survey used an open-ended response option, which enabled greater detail to be gathered, while the 2019 survey used a categorical variable with hour ranges (less than 40 hours, 41–80 hours, 81–120 hours, etc.).

Figure 4.3.1. Level of effort by applicants to prepare and submit an NSF proposal compared with other federal agencies



Note: $N = 8,493$

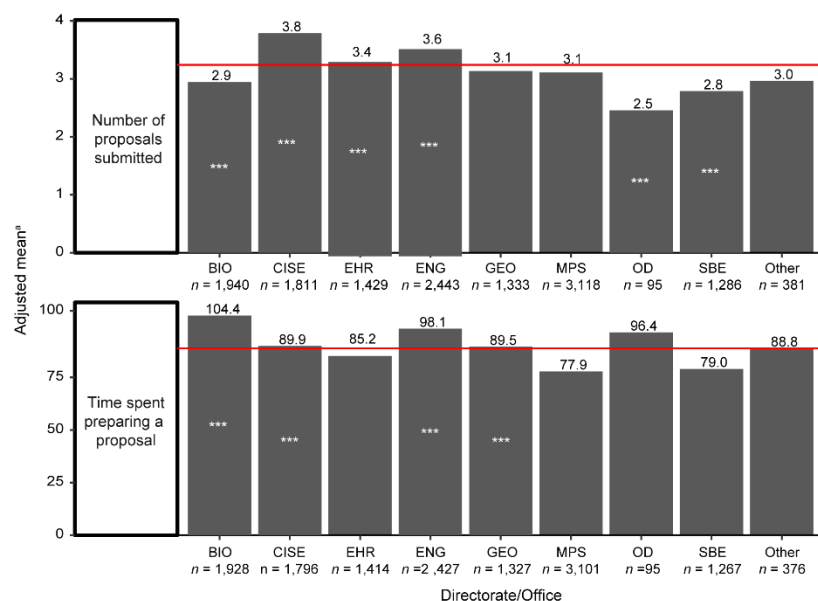
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.3.2 for detailed regression results.

Source: 2021 Merit Review Survey

Applicants were also asked to evaluate if preparing and submitting a proposal to NSF required more or less effort compared with other federal agencies. As figure 4.3.1 illustrates, a majority of applicants perceived the level of effort required of NSF's proposal process to be comparable to other federal agencies, nearly one-third perceived NSF's process to require more effort, and 1 in 10 perceived the process to require less effort.

As figures 4.3.2 and 4.3.3 illustrate, applicant burden varied by NSF directorate or office affiliation.

Figure 4.3.2. Number of proposals submitted and hours spent to prepare most recent proposal, by directorate



Note: Number of proposals submitted $N = 13,836$; time to prepare a proposal $N = 13,731$

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; OD = Office of the Director; SBE = Social, Behavioral, and Economic Sciences; Other = Directorate/office not listed in the survey

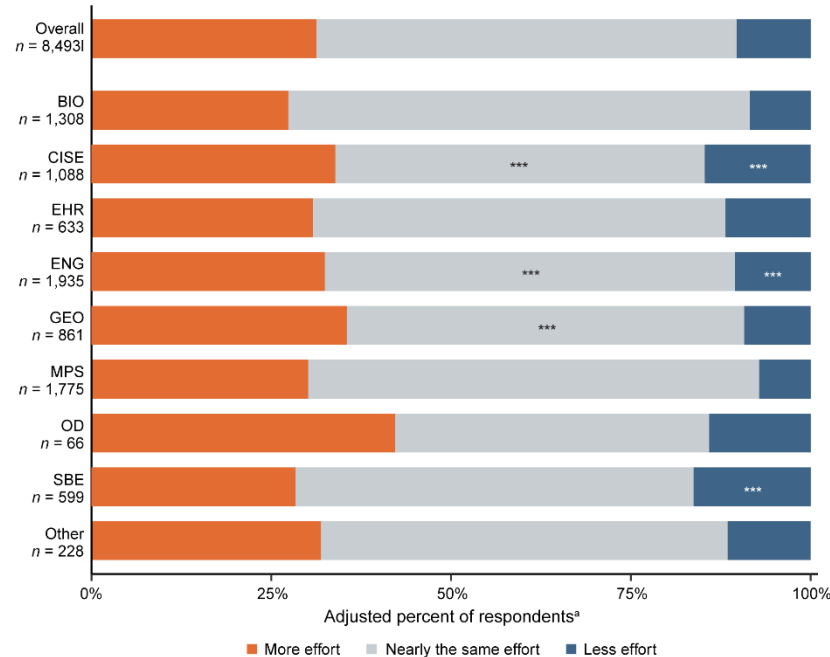
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.3.1 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (MPS).

— indicates the overall adjusted mean number of proposals submitted and time spent preparing a proposal, respectively.

Source: 2021 Merit Review Survey

Figure 4.3.3. Level of effort to prepare and submit an NSF proposal compared with other federal agencies, by directorate



Note: $N = 8,493$

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; OD = Office of the Director; SBE = Social, Behavioral, and Economic Sciences; Other = Directorate/office not listed in the survey

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.3.2 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (MPS).

Source: 2021 Merit Review Survey

The number of proposals applicants submitted varied by NSF directorates/offices, ranging from 2.5 for those associated with OD (including OIA and OISE) to 3.8 for those associated with CISE (figure 4.3.2). The time spent preparing and submitting proposals ranged from 78 hours to 104 hours for those associated with MPS and Biological Sciences (BIO), respectively. Applicant perceptions of the level of effort to submit proposals to NSF compared with other federal agencies varied by directorate/office. Applicants affiliated with CISE, EHR, ENG, and the Directorate for Social, Behavioral, and Economic Sciences (SBE) were more likely than their MPS peers (the most common directorate) to perceive NSF's merit review process as requiring less effort than other federal agencies' processes (figure 4.3.3).

Results provide some evidence to suggest applicant burden may also vary by race and ethnicity, career stage, gender, institutional affiliation, and role in the merit review process. For example, applicants who identified as Asian or early career submitted more proposals and spent more time preparing them than their White or later career counterparts. Female applicants or applicants from for-profit organizations submitted fewer proposals but spent more time preparing them than their male or R1/non-MSI-affiliated counterparts. Applicants who had also served as reviewers submitted more proposals but spent less time preparing them than those who had not also served as reviewers. See appendix table C.3.1 for details.

4.3.2 Reviewer burden

The survey asked reviewers about the number of proposals they had reviewed since October 1, 2018, and the number of hours they spent conducting their most recent review (including reading the proposal and writing and submitting the review).¹³ Overall, reviewers reviewed an average of 10 proposals and spent an average of 7 hours reviewing the most recent proposal (see appendix table C.1.4). The latter finding represents a decrease in hours spent conducting a review since the 2019 survey (see appendix table C.3.7; Hare et al. 2020). Sixty-eight percent of reviewers declined at least once to serve as a reviewer for NSF. Compared with 2019, reviewers were 10 percentage points more likely to decline a review request in 2021 (see appendix table C.3.9; Hare et al. 2020). A common theme among open-ended comments from respondents who declined to review was lack of time, including challenges with work-life balance because of changes brought on by the COVID-19 pandemic.

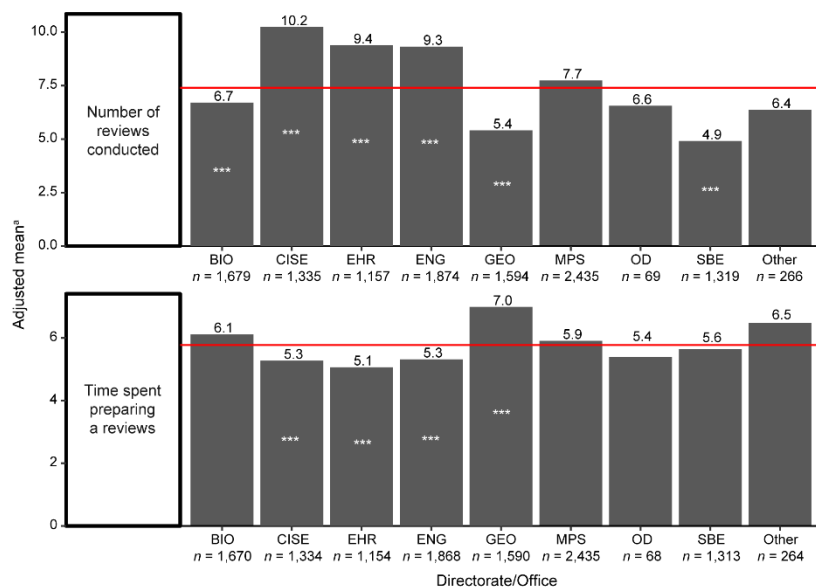
"Reviewing is a large time commitment. In person is easily 12 to 14 hours per day. No one has that sort of time while working at home."

"[The] COVID pandemic brought high workload for work-family balance, especially for faculty with young children."

Similar to applicant burden, reviewer burden varied by NSF directorate or office, as figures 4.3.4 and 4.3.5 illustrate.

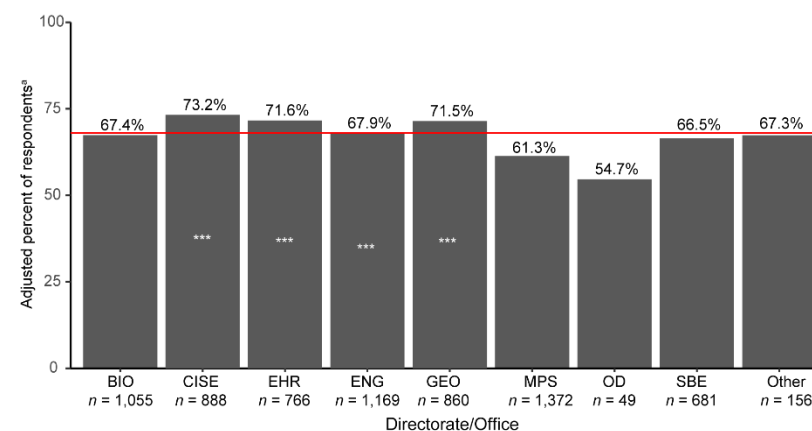
¹³ The number of proposals reviewed and hours spent conducting a proposal review were log transformed, so the findings presented in text are exponentiated regression coefficients.

Figure 4.3.4. Number of reviews conducted and hours spent conducting a review, by directorate



Note: Number of reviews conducted $N = 11,728$; time spent conducting a review $N = 11,696$
 BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; OD = Office of the Director; SBE = Social, Behavioral, and Economic Sciences; Other = Directorate/office not listed in the survey.
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.3.3 for detailed regression results.
 *** indicates significant at $p < 0.001$ relative to the reference group (MPS).
 — indicates the overall adjusted mean number of reviews conducted and hours spent conducting most recent review, respectively.
 Source: 2021 Merit Review Survey

Figure 4.3.5. Percentage of respondents who declined to serve as a reviewer, by directorate



Note: $N = 6,996$
 BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; OD = Office of the Director; SBE = Social, Behavioral, and Economic Sciences; Other = Directorate/office not listed in the survey.
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.3.5 for detailed regression results.
 *** indicates significant at $p < 0.001$ relative to the reference group (MPS).
 — indicates the overall adjusted percentage of respondents who declined to serve as a reviewer.
 Source: 2021 Merit Review Survey

The number of reviews conducted by reviewers varied by NSF directorates/offices, ranging from 4.9 for those associated with SBE to 10.2 for those associated with CISE (figure 4.3.4). The time spent reviewing their most recent proposal ranged from 5.1 hours to 7.0 hours for those associated with EHR and Geosciences (GEO), respectively (figure 4.3.5). The percentage of respondents who declined to serve as a reviewer ranged from 54.7 for those associated with OD to 73.2 for those associated with CISE (figure 4.3.5).

Results provide some evidence to suggest reviewer burden may also vary by gender, race and ethnicity, career stage, and institution. For example, female reviewers conducted more reviews than their male counterparts but spent similar amounts of time conducting their most recent review. Conversely, Asian and Black reviewers conducted similar numbers of reviews to their White, non-Hispanic counterparts but spent more time conducting their most recent review. Finally, early career reviewers or those affiliated with non-R1/MSIs, for-profit organizations, and government agencies conducted fewer reviews than their later-career or R1/non-MSI-affiliated counterparts but spent similar amounts of time conducting their most recent review. See appendix table C.3.3 for details.

4.4 Proposal and Review Quality in the Merit Review Process

Quality Highlights

- Across all respondents (applicants and reviewers), improving the quality of feedback to PIs in the form of comments in written reviews was the most commonly endorsed way to improve the merit review process.
- Less than half of applicants agreed or strongly agreed that written reviews were of high quality.
- About three out of four reviewers indicated the majority of proposals they reviewed in recent years have been of high quality; however, this number was higher among respondents in 2019.
- Applicants who had also reviewed proposals rated written reviews as higher quality than their applicant-only counterparts, while reviewers who had also submitted proposals rated proposal quality lower than their reviewer-only peers.
- Female respondents rated written reviews and proposals as higher quality than males.
- Asian, non-Hispanic respondents rated written reviews as higher quality but proposals as lower quality than White, non-Hispanic respondents.

This study examined the quality of the merit review process in three ways. First, the survey asked all respondents a general question about factors they thought would have the most significant effect on improving the merit review process. Second, applicants were asked about the quality of the review process, including the quality of written reviews. Finally, reviewers were asked about the quality of the proposals they had reviewed.

All analyses are modelled controlling for key respondent characteristics (i.e., role in the merit review process, race and ethnicity, gender, disability status, NSF directorate or office, institution affiliation, and early career status).

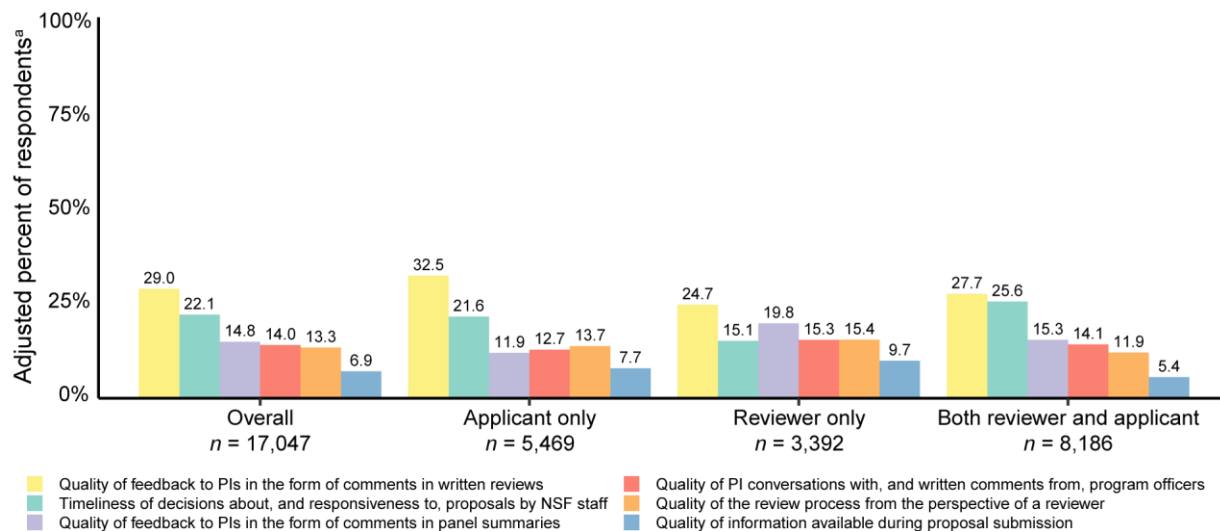
4.4.1 Primary factors that would improve the merit review process

Applicants and reviewers were asked to indicate which of the following factors would have the most significant effect on improving the merit review process:

- timeliness of decisions about and responsiveness to proposals by NSF staff;
- quality of feedback to PIs in the form of comments in written reviews;

- quality of feedback to PIs in the form of comments in panel summaries;
- quality of PI conversations with and written comments from program officers;
- quality of information available during proposal submission; and
- quality of the review process from the perspective of reviewers.

Figure 4.4.1. Areas of the merit review process to improve, by role in the merit review process



Note: N = 17,047

^a Estimates are adjusted after controlling for key respondent characteristics. Modeled estimates are used descriptively in figure 4.4.1 to illustrate the most selected improvement area; see appendix table C.4.5 for detailed regression results.

Estimates are adjusted after controlling for key respondent characteristics.

Source: 2021 Merit Review Survey

As figure 4.4.1 illustrates, the most frequently cited area for improvement was the quality of feedback to PIs in the form of comments in written reviews (reported by 29 percent of respondents). This number was consistent with responses provided by 2019 survey respondents, most of whom endorsed improving the quality of written reviews (31 percent; Hare et al. 2020). Although the survey did not directly ask respondents to expand on their reports related to review quality, lack of review quality was a common theme among open-ended comments:

"Proposals that are funded are not transformative and reward the same people for incremental work. Comments are not useful and sometimes reveal that reviewers did not thoroughly read the proposal."

"Many of my proposals received irrelevant and somewhat unprofessional comments and feedback. In one case it seemed that they totally reviewed a different proposal!"

"Not all reviewers are diligent in reading applications carefully or have the needed expertise. This leads to wide discrepancies in scores among reviewers and conflicting feedback from reviewers."

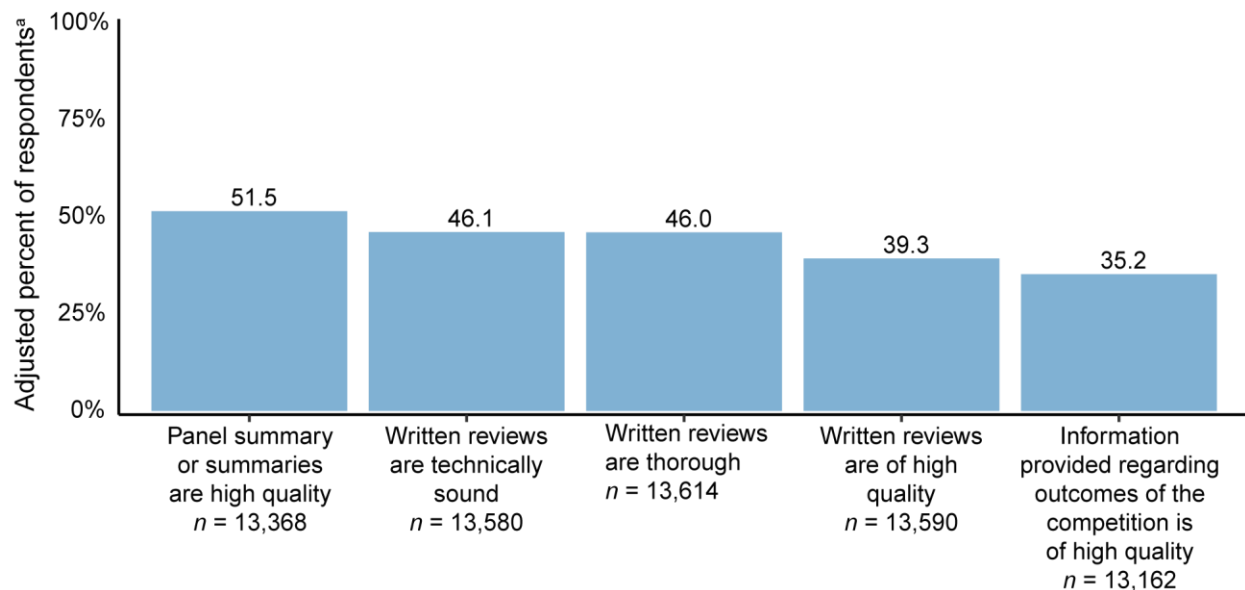
"The process gives a reviewer power but they do not always have the background or knowledge needed to make an informed decision."

4.4.2 Applicant perceptions of quality

Applicants were asked about the quality of information they received from NSF during submission and the reviews they received in response to their proposals. Overall, 70 percent of applicants were somewhat or very satisfied with the quality of information provided by NSF during the proposal submission process, which is 5 percentage points lower compared with 2019 (see appendix table C.4.6.B; Hare et al. 2020).

Applicants were asked to which extent they agreed or disagreed that the reviews they received in response to their proposals were thorough, technically sound, and of high quality and that the information on outcomes of the competition was of high quality.

Figure 4.4.2. Overall applicant perceptions of quality: percentage of respondents who agreed or strongly agreed



^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.4.2.A–C.4.2.C for detailed regression results.

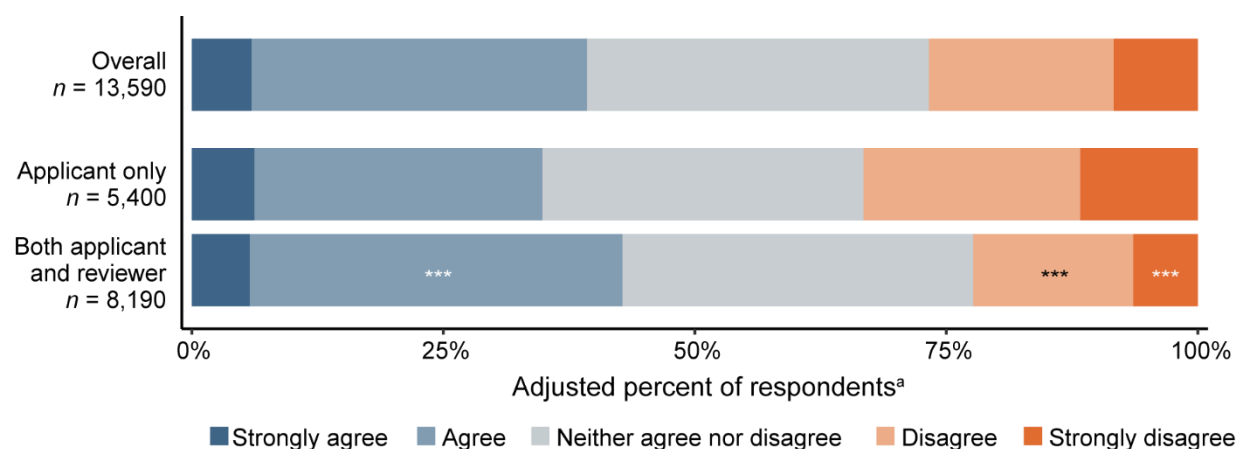
Source: 2021 Merit Review Survey

As figure 4.4.2 shows, just over half of applicants agreed or strongly agreed that panel summaries were of high quality (52 percent). Nearly half of applicants agreed or strongly agreed that the written reviews were thorough, technically sound, and of high quality (46, 46, and 39 percent, respectively). About one-third agreed or strongly agreed that the information provided by NSF regarding the outcomes of the competition was of high quality (35 percent).¹⁴

¹⁴ The addition of a middle neutral value, resulting in strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, and not applicable, limited the cross-year comparison for these items.

As figures 4.4.3–4.4.6 illustrate, applicant perceptions of quality varied by role in the merit review process, directorate, gender, and race and ethnicity.

Figure 4.4.3. Applicant agreement that written reviews were of high quality, by role in the merit review process



Note: N = 13,590

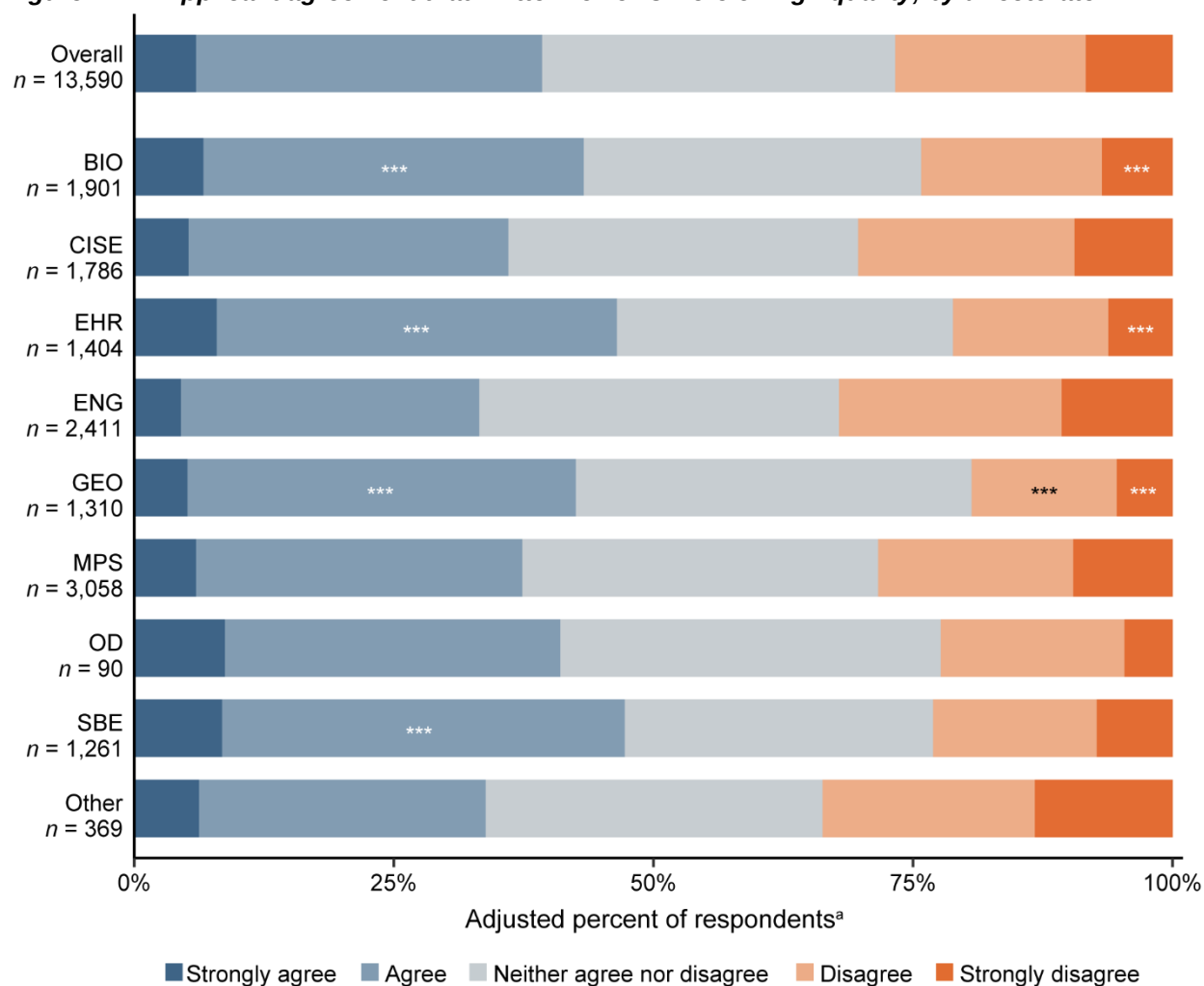
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.4.2.B for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (applicant only).

Source: 2021 Merit Review Survey

As figure 4.4.3 illustrates, applicants who had also reviewed proposals were more likely to agree that written reviews were of high quality than their applicant-only counterparts (by 8 percentage points).

Figure 4.4.4. Applicant agreement that written reviews were of high quality, by directorate



Note: N = 13,590

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.4.2.B for detailed regression results.

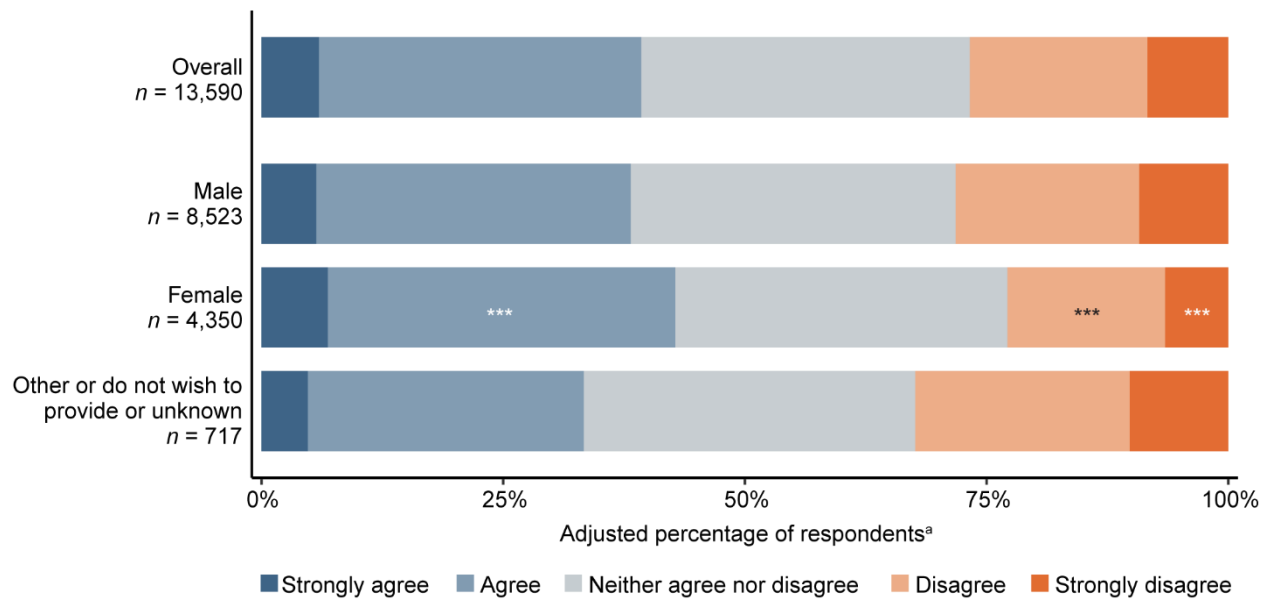
BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; OD = Office of the Director; SBE = Social, Behavioral, and Economic Sciences; Other = Directorate/office not listed in the survey

*** indicates significant at $p < 0.001$ relative to the reference group (MPS).

Source: 2021 Merit Review Survey

Applicant perceptions of review quality varied by directorate/office. They ranged from 33 percent to 47 percent of respondents (associated with ENG and SBE, respectively), reporting they strongly agree or agree that written reviews were of high quality (figure 4.4.4).

Figure 4.4.5. Applicant agreement that written reviews were of high quality, by gender



Note: N = 13,590

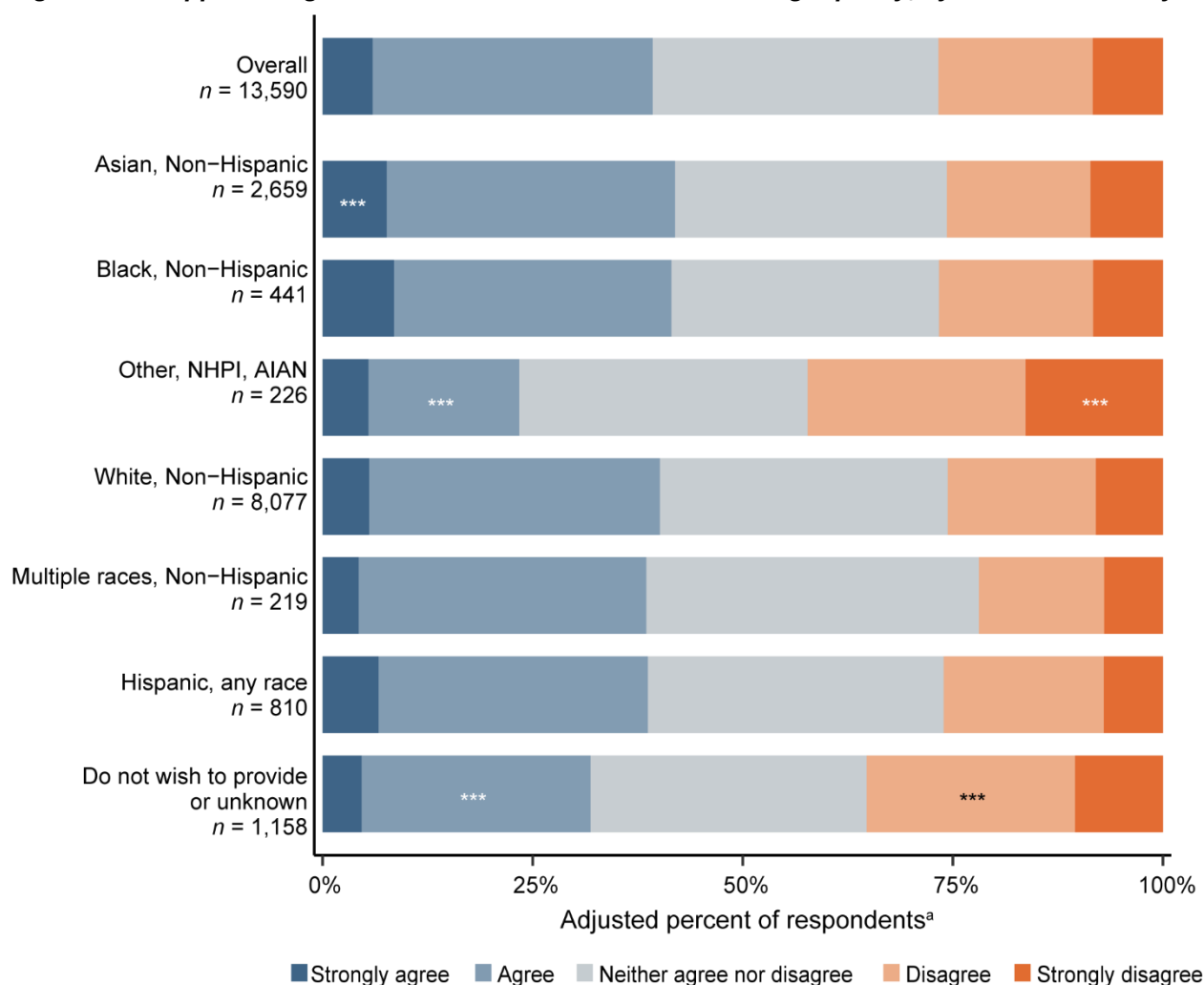
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.4.2.B for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (male).

Source: 2021 Merit Review Survey

As figure 4.4.5 illustrates, female applicants were more likely to agree reviews were of high quality than male applicants (by 3 percentage points).

Figure 4.4.6. Applicant agreement that written reviews were of high quality, by race and ethnicity



Note: N = 13,590

AIAN = American Indian or Alaska Native; NHPI = Native Hawaiian or Other Pacific Islander

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.4.2.B for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (White, non-Hispanic).

Source: 2021 Merit Review Survey

As figure 4.4.6 illustrates, applicants who identified as Asian, non-Hispanic were more likely to strongly agree that reviews were of high quality than White, non-Hispanic applicants (by 2 percentage points).

Results provide some evidence to suggest applicant perceptions of the quality of reviews also varied by early career status, institutional affiliation, and disability status. For example, early career applicants were more likely to agree or strongly agree reviews were of high quality than later career applicants (by 7 and 2 percentage points, respectively). Applicants affiliated with for-profit organizations were more likely to strongly disagree reviews were of high quality than those from R1/non-MSIs (by 5 percentage points). Applicants who reported having a vision, hearing, mobility, or other serious disability were more likely to strongly disagree reviews were of high

quality than those without a disability (by 7 percentage points). See appendix table C.4.2.B for details.

4.4.2.1 Declined applicants' perceptions of quality

Ninety percent of applicants reported they submitted a proposal to NSF that was declined during the survey timeframe (i.e., since October 1, 2018). These individuals were asked about the quality of the reviews accompanying their most recently declined proposal application (see tables C.4.3.A and C.4.3.B for detailed regression findings). They reported the following:

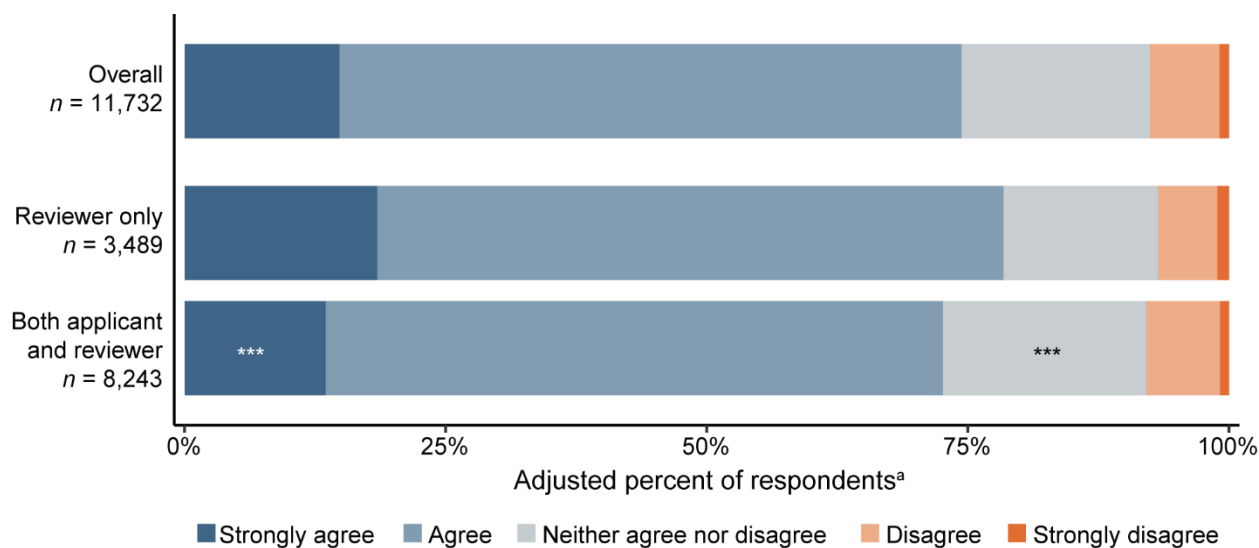
- 55 percent of respondents indicated the reviews moderately or greatly provided useful information for improving their next proposal submission.
- 38 percent of respondents indicated the reviews moderately or greatly improved their understanding of the proposal process.
- 43 percent of respondents indicated the reviews moderately or greatly discouraged them from revising and resubmitting their proposal to NSF.
- 32 percent of respondents indicated the reviews moderately or greatly influenced their decision to submit to another funding agency.

4.4.3 Reviewer perceptions of quality

Seventy-four percent of reviewers agreed or strongly agreed that the majority of proposals they reviewed in recent years were of high quality (see appendix table C.1.4). In comparison, 87 percent of reviewers felt this way in 2019 (Hare et al. 2020).

As with applicant perceptions of quality, reviewer perceptions of quality varied by role in the merit review process, directorate, gender, and race and ethnicity (see figures 4.4.7–4.4.10).

Figure 4.4.7. Reviewer agreement that proposals were of high quality, by role in the merit review process



Note: N = 11,732

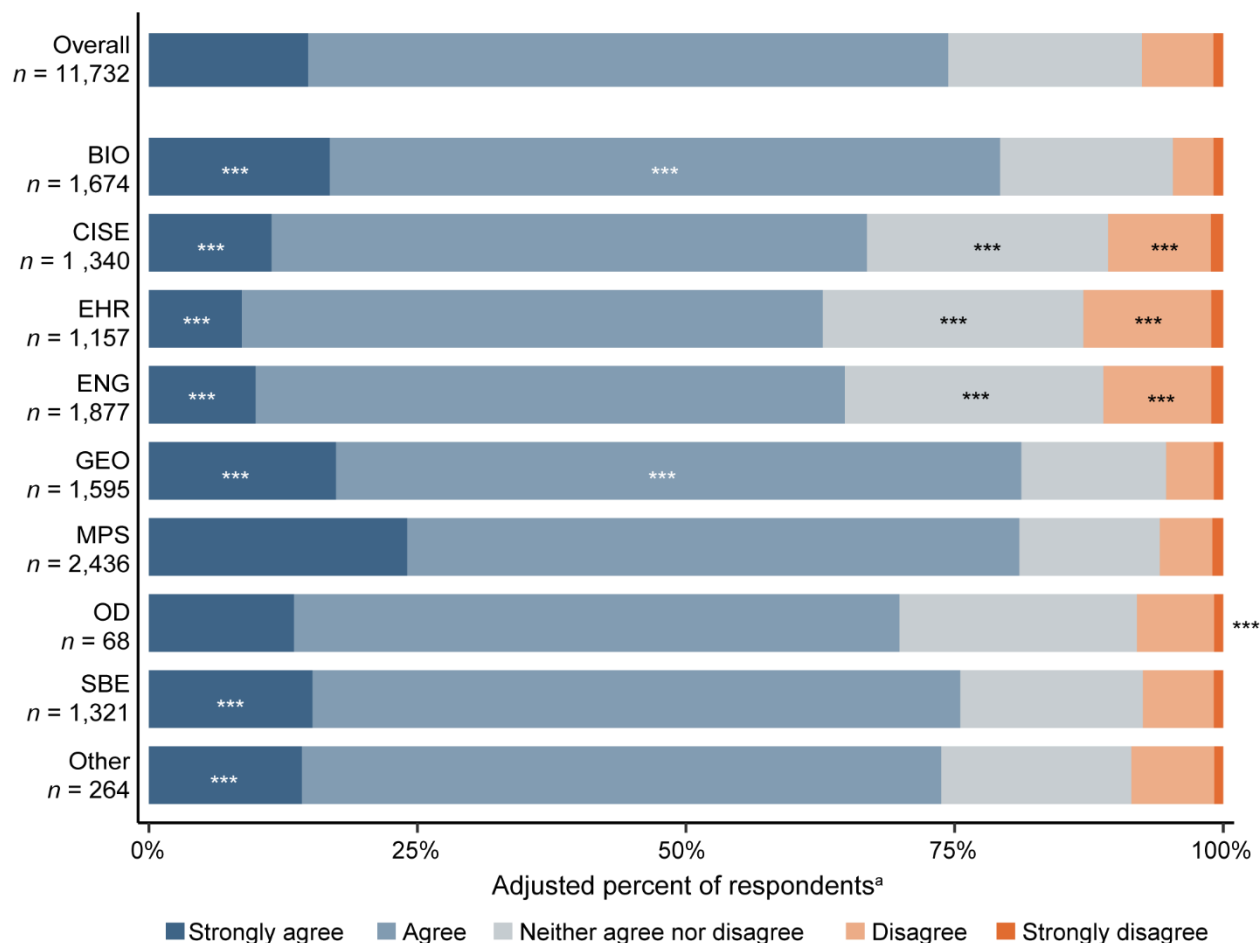
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.4.4 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (reviewer only).

Source: 2021 Merit Review Survey

As figure 4.4.7 shows, reviewers who had also submitted proposals were less likely to strongly agree that written proposals were of high quality than their reviewer-only counterparts (by 5 percentage points).

Figure 4.4.8. Reviewer agreement that proposals were of high quality, by directorate



Note: N = 11,732

BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; MPS = Mathematical and Physical Sciences; OD = Office of the Director; SBE = Social, Behavioral, and Economic Sciences; Other = Directorate/office not listed in the survey.

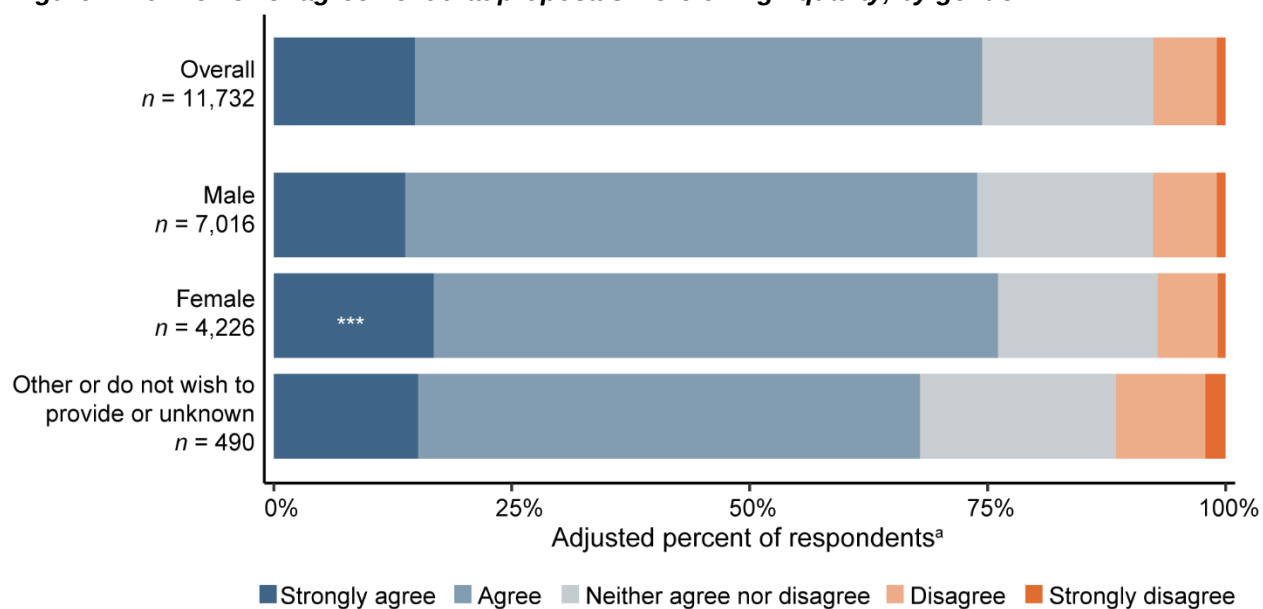
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.4.4 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (MPS).

Source: 2021 Merit Review Survey

Reviewer perceptions of proposal quality varied by directorate/office. They ranged from 63 percent of respondents associated with EHR to 81 percent of respondents associated with GEO reporting they strongly agree or agree that written proposals were of high quality (figure 4.4.8).

Figure 4.4.9. Reviewer agreement that proposals were of high quality, by gender



Note: N = 11,732

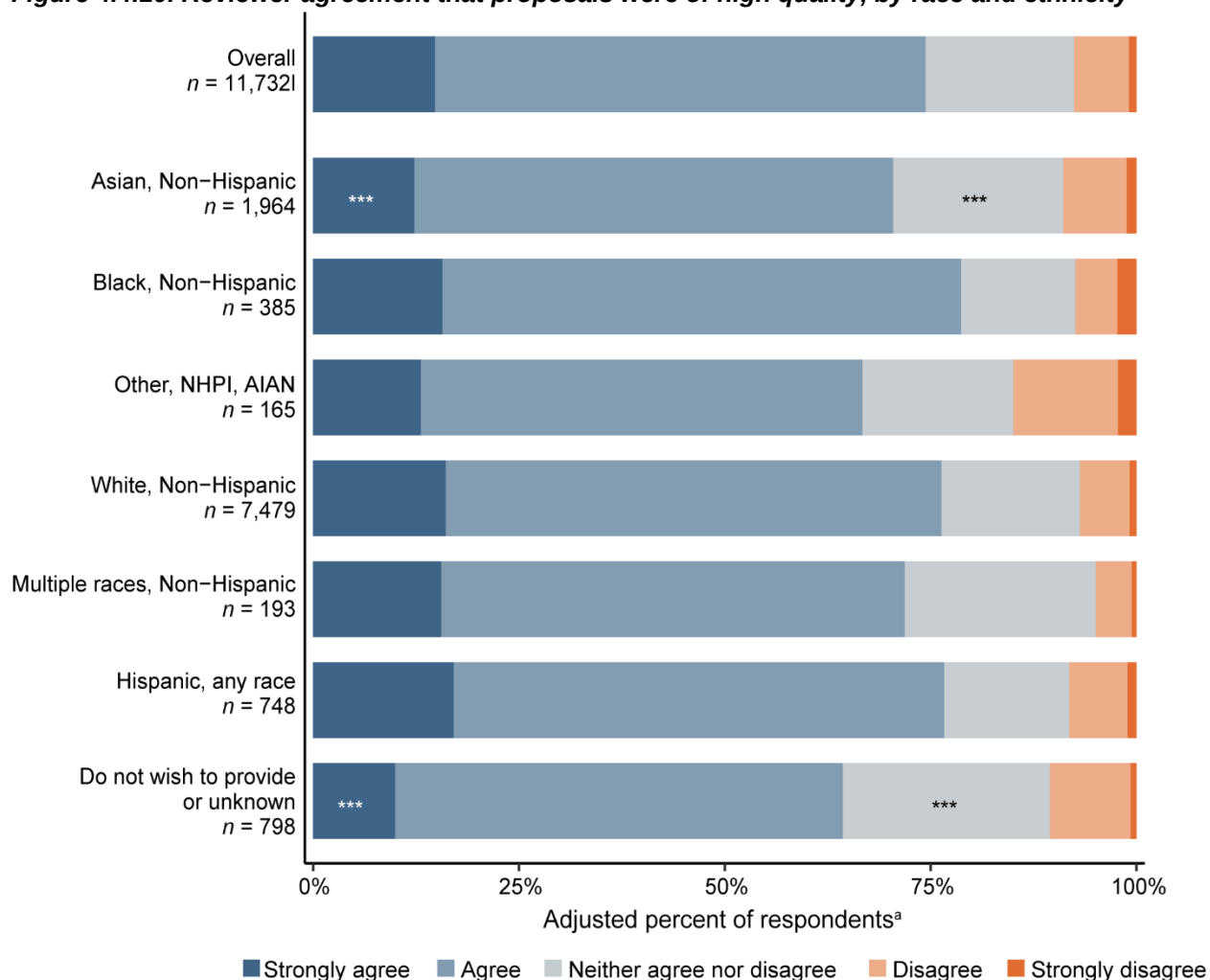
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.4.4 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (male).

Source: 2021 Merit Review Survey

Consistent with applicant perceptions of quality, figure 4.4.9 illustrates that female reviewers were more likely to strongly agree proposals were of high quality than male reviewers (by 3 percentage points).

Figure 4.4.10. Reviewer agreement that proposals were of high quality, by race and ethnicity



Note: N = 11,732

AIAN = American Indian or Alaska Native; NHPI = Native Hawaiian or Other Pacific Islander

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.4.4 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (White, non-Hispanic).

Source: 2021 Merit Review Survey

As figure 4.4.10 illustrates, reviewers who identified as Asian, non-Hispanic were less likely to strongly agree that proposals were of high quality than White, non-Hispanic reviewers (by 4 percentage points).

Results provide some evidence to suggest reviewer perceptions of proposal quality also varied by institutional affiliation. Specifically, reviewers from for-profit organizations were less likely to agree proposals were of high quality than reviewers from R1/non-MSIs (by 10 percentage points). See appendix table C.4.4 for more details.

4.5 Fairness in the Merit Review Process

Fairness Highlights

- Overall, two out of three respondents perceived the merit review process to be fair, which was fewer than in 2019.
- Black, non-Hispanic and Hispanic applicants were less likely to think the process was fair than White, non-Hispanic applicants.
- Respondents who had only reviewed proposals were most likely to think the process was fair; those who had only submitted proposals were least likely to think the process was fair.

This study examined fairness of the merit review process in three ways. First, the survey asked all respondents a general question about whether they thought the merit review process was fair. Second, applicants were asked about the extent to which they agreed or disagreed that individuals submitting proposals were treated fairly. Third, reviewers were asked to what extent they agreed or disagreed that individuals submitting proposals were treated fairly. Individuals who were both applicants and reviewers were asked the question twice.

All analyses are modelled controlling for key respondent characteristics (i.e., role in the merit review process, race and ethnicity, gender, disability status, NSF directorate or office, institution affiliation, and early career status).

4.5.1 Overall perceptions of fairness

Overall, 66 percent of respondents agreed or strongly agreed that NSF's merit review process was fair (see appendix table C.1.2). This proportion was 15 percentage points higher among respondents in 2019 (see appendix table C.5.4.B). Although the survey did not directly ask respondents to expand on their reports related to fairness, perceived biases were a common theme among open-ended comments.

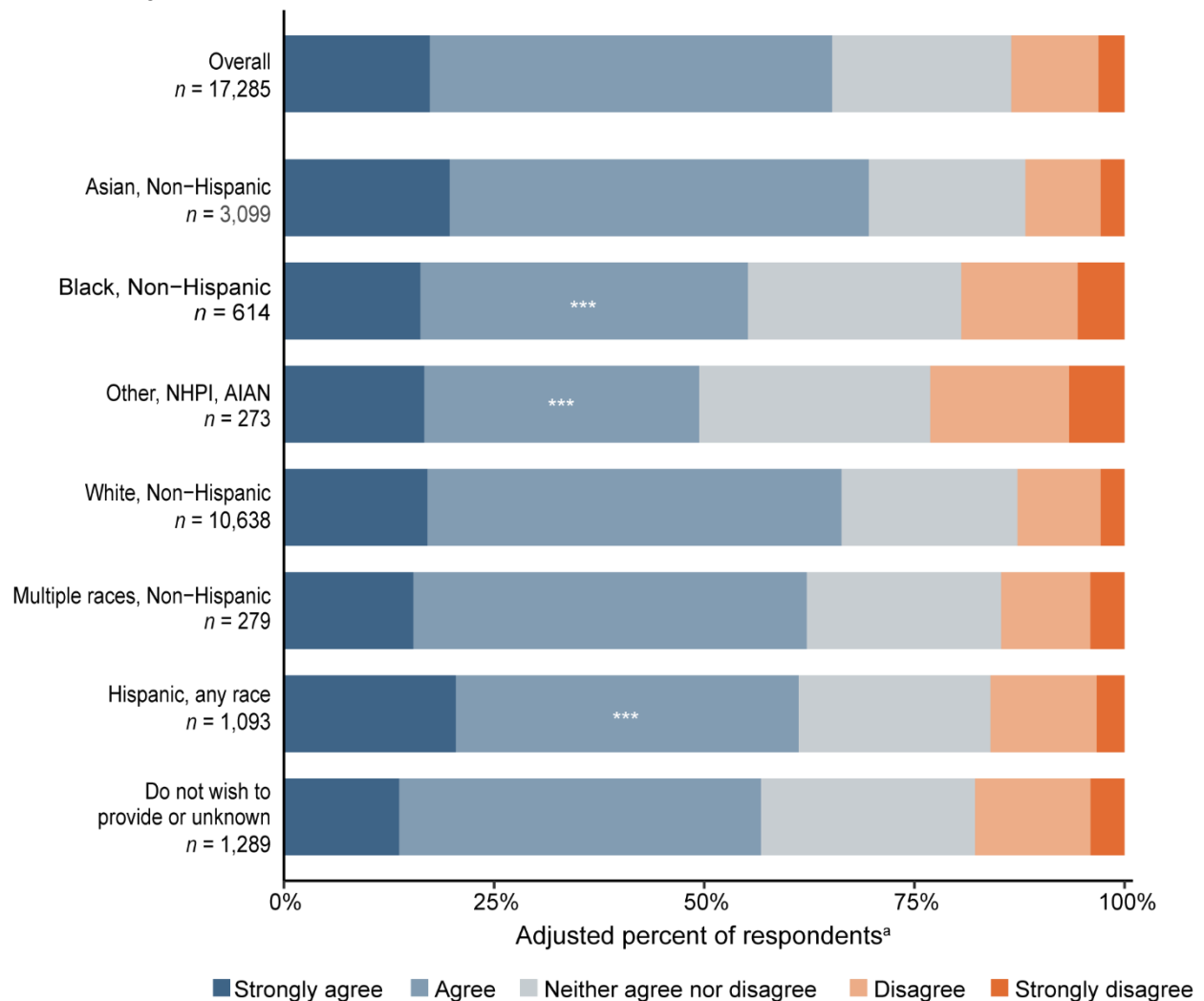
"The NSF review process is usually biased by the views of individual program directors. This affects the selection of reviewers, the quality of the review process and panels, and the reasons that a proposal will be declined or funded."

"NSF review process is biased with lots of hidden discriminations. They favor bigger schools and names rather than great ideas."

"As an 'older' scientist, I often feel a strong ageism bias in the reviews. From some of the comments in the individual reviews, it becomes clear that many reviewers feel that preference should be given to younger early career investigators, and that somebody at my career stage does not need NSF support."

As figures 4.5.1 and 4.5.2 illustrate, overall perceptions of fairness depended on respondents' race and ethnicity and role in the merit review process.

Figure 4.5.1. Applicant and reviewer agreement that the NSF merit review process is fair, by race and ethnicity



Note: N = 17,285

AIAN = American Indian or Alaska Native; NHPI = Native Hawaiian or Other Pacific Islander

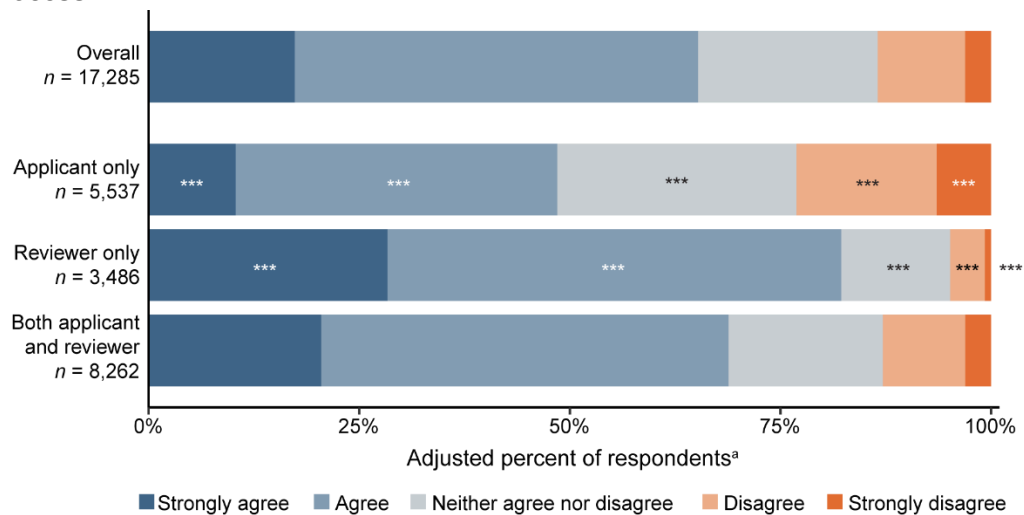
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.5.1 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (White, non-Hispanic).

Source: 2021 Merit Review Survey

As figure 4.5.1 shows, respondents who identified as Black, non-Hispanic; Hispanic; or other, NHPI, AIAN were less likely than White, non-Hispanic respondents to agree that the merit review process was fair (10, 9, and 17 percentage point difference, respectively).

Figure 4.5.2. Applicant and reviewer agreement that the NSF merit review process is fair, by role in the merit review process



Note: N = 17,285

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.5.1 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (both reviewer and applicant).

Source: 2021 Merit Review Survey

Figure 4.5.2 shows that respondents who were only applicants were less likely, and those who were only reviewers were more likely, to strongly agree that the merit review process was fair relative to those who had both submitted and reviewed proposals (by 10 and 8 percentage points, respectively).

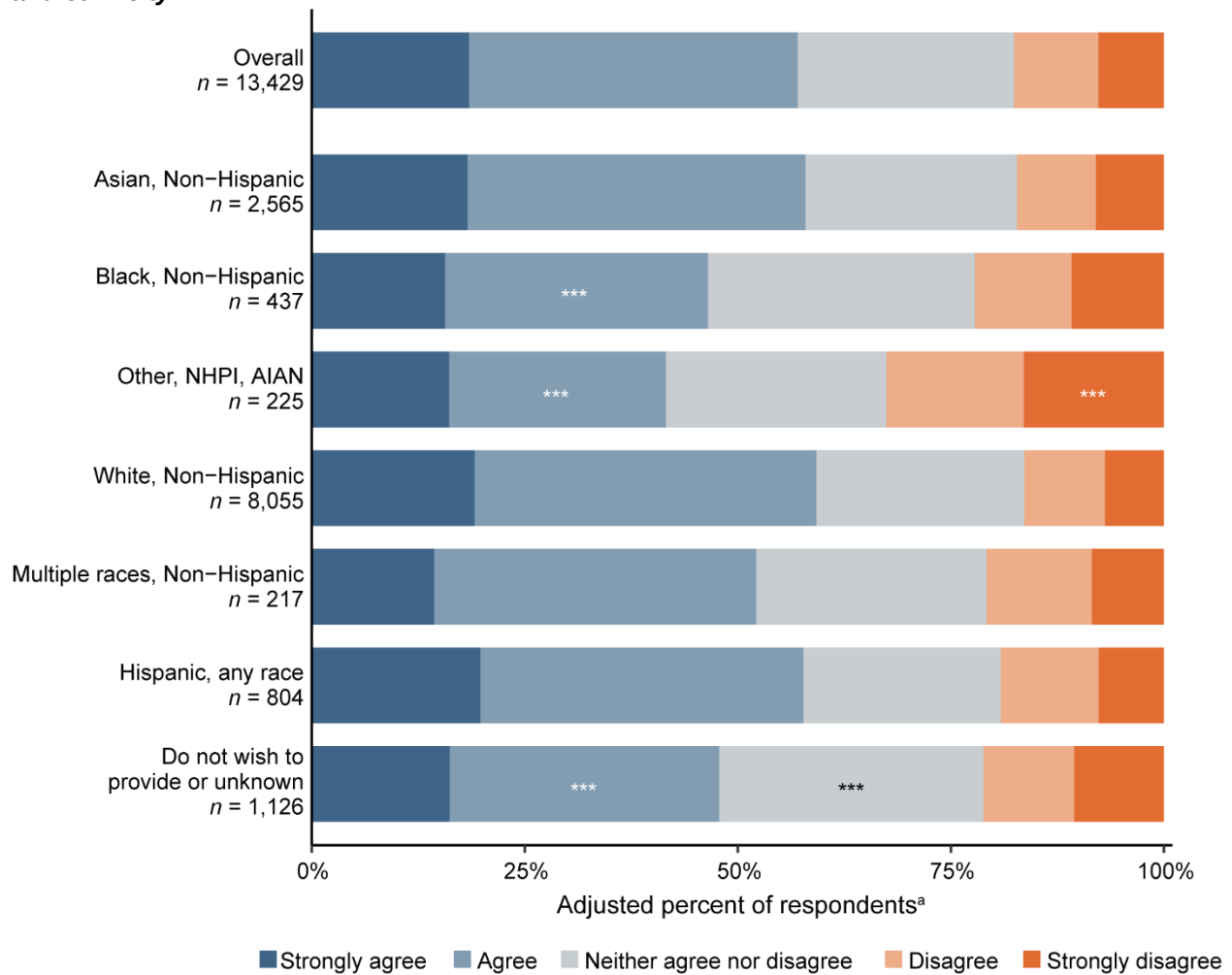
Results provide some evidence to suggest perceptions of overall fairness also varied by gender and disability status. Female respondents were less likely than male respondents to strongly agree that the merit review process was fair (by 2 percentage points). Respondents with hearing, mobility, vision, or other disabilities were less likely to strongly agree or agree that the merit review process was fair compared with their counterparts without disabilities (by 5 and 8 percentage points, respectively). See appendix table C.5.1 for detailed regression results.

4.5.2 Applicant perceptions of fairness

Based on their experience submitting proposals to NSF, applicants were asked to which extent they agreed or disagreed that individuals submitting proposals were treated fairly. Overall, 57 percent agreed or strongly agreed that applicants were treated fairly (see appendix table C.1.3).

Similar to overall perceptions of fairness, figures 4.5.3 and 4.5.4 show applicant perceptions of fairness varied by race and ethnicity and role in the merit review process.

Figure 4.5.3. Applicant agreement that individuals submitting proposals are treated fairly, by race and ethnicity



Note: N = 13,429

AIAN = American Indian or Alaska Native; NHPI = Native Hawaiian or Other Pacific Islander

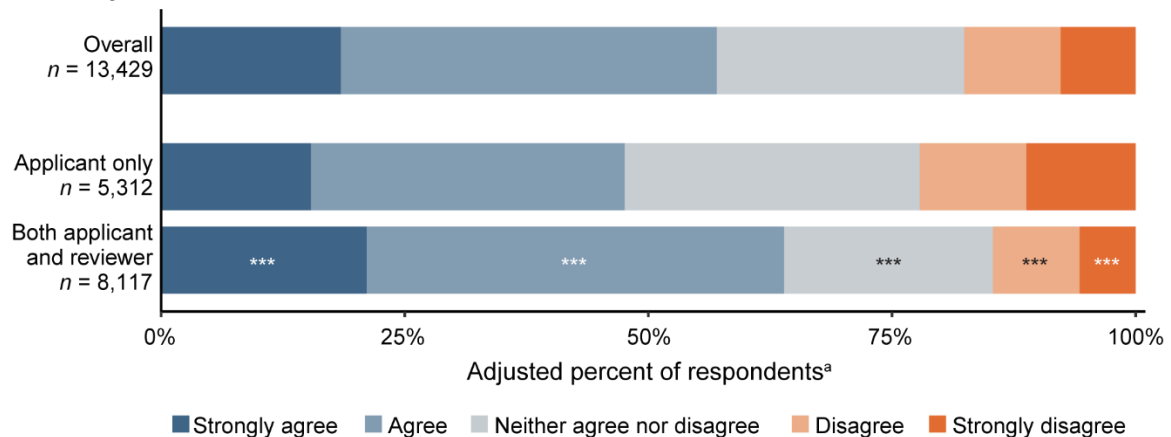
^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.5.2 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (White, non-Hispanic).

Source: 2021 Merit Review Survey

As figure 4.5.3 illustrates, Black, non-Hispanic applicants and those identifying as other, NHPI, or AIAN were less likely than their White, non-Hispanic peers to agree that applicants were treated fairly (by 9 and 15 percentage points, respectively).

Figure 4.5.4 Applicant agreement that individuals submitting proposals are treated fairly, by role in merit review process



Note: N = 13,429

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.5.2 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (applicant only).

Source: 2021 Merit Review Survey

Figure 4.5.4 shows that compared with applicants who only submitted NSF proposals, those who also reviewed proposals were more likely to agree and strongly agree that applicants were treated fairly (by 11 and 6 percentage points, respectively).

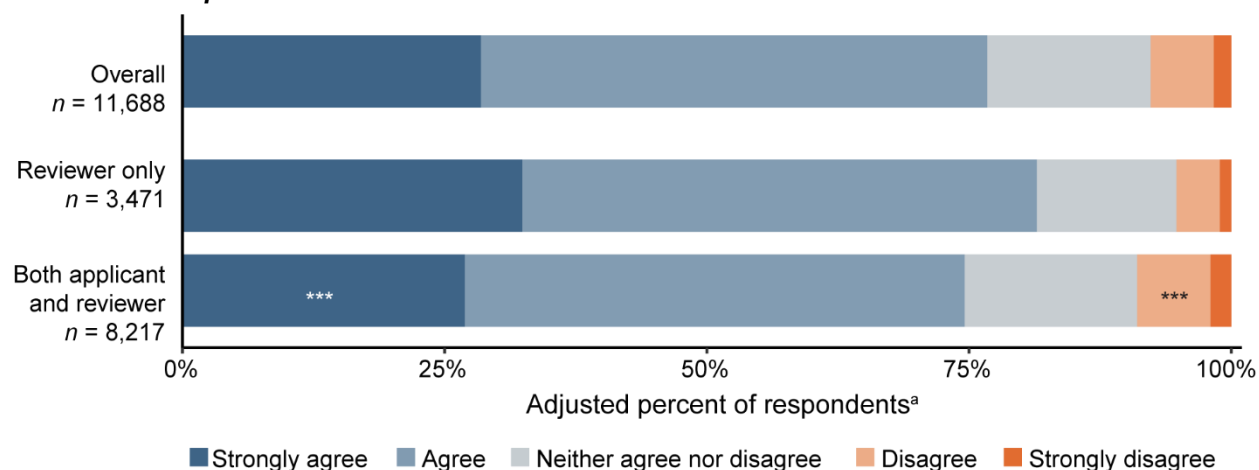
Results provide some evidence to suggest applicant perceptions of fairness also varied by disability status. Applicants with a hearing, mobility, vision, or other disability were 8 percentage points more likely to strongly disagree that applicants were treated fairly, compared with those without a disability. See appendix table C.5.2 for details.

4.5.3 Reviewer perceptions of fairness

Similar to applicants, reviewers were also asked to which extent they agreed or disagreed that individuals submitting proposals were treated fairly, based on their experience reviewing proposals for NSF. Overall, 77 percent agreed or strongly agreed that individual applicants were treated fairly (see appendix table C.1.4).

Figure 4.5.5 shows that reviewer perceptions of fairness varied by role in the merit review process.

Figure 4.5.5. Reviewer agreement that individuals submitting proposals are treated fairly, by role in merit review process



Note: N = 11,688

^a Estimates are adjusted after controlling for key respondent characteristics; see appendix table C.5.3 for detailed regression results.

*** indicates significant at $p < 0.001$ relative to the reference group (reviewer only).

Source: 2021 Merit Review Survey

As figure 4.5.5 illustrates, reviewers who had experience as applicants were less likely than reviewers who had no experience submitting proposals to strongly agree that applicants were treated fairly (by 5 percentage points).

Results provide some evidence to suggest reviewer perceptions of fairness also varied by gender, directorate, race and ethnicity, early career status, and institutional affiliation. Female reviewers were less likely than male reviewers to strongly agree that those submitting proposals were treated fairly (by 4 percentage points). The proportion of reviewers agreeing or strongly agreeing that applicants were treated fairly ranged from 72 percent for GEO reviewers to 85 percent for EHR reviewers. Reviewers in the other, NHPI, AIAN race and ethnicity category were less likely than White, non-Hispanic reviewers to agree that applicants were treated fairly (by 13 percentage points). Early career reviewers were less likely than those further along in their career to strongly agree that applicants were treated fairly (by 6 percentage points). Reviewers associated with for-profit organizations were more likely than those from R1/non-MSI to strongly agree that those submitting proposals were treated fairly (by 14 percentage points). See appendix table C.5.3 for details.

5 Conclusions

This chapter summarizes applicant and reviewer perceptions of the NSF merit review process related to satisfaction, burden, quality, and fairness. This chapter also provides the study team's assessment of the potential policy and research implications of these findings.

5.1 Summary of Key Findings

In general, applicants and reviewers were satisfied with the merit review process. Assessments of satisfaction, burden, quality, and fairness were mostly positive, particularly for those individuals who both applied for funding and served as reviewers.

5.1.1 *Satisfaction with the merit review process*

Assessments of satisfaction were mostly positive. Over half of respondents were satisfied with the merit review process. Satisfaction varied by respondent role, institutional affiliation, disability status, and race and ethnicity. Respondents who had only reviewed proposals were more satisfied and those who had only submitted proposals were less satisfied than respondents who had done both. Respondents from non-R1/MSIs were more satisfied than those from other organizations. Respondents with disabilities were less satisfied than those without disabilities. Asian, non-Hispanic respondents were more satisfied with the merit review process than White, non-Hispanic respondents. In contrast, respondents who identified as American Indian, Native Hawaiian, or selected other race were less satisfied than White, non-Hispanic respondents.

5.1.2 *Assessment of burden of the merit review process*

Applicants reported submitting an average of three proposals since October 1, 2018, and spending an average of 126 hours preparing their most recent proposal. Reviewers reported conducting an average of 10 reviews since October 1, 2018, and spending an average of 7 hours conducting their most recent review. Assessment of burden varied by NSF directorate/office affiliation. Relative to their peers, applicants affiliated with CISE, EHR, and ENG submitted more proposals and spent more hours preparing them and conducted more reviews but spent fewer hours preparing them.

5.1.3 *Perceptions of quality within the merit review process*

Perceptions of quality were mixed. Most reviewers (74 percent) indicated the majority of proposals they reviewed in recent years were of high quality. Less than half of applicants agreed or strongly agreed that the written reviews were thorough. Perceptions of quality varied by role in the merit review process, NSF directorate/office affiliation, gender, and race and ethnicity. Applicants who had also reviewed proposals rated the quality of written reviews more highly than their applicant-only counterparts, while reviewers who had submitted proposals rated proposal quality less highly than their reviewer-only peers. Respondents affiliated with EHR, SBE, GEO, and BIO rated written reviews as higher quality but proposals as lower quality compared with respondents in other directorates. Female respondents rated the quality of written reviews and proposals more highly than male respondents. Asian respondents rated the quality of written reviews higher but the quality of proposals lower than White, non-Hispanic respondents.

5.1.4 *Perceptions of fairness within the merit review process*

In general, perceptions of fairness were positive. Two out of three respondents perceived the process to be fair. Perceptions of fairness patterns varied by race and ethnicity and role in the merit review process. Respondents who identified as Hispanic and those who identified as Black, non-Hispanic were less likely to think the process was fair compared with White, non-

Hispanic respondents. Respondents who had only reviewed proposals were most likely to think the process was fair, and those who had only submitted proposals were least likely to think the process was fair.

5.1.5 Trend differences compared with 2019 survey

Respondent satisfaction and perception of fairness in the merit review process were both lower in 2021 compared with 2019. Respondents were significantly more likely to report declining to serve as a reviewer in 2021 than in 2019. Respondents in 2021 were also less likely to perceive proposals as high quality than in 2019. This study could not compare perceptions of quality of written reviews or burden across survey administration because of variation in survey items.

5.2 Implications

Based on the findings discussed in this report, Team RIVA identified a series of potential policy and research implications designed to enable NSF to continue to assess and improve the merit review process.

5.2.1 Facilitating applicants' exposure to or transparency around the merit review process may increase satisfaction and perceptions of fairness

Survey findings demonstrated consistent differences in the experiences of applicants who had also served as reviewers and those who had not. For example, individuals who were applicants only were less satisfied with the merit review process and perceived the process to be less fair. It is possible that exposure to the merit review process as a reviewer may give applicants a better understanding of how challenging funding decisions are made, thereby reducing negative perceptions among those whose funding requests are declined. While not part of this study, future research could examine funding and satisfaction to better understand this relationship. (See section 5.2.6 for additional information.) A supplemental area for analysis may examine if applicants who also serve as reviewers have received more funding than those who have not served as reviewers. This factor could also affect their perceptions of the merit review process.

Increasing applicant exposure to or transparency around the merit review process may increase satisfaction and perceptions of fairness. One way to achieve this may be by working to expand some of the less represented (and possibly less satisfied) subgroups among the reviewer pool. This may require further examination of the involvement of underrepresented groups (e.g., individuals with disabilities).¹⁵

5.2.2 Redesigning future survey sampling and recruitment efforts may enhance the conclusions that can be drawn by population subgroups and other relevant process factors

As noted in chapter 3.2, the response rate for the survey (27 percent) may limit the generalizability of the results of this study. The nonresponse bias analyses and construction of survey weights were applied to mitigate this issue, but respondents and nonrespondents may have differed on other, unobservable factors that may have influenced willingness to participate in the survey. Relatedly, as noted in chapter 3.3, the conclusions the study team could draw were limited by small sample sizes for some subpopulations. This was particularly problematic

¹⁵ Question 9 of the NSF *Learning Agenda FY 2022–FY 2026* examines the characteristics of proposals (including those of individual applicants) submitted to the merit review process and whether these characteristics are associated with different review and funding outcomes (NSF 2022d). A similar examination of reviewer representation may further inform this topic.

for small race and ethnicity groups, who are historically overlooked or marginalized when small subgroups are collapsed to increase cell sizes (Shotton et al. 2013).

Alternative survey design approaches could help increase response rates, which would help further mitigate bias and increase sample sizes for subgroup analyses and improve the reliability of the data. First, this biennial survey could be redesigned as a stratified sample based on priority characteristics of applicants and reviewers rather than a census, which would enable identification of required response rate thresholds to create representative sample by key characteristics. A smaller sample would also make additional nonresponse outreach strategies (e.g., telephone follow-up, text message) more feasible. This approach could be pilot tested to determine effectiveness. The result could be improvement in overall data quality and ability to conduct more robust analyses for specific population subgroups.

Second, NSF could implement an ongoing survey for collecting immediate reactions to the merit review process after submission of proposals or receipt of award decisions for applicants and after submission of reviews or participation in a panel for reviewers. These efforts could be designed to be very brief, targeting one or two key elements of interest to NSF—for example, overall satisfaction with written proposal reviews and likelihood to submit a proposal to NSF again. These data would be associated with specific proposal or review activities, facilitating more detailed analyses by NSF division or program, for example, rather than a broader retrospective of previous experiences.

5.2.3 Examination or expansion of efforts to support underrepresented groups may be warranted

Patterns of survey findings by institutional affiliation suggested respondents affiliated with non-R1/MSI may be more satisfied with the merit review process, both as applicants and reviewers, than those respondents affiliated with the more commonly funded R1/non-MSIs (NSF N.d.a). These individuals also conducted fewer reviews and submitted fewer proposals but were more likely to agree that reviews of proposals were of high quality. While individuals associated with these often underrepresented institutions may not be participating as broadly as their large research institution counterparts, their experiences appear to be more positive. Although this study is descriptive in nature and findings should be interpreted with caution, one possible explanation may be NSF's intentional efforts to support MSIs (e.g., attendance at workshops for tribal colleges and other MSIs; NSF 2020) and non-R1s (e.g., initiatives such as the Community College Innovation Challenge; NSF 2014). Evaluation of these efforts may be warranted because expansion of similar efforts may have the potential to help improve the experiences of other underrepresented groups, such as Black scientists who perceive the process to be unfair relative to White respondents.

5.2.4 Increasing awareness of available accommodations and supports for individuals with disabilities may increase satisfaction and perceptions of fairness

Survey results showed that respondents with disabilities were less satisfied with the merit review process than their counterparts who did not report having a disability and less likely to perceive the merit review process as fair. All NSF systems are required to comply with Section 508 of the Rehabilitation Act of 1973, designed to ensure access for people with physical, sensory, or cognitive disabilities (U.S. Access Board N.d.). NSF's Office of Equity and Civil Rights is also charged with removing barriers and providing reasonable accommodations to individuals with disabilities (NSF N.d.b). While this study does not examine the topic, awareness of and access to accommodations may affect the experience individuals with disabilities have

with the merit review process. To increase perceived fairness, NSF might consider dissemination efforts to increase awareness of such accommodations and supports among the scientific community. There may also be other reasonable accommodations not currently offered that individuals with disabilities need. Efforts to collect data from applicants and reviewers with disabilities would help NSF better understand challenges and limitations of the current process and develop supports to provide reasonable accommodation when needed.

5.2.5 Establishment of ongoing assessment methods can help determine how satisfaction with and perceptions of the merit review process change over time

Survey results showed that respondents were generally less satisfied with the merit review process in 2021 compared with 2019. Perceptions of review quality and fairness of the process were also lower in 2021. While these data provide a high-level first look at changes over time, collected data cannot offer a more detailed understanding of these declines. The period between 2019 and 2021 saw a global health pandemic and deep social unrest in the United States surrounding race and politics. The factors make it difficult to disentangle what additional factors outside NSF control may have affected these changes in the satisfaction with and perceptions of the merit review process. A more targeted data collection on how NSF procedures and policies relate to applicant and reviewer experiences may help clarify the reason behind these changes.

5.2.6 Investigate if applicant funding success may explain differences between satisfaction with and perceptions of the merit review process across subgroups

This study does not control for overall applicant success rate or the most recent proposal award outcome when examining respondent satisfaction and perceptions of burden, quality, and fairness. However, applicant success could be related to these outcomes to some degree. Relatedly, the *2019 Merit Review Report* did find that applicants whose most recent proposal was accepted (or, for preliminary proposals, resulted in an invitation to submit a full proposal) were more likely to be satisfied with the merit review process (Hare et al. 2020). Additional analyses of the 2021 Merit Review Survey data may benefit from understanding if applicant funding success explains the observed differences in subgroups, such as applicants' and reviewers' perceptions of the merit review process experience. For example, applicants with successful proposals may be more likely to be satisfied with the process, so controlling for NSF success may either partially or totally explain that association.

Acknowledgements

The authors wish to thank the researchers within the Evaluation and Assessment Capability Section of the Office of Integrative Activities at the National Science Foundation for their guidance on the 2021 Merit Review Survey study design, data collection, and analysis activities. Specifically, thank you to Erika Rissi and Cecilia Speroni for their collaboration on the study design, questionnaire development, and data analysis approach. Additional thanks to Gregory Anderson and Christina Freyman for feedback on report analyses and Christopher Monk and Robyne McRey for contract management and support. Finally, we express appreciation to the applicants and reviews who took the time to complete the survey to provide valuable data on their experiences with the merit review process.

References

- American Association for Public Opinion Research. 2016. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys." Retrieved July 31, 2017 (http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf).
- Carnegie Classification of Institutions of Higher Education. N.d. "Basic Classification Description." Retrieved August 12, 2022 (https://carnegieclassifications.acenet.edu/classification_descriptions/basic.php).
- Dhingra, Pawan. 2007. *Managing Multicultural Lives: Asian American Professionals and the Challenges of Multiple Identities*. Stanford, CA: Stanford University Press.
- Hare, Amanda, Meg Trucano, Elaine Wilcox-Cook, Sarah Hong, and Josh Rozen. 2020. "Merit Review—2019 Assessment of Principal Investigator and Reviewer Experiences." Arlington, VA: National Science Foundation.
- Kim, Nadia Y. 2007. "Critical Thoughts on Asian American Assimilation in the Whitening Literature." *Social Forces* 86(2):561–74.
- Min, Pyong Gap, ed. 2002. *The Second Generation: Ethnic Identity among Asian Americans*. Walnut Creek, CA: AltaMira Press.
- National Science Foundation. N.d.a. "Award Summary: Top 50 Institutions FY 2021." Retrieved August 4, 2022 (<https://dellweb.bfa.nsf.gov/Top50Inst2/default.asp>).
- National Science Foundation. N.d.b. "Welcome to the Office of Equity and Civil Rights." Retrieved August 9, 2022 (<https://www.nsf.gov/od/oecr/>).
- National Science Foundation. 2014. "News Release 14-123: NSF Announces the Community College Innovation Challenge." Retrieved August 4, 2022 (https://www.nsf.gov/news/news_summ.jsp?cntn_id=132689).
- National Science Foundation. 2020. "Merit Review Process: Fiscal Year 2019 Digest." Retrieved August 4, 2022 (https://nsf.gov/nsb/publications/2020/merit_review/FY-2019/nsb202038.pdf).

- National Science Foundation. 2021a. "Proposal and Award Policies and Procedures Guide (NSF 20-1)." Retrieved August 9, 2022 (https://www.nsf.gov/pubs/policydocs/pappg22_1/nsf22_1.pdf).
- National Science Foundation. 2021b. "Merit Review Process: Fiscal Year 2020 Digest." Retrieved February 14, 2023 (https://www.nsf.gov/nsb/publications/2021/merit_review/FY-2020/nsb202145.pdf).
- National Science Foundation. 2022a. "NSF by the Numbers." Retrieved August 9, 2022 (https://tableau.external.nsf.gov/views/NSFbyNumbers/Trends?%3AisGuestRedirectFromVizportal=y&%3Aembed=y&%3Alinktarget=_blank&%3Atoolbar=top).
- National Science Foundation. 2022b. "NSF FY 2023 Budget Request to Congress." Retrieved August 9, 2022 (<https://www.nsf.gov/about/budget/fy2023/pdf/fy2023budget.pdf>).
- National Science Foundation. 2022c. "Leading the World in Discovery and Innovation, STEM Talent Development and the Delivery of Benefits from Research: NSF Strategic Plan for Fiscal Years (FY) 2022-2026." Retrieved August 9, 2022 (<https://www.nsf.gov/pubs/2022/nsf22068/nsf22068.pdf>).
- National Science Foundation. 2022d. "Learning Agenda FY 2022–FY 2026." Retrieved August 9, 2022 (https://www.nsf.gov/od/oia/eac/PDFs/NSF_FY22-FY26%20Learning%20Agenda%20Final.pdf).
- Office of Management and Budget. 1997. "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity." *Federal Register* 62(210):58782-90. Retrieved August 9, 2022 (<https://www.govinfo.gov/content/pkg/FR-1997-10-30/pdf/97-28653.pdf>).
- Shotton, Heather J., Shelly C. Lowe, and Stephanie J. Waterman, eds. 2013. *Beyond the Asterisk: Understanding Native Students in Higher Education*. Sterling, VA: Stylus Publishing.
- STATA.com. N.d. "Margins – Marginal Means, Predictive Margins, and Marginal Effects." Retrieved August 4, 2022 (<https://www.stata.com/manuals/rmargins.pdf>).
- U.S. Access Board. N.d. *Text of Section 508 of the Rehabilitation Act of 1973, as Amended (29 U.S.C. §794d)*. Retrieved August 9, 2022 (<https://www.access-board.gov/law/ra.html#section-508-federal-electronic-and-information-technology>).

Appendix A. Supplemental Information on Data

A.1 Administrative Data

Team RIVA used proposal and review-level administrative data from NSF to construct the 2021 survey universe. Tables A.1.1 and A.1.2 provide descriptions of the data fields that were present in the proposal- and review-level datasets, respectively.

Table A.1.1. NSF Proposal-level administrative data variables

Variable name	Description	Data Source
revr_id	Applicant's reviewer ID (if applicable)	RPTSQL (csd.revr)
pi_id	Applicant's PI ID	RPTSQL (csd.pi)
pi_frst_name	Applicant's first name	RPTSQL (csd.pi)
pi_last_name	Applicant's last name	RPTSQL (csd.pi)
pi_ethn_code	Applicant's ethnicity	RPTSQL (csd.pi)
pi_gend_code	Applicant's gender	RPTSQL (csd.pi)
pi_degr_yr	Applicant's highest degree year	RPTSQL (csd.pi)
pi_mid_init	Applicant's middle initial	RPTSQL (csd.pi)
pi_sufx_name	Applicant's name suffix (if applicable)	RPTSQL (csd.pi)
pi_email_addr	Applicant's email address	RPTSQL (csd.pi_vw)
DemogDescrip	Description of either applicant's race or applicant's disability status	RPTSQL (csd.dmog_code_vw)
dmog_tbl_code	Indicator for whether DemogDescrip refers to applicant's race or disability status	RPTSQL (csd.pi_dmog)
DemogDate	Date demographic information was collected	RPTSQL (csd.pi_dmog)
email2	Alternative applicant email address	RPTSQL (csd.pi_addr)
Emaildate	Date the email address was last used	RPTSQL (csd.pi_addr)
TotalNumProps	Lifetime number of proposals submitted	RPTSQL (csd.prop)
Tot_prop_awd	Lifetime number of proposals awarded	RPTSQL (csd.prop)
Tot_funding_rcd	Lifetime amount of funding received	RPTSQL (flfldb.flp.awd)
prop_id	Proposal ID	RPTSQL (csd.prop)
DUNS	Institution's DUNS number	RPTSQL (csd.inst)
inst_name	Institution name	RPTSQL (csd.inst)
inst_id	Institution ID	RPTSQL (csd.prop)
MSI_flag	Indicator for minority-serving institution	RPTSQL (csd.inst)
nsf_rcvd_date	Date NSF received the submission	RPTSQL (csd.prop)
prop_titl_txt	Proposal title	RPTSQL (csd.prop)
pgm_ele_code	NSF program element code	RPTSQL (csd.prop)
perf_org_txt	Description of the applicant's institution type	RPTSQL (csd.perf_org)
DIR_DIV_ABBR	NSF directorate division abbreviation	RPTSQL (flfldb.dbo.prop)

Table A.1.2. NSF review-level administrative data variables

Variable name	Description	Data Source
revr_id	Reviewer's reviewer ID	RPTSQL (csd.revr)
pi_id	Reviewer's applicant ID (if applicable)	RPTSQL (csd.revr)
revr_frst_name	Reviewer's first name	RPTSQL (csd.revr)
revr_last_name	Reviewers last name	RPTSQL (csd.revr)
revr_ethn_code	Reviewer's ethnicity	RPTSQL (csd.revr_dmog_hdr_vw)
revr_gend_code	Reviewer's gender	RPTSQL (csd.revr_dmog_hdr_vw)
pi_degr_yr	Reviewer's highest degree year (if applicable)	RPTSQL (csd.pi)
revr_mi_name	Reviewer's middle initial	RPTSQL (csd.revr)
revr_sufx_name	Reviewer's name suffix (if applicable)	RPTSQL (csd.revr)
revr_emai_addr	Reviewer's email address	RPTSQL (csd.revr_emai_pkg)
DemogDescrip	Description of either reviewer's race or reviewer's disability status	RPTSQL (csd.dmog_code_vw)
DemogDate	Date demographic information was collected	RPTSQL (csd.pi_dmog)
DemogCode	Indicator for whether DemogDescrip refers to reviewer's race or disability status	RPTSQL (csd.pi_dmog)
emaildate	Date the email address was last used	RPTSQL (csd.revr_emai_pkg)
TotalNumRevs	Lifetime number of reviews submitted	RPTSQL (csd.rev_prop)
prop_id	Proposal ID	RPTSQL (csd.rev_prop)
DUNS	Institution's DUNS number	RPTSQL (csd.inst)
inst_name	Institution name	RPTSQL (csd.inst)
inst_id	Institution ID	RPTSQL (csd.inst)
MSI_flag	Indicator for minority-serving institution	RPTSQL (csd.inst)
rev_sent_date	Date that reviewer submitted the review	RPTSQL (csd.rev_prop)
prop_titl_txt	Proposal title	RPTSQL (csd.prop)
pgm_ele_code	NSF program element code	RPTSQL (csd.prop)
perf_org_txt	Description of the reviewer's institution type	RPTSQL (csd.inst)
Rev_Type	Type of review performed (i.e., panelist, ad hoc, or both)	RPTSQL (csd.rev_prop)
DIR_DIV_ABBR	NSF directorate division abbreviation	RPTSQL (fflfpdb.dbo.prop)

Note: PI = principal investigator

These data were combined into a single dataset. Deduplication was then performed using a series of deduplication keys, which were combinations of variables describing applicant and reviewer characteristics. Examples of deduplication keys follow:

- concatenation of the individual's full first name, middle initial, last name, institution ID, institution name, race, ethnicity, and gender;
- concatenation of the individual's first name, last name, and institution ID;
- concatenation of the individual's first name, last name, and institution's DUNS number;

- concatenation of the individual's first name, last name, and truncated proposal title; and
- individual's email, converted entirely to lowercase.

Deduplication was carried out by assigning individuals a preliminary universe ID. Then, the keys were used to identify individuals within the dataset who were likely to be the same person by identifying records with matching values for the keys. If two or more records were found to belong to the same person, the universe ID corresponding to the most recent record within the dataset was retained and the other universe ID(s) for that individual were set equal to the retained ID.

After the first round of deduplication was completed, a third-party vendor flagged any email addresses that were invalid or no longer active. The team then conducted searches of publicly available websites (e.g., professional academic profiles, LinkedIn) to identify alternative emails for individuals whose initial address was invalid or unreachable or for whom no email address was available in the NSF administrative data. After incorporating these new email addresses, another round of deduplication was conducted in case the addition of new email addresses enabled more duplicates to be identified. Once this process was complete, the deduplicated universe file contained 66,585 unique individuals.

In March 2023, four additional variables were pulled from NSF's DataLake in response to an identified issue of missing demographic data in NSF's RPTSQL database. Table A.1.3 provides descriptions of the four variables pulled from DataLake and incorporated into the analysis file. For all universe members who have served as applicants, these data were used to construct new versions of the administrative race and ethnicity, gender, and disability status variables used in nonresponse bias analysis.

Table A.1.3. NSF applicant demographic data variables pulled for demographic update

Variable name	Description	Data Source
gend_code	Applicant's gender	DataLake (pars.nsf_user_gend)
ethn_code	Applicant's ethnicity	DataLake (pars.nsf_user_ethn)
rno_code	Applicant's race	DataLake (pars.nsf_user_race)
hdcp_code	Applicant's disability status	DataLake (pars.nsf_user_hdcp)

A.2 2021 Survey Instrument

Survey Instrument

[ALL, NOT REQUIRED]

Welcome to the 2021 Merit Review Survey

Your responses will help the National Science Foundation (NSF) understand researchers' experiences with proposing to and/or reviewing proposals for our programs. NSF leadership has used past results to revise NSF proposal submission and review processes. We need your participation now to inform future decisions.

This survey focuses on your experiences as someone who has **reviewed** proposals for NSF and/or **submitted** proposals to NSF. It also has some demographic questions to help NSF understand how experiences may vary for different groups of individuals, such as for underrepresented minorities, women, and early career faculty.

Your participation is voluntary but critical to the success of NSF improvement efforts. Your decision to participate or not to participate in this survey will not adversely affect consideration of your pending or future proposals or your status as a reviewer.

This survey should take about 20 minutes to complete. Your responses will be captured and saved as they are entered. You may exit the survey at any time, and all of your previous responses will be retained. When you return, you will start back up at the point at which you left. Any responses entered may be used for analysis regardless of completion status. **All responses will be kept confidential** and only aggregate findings will be reported.

Once you complete a section, you will **not** be able to return and edit your responses. Please make sure your responses are final before completing each section and advancing to the next one.

Note: Do not use your browser "back" button while taking this survey, as it may result in the loss of your responses. Instead, use the blue navigation arrows that appear at the bottom of the survey page.

OMB Number: 3145-0257

Expiration Date: 04/30/2024

[PAPERWORK BURDEN STATEMENT WILL BE LINKED IN THE FOOTER OF EACH PAGE]

Paperwork Burden Statement

This information is collected under the authority of the National Science Foundation Act of 1950, as amended. According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number. The valid OMB control number for this information collection is 3145-0257. The time required to complete this voluntary information collection is estimated to average 20 minutes, including the time to review instructions, and complete and review responses.

[APPEARS IN FOOTER ON WEB SCREEN]

Please contact MeritReviewSurvey@nsf.gov for assistance with or questions about this survey.

[SCREEN_READER]_Are you using any technology assistance, such as a screen reader, to complete this survey?

- ☐ Yes **1**
- ☐ No **0**

[ALL, REQUIRED]

Q1A.* Since October 1, 2018, have you reviewed a proposal for NSF?

- ☐ Yes **1**
- ☐ No **0**

[ALL, REQUIRED]

Q1B.* Throughout the survey, someone who submits a proposal to NSF, a proposer, is also called a Principal Investigator (PI).

Since October 1, 2018, have you submitted a proposal to NSF as PI? (Do not include your experience as a co-PI.) For the purpose of this survey, please do not count post-doctoral fellowship applications or student fellowship applications as proposals.

- ☐ Yes **1**
- ☐ No **0**

[IF Q1A=0 AND Q1B=0, GO TO THANK YOU]

[Q1A=1 OR Q1B=1, NOT REQUIRED]

[SELECT ONE; DIVISION LIST WILL APPEAR AS A LINKED FILE UNDER THE QUESTION STEM]

Q2. Since October 1, 2018, with which NSF Directorate have your proposal submission or review activities been most often affiliated?

If your work aligns with more than one, select the directorate or office with which your activities have been most often affiliated.

Click here to view a list of divisions by directorate.

- ☐ Biological Sciences 1
- ☐ Computer & Information Science & Engineering 2
- ☐ Education & Human Resources 3
- ☐ Engineering 4
- ☐ Geosciences 5
- ☐ Mathematical & Physical Sciences 6
- ☐ Social, Behavioral & Economic Sciences 7
- ☐ Office of Integrative Activities 8

- International Science & Engineering 9
○ Other, please specify _____ 10____

[PDF FILE IN REFERENCE WINDOW]

Directorate/Office	Division
BIO = Biological Sciences (01)	DBI = Biological Infrastructure DEB = Environmental Biology EP = Emerging Frontiers IOS = Integrative Organismal Systems MCB = Molecular & Cellular Biosciences
CISE = Computer & Information Science & Engineering (02)	OAC = Office of Advanced Cyberinfrastructure (formerly, Division of Advanced Cyberinfrastructure) CNS = Computer & Networking Systems CCF = Computing & Communication Foundations CCF = Computing & Communication Foundations IIS = Information & Intelligent Systems
EHR = Education & Human Resources (03)	DGE = Graduate Education HRD = Human Resource Development DRL = Research on Learning in Formal & Informal Settings DUE = Undergraduate Education
ENG = Engineering (04)	CBET = Chemical, Bioengineering, Environmental, and Transport Systems CMMI = Civil, Mechanical & Manufacturing Innovation ECCS = Electrical, Communications & Cyber Systems EEC = Engineering Education & Centers EFMA = Merging Frontiers and Multidisciplinary Activities IIP = Industrial Innovation & Partnerships
GEO = Geosciences (05)	AGS = Atmospheric & Geospace Sciences EAR = Earth Sciences OCE = Ocean Sciences OPP = Office of Polar Programs (formerly, Division of Polar Programs)
MPS = Mathematical & Physical Sciences (06)	AST = Astronomical Sciences CHE = Chemistry DMR = Materials Research DMS = Mathematical Sciences PHY = Physics
SBE = Social, Behavioral & Economic Sciences (07)	BCS = Behavioral & Cognitive Sciences NCSES=National Center for Science and Education Statistics SES = Social & Economic Sciences
OIA = Office of Integrative Activities (08)	
OISE = Office of International Science and Engineering (09)	
Other, please specify	

[ALL, NOT REQUIRED]

[X] If you wish to modify any of your responses to this section, click the back arrow below. **After clicking the forward arrow, you will not be able to navigate back to this section of the survey.**

[IF Q1A=0 AND Q1B=1, GO TO PI_Intro]

[SECTION: EXPERIENCES AS A REVIEWER]

[Q1A=1, NOT REQUIRED]

[Reviewer_Intro] The following questions ask about your experiences reviewing NSF proposals for the **[autofill directorate] directorate** since October 2018. For these questions, please use the definitions below.

- An **ad hoc reviewer** is someone who submits a written review of a proposal but does not participate in a discussion of the proposal with other reviewers.
- A **panelist, or panel reviewer**, is someone who participates in a discussion of a proposal (usually more than one proposal) with other reviewers. A panelist may or may not prepare a written review but has access to the reviews written by others. Panelists may meet face to face or remotely.

[Q1A=1, NOT REQUIRED; IF SCREEN_READER=1 SHOW ALTERNATE ACCESSIBLE GRID FORMAT]

Q3. Since October 2018, for NSF, have you served as ...

Yes 1	No 0
----------	---------

Q3a. An ad hoc reviewer?

Q3b. A panelist/panel reviewer?

[Q1A=1, NOT REQUIRED]

Q4. Since October 1, 2018, approximately what is the total number of reviews of individual proposals that you have written for NSF, as both an ad hoc reviewer or a panelist?

[text box] Number of written reviews of individual proposals [RANGE CHECK 0 TO 100]

[NUMBER VALIDATION: "Please enter a value less than 100 and greater than zero; no decimals, please."]

[Q1A=1, REQUIRED; IF SCREEN_READER=1 SHOW ALTERNATE ACCESSIBLE GRID FORMAT]

Q5.* Since October 1, 2018, have you declined to...

Yes 1	No 0	Was not asked to participate 2
----------	---------	-----------------------------------

- Q5A. Serve as an ad hoc reviewer for NSF?
- Q5B. Serve as a face-to-face panelist on an NSF review panel?
- Q5C. Serve as a remote panelist on an NSF review panel?

[Q1A=1 AND Q5A=1, Q5B=1, OR Q5C=1, NOT REQUIRED; QUESTION WILL BE SPLIT ACROSS TWO SCREENS Q6A-E & Q6F-J; IF SCREEN_READER=1 SHOW ALTERNATE ACCESSIBLE GRID FORMAT]]

- Q6. Thinking about the most recent time you declined to participate in a review, to what extent did the following factors influence your decision?

To a Great Extent 3	To a Moderate Extent 2	To a Small Extent 1	Not at all 0
------------------------	---------------------------	------------------------	-----------------

- Q6A. Proposal or program was not related to my professional interests
- Q6B. Lack of time
- Q6C. Conflict of interest
- Q6D. Too many NSF review requests
- Q6E. Competing professional pressures (including teaching, organizational administration service, etc.)
- Q6F. Dissatisfaction with the proposal review process
- Q6G. Commitments as a reviewer to other funding agencies
- Q6H. [SHOW IF Q5B=1] Unable to travel to a face-to-face panel
- Q6I. [SHOW IF Q5C=1] Dislike participating in discussions over phone, video-conference, or web-based meeting technology
- Q6J. Some other factor (Specify): *[text box]*

[Q1A=1, NOT REQUIRED]

- Q7. Thinking about the most recent time you wrote a review of an NSF proposal, approximately how many hours did it take for you to read the proposal, write, and submit that individual review? Please do not count time spent travelling to or sitting in panels.

[text box] hours [NUMBER VALIDATION: "Please enter a value less than 1,000 and greater than zero; no decimals, please."]

[Q1A=1, NOT REQUIRED]

- Q8. When do you typically read proposals and write reviews of NSF proposals?

- ☐ During normal working hours. 1
- ☐ Outside of normal working hours. 2
- ☐ Both during and outside normal working hours equally. 3

[Q1A=1, NOT REQUIRED]

Q9. Does your employer view your participation as a reviewer (for NSF or other funding agencies) to be within the scope of your normal work duties, or outside the scope of your normal work duties?

- ☐ Within the scope of my normal work duties. **1**
- ☐ Outside the scope of my normal work duties. **2**
- ☐ I don't know. **3**

[Q1A=1, NOT REQUIRED]

[OverallPropQual_Intro] The following questions will ask you about your perceptions about the quality of the proposals you have reviewed.

[Q1A=1, NOT REQUIRED; IF SCREEN_READER=1 SHOW ALTERNATE ACCESSIBLE GRID FORMAT]

Q10. Based on your experience reviewing proposals for NSF, to what extent do you agree or disagree with the following statements?

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1	2	3	4	5

Q10A. Overall, the majority of proposals I have reviewed in recent years have been of high quality.

Q10B. Individuals submitting proposals are treated fairly.

[Q1A=1, NOT REQUIRED; IF SCREEN_READER=1 SHOW ALTERNATE ACCESSIBLE GRID FORMAT]

Q11. To what extent has participating as an NSF reviewer...

To a Great Extent	To a Moderate Extent	To a Small Extent	Not at All	Not Applicable
3	2	1	0	4

Q11A. Improved your understanding of the proposal process?

Q11B. Provided useful information for improving your next proposal?

Q11C. Influenced you to submit to another funding agency?

Q11D. Discouraged you from submitting your proposals to NSF?

[X] If you wish to modify any of your responses to this section, click the back arrow below. **After clicking the forward arrow, you will not be able to navigate back to this section of the survey.**

[IF Q1A=1 AND Q1B=0, GO TO Q21]

[SECTION: EXPERIENCES AS A PRINCIPAL INVESTIGATOR]

[Q1B=1, NOT REQUIRED]

[PI_Intro] NSF is interested in your experience seeking funding from NSF and other sources. Please answer the following questions based on your experience as a principal investigator (PI), not on any experience that you may have had as a co-PI. **Please think only of the proposals you have submitted to NSF since October 1, 2018.**

[Q1B=1, REQUIRED]

Q12.* Since October 1, 2018, how many proposals have you submitted to NSF's [autofill directorate] directorate?

[textbox] Number of proposals [RANGE CHECK 0 TO 99]

[NUMBER VALIDATION: "Please enter a value greater than zero; no decimals, please."]

[Q1B=1, NOT REQUIRED]

Q13. Thinking about the most recent proposal you submitted to NSF, how many hours did you spend preparing (writing, formatting, and submitting) the proposal?

[textbox] hours

[Q1B=1, REQUIRED]

Q14.* Since October 1, 2018, have you applied for funding (i.e., proposals or contracts) from a federal agency other than NSF?

- ☐ Yes 1
- ☐ No 0 [GO TO Q16]

[NUMBER VALIDATION: "Please enter a value greater than zero; no decimals, please."]

[Q1B=1, NOT REQUIRED]

Q15. Compared to other federal agencies' proposal submission systems, how much effort does it take to write and complete a proposal in the required format and submit it to NSF?

- ☐ More Effort 3
- ☐ Nearly the Same Effort 2
- ☐ Less Effort 1

[Q1B=1, REQUIRED]

Q16.* Since October 1, 2018, have you received a funding decision for any proposals you submitted to NSF?

- ☐ Yes **1**
☐ No **0**

[Q1B=1, REQUIRED]

Q17.* Have you ever submitted a proposal to NSF that was declined?

- ☐ Yes **1**
☐ No **0** [GO TO Q19]

[Q1B=1 AND Q17=1, NOT REQUIRED; IF SCREEN_READER=1 SHOW ALTERNATE ACCESSIBLE GRID FORMAT]

Q18. To what extent did the written reviews that accompanied the MOST RECENT decline of one of your NSF proposals...

To a Great Extent	To a Moderate Extent	To a Small Extent	Not at All
3	2	1	0

Q18A. Improve your understanding of the proposal process?

Q18B. Provide useful information for revising and improving your next proposal?

Q18C. Influence you to submit to another funding agency?

Q18D. Discourage you from revising and submitting your proposals to NSF?

[Q1B=1 AND Q16=1, NOT REQUIRED; IF SCREEN_READER=1 SHOW ALTERNATE ACCESSIBLE GRID FORMAT]

Q19. Thinking of the MOST RECENT proposal you submitted to NSF, how satisfied or dissatisfied were you with...

Very Dissatisfied	Somewhat Dissatisfied	Neither Dissatisfied nor Satisfied	Somewhat Satisfied	Very Satisfied	Not Applicable
1	2	3	4	5	0

Q19A. The quality of the information NSF provided during the proposal submission process (i.e., FastLane, FAQs, website content)

Q19B. The timeliness of the decision to award or decline funding

Q19C. Your interaction with NSF staff

[Q1B=1, NOT REQUIRED; SPLIT ACROSS TWO SCREENS-Q19A-E & Q19F-H; IF SCREEN_READER=1 SHOW ALTERNATE ACCESSIBLE GRID FORMAT]

Q20. Based on your experience submitting proposals to NSF, to what extent do you agree or disagree with the following statements?

Strongly Agree 5	Agree 4	Neither Agree nor Disagree 3	Disagree 2	Strongly Disagree 1	Not Applicable 0
---------------------	------------	---------------------------------	---------------	------------------------	---------------------

Q20A. Written reviews are thorough.

Q20B. Written reviews are technically sound.

Q20C. Written reviews are of high quality.

Q20D. The panel summary or summaries are of high quality.

Q20E. The information provided regarding the outcomes of the competition is of high quality.

Q20F. The program officer comments I view in FastLane help me understand the decision to decline or award my proposal.

Q20G. The conversations (email, phone, face-to-face) I have with my program officer provide me with helpful feedback about my proposal.

Q20H. Individuals submitting proposals are treated fairly.

[SECTION: OVERALL EXPERIENCES]

[ALL, NOT REQUIRED; IF SCREEN_READER=1 SHOW ALTERNATE ACCESSIBLE GRID FORMAT]

Q21. Please indicate whether you agree or disagree with the following statements:

Strongly Agree 5	Agree 4	Neither Agree nor Disagree 3	Disagree 2	Strongly Disagree 1
---------------------	------------	---------------------------------	---------------	------------------------

Q21A. I think NSF's merit review process is fair.

Q21B. I think NSF's merit review process is effective.

Q21C. [SHOW IF Q1A=1] I intend to continue to review proposals for NSF in the future.

Q21D. [SHOW IF Q1B=1] I intend to continue to submit proposals to NSF in the future.

Q21E. I am satisfied with NSF's merit review process.

[ALL, NOT REQUIRED]

Q22. This survey has asked about your experiences with NSF's merit review process. In your opinion, improving which **one** of the following factors will have the most significant effect on improving the merit review process?

- ☐ Timeliness of decisions about, and responsiveness to, proposals by NSF staff **1**
- ☐ Quality of feedback to PIs in the form of comments in written reviews **2**
- ☐ Quality of feedback to PIs in the form of comments in panel summaries **3**

- ☐ Quality of PI conversations with, and written comments from, program officers 4
- ☐ Quality of information available during proposal submission 5
- ☐ Quality of the review process from the perspective of a reviewer 6

[ALL, NOT REQUIRED]

[ALL, NOT REQUIRED]

[SECTION: EDUCATION HISTORY & EMPLOYMENT]

The following questions will prompt you to provide basic information on your career experience and your institution/organization. (Data will be reported at an aggregate level and are requested to help us understand the experiences of different groups.)

[ALL, NOT REQUIRED]

Q23. In what year did you receive your highest degree? (Please do not count honorary degrees.)

[text box] YYYY [RANGE CHECK 1950 TO 2021]

[ALL, NOT REQUIRED]

Q24. What type of degree did you earn for your highest degree?

- ☐ Bachelor's degree (e.g., BS, BA, AB) 1
- ☐ Master's degree (e.g., MS, MA, MBA) 2
- ☐ Doctorate (e.g., PhD, DSc, EdD) 3
- ☐ Other professional degree (e.g., JD, LLB, MD, DDS, DVM) - please specify 4 [... Text box ...]
- ☐ Other degree - please specify 5 [... Text box ...]

[ALL, NOT REQUIRED]

Q25. Did you receive any financial support (e.g., research assistantship, fellowship, traineeship, scholarship, other grants) from NSF as an undergraduate or graduate student?

- ☐ Yes 1
- ☐ No 0 [GO TO Q27]
- ☐ Don't know 2 [GO TO Q27]

[Q25=1, NOT REQUIRED]

Q26. What type of financial support did you receive from NSF while you were an undergraduate or graduate student?

Please select all that apply.

- ☐ REU (research experience for undergraduates) support 1
- ☐ Research assistantship 2
- ☐ Fellowship support 3
- ☐ Traineeship support 4
- ☐ Scholarship 5
- ☐ Travel grant 6
- ☐ Other (please specify): [... Text box ...] 7

[ALL, REQUIRED]

Q27.* Do you work for an institution of higher education?

If you work for multiple organizations, please pick the one you consider to be your primary employer and answer in terms of that organization.

- ☐ Yes 1
- ☐ No 0 [GO TO Q29]

[Q27=1, NOT REQUIRED]

Q28. What is your position? ■■

- ☐ Post-doctoral fellow 1
- ☐ Assistant Professor 2
- ☐ Associate Professor 3
- ☐ Full Professor 4
- ☐ Research Faculty or Scientist 5
- ☐ Adjunct Professor 6
- ☐ Emeritus/Emerita Professor 7
- ☐ Retired 8
- ☐ Other (please specify): 9 [... Text box ...]

[IF Q27=1, GO TO Demographics_Info]

[Q27=0, REQUIRED]

Q29.* Which of the following best describes your organization?

- ☐ Primary or secondary academic institution 1
- ☐ Non-profit organization 2
- ☐ For-profit organization 3
- ☐ Federally-funded R&D center 4
- ☐ Government (local, state, federal, or tribal) 5
- ☐ Other (please specify) 6

[SECTION: DEMOGRAPHICS]

[ALL, NOT REQUIRED]

The next questions request demographic information. NSF will only use these data to generate statistics to inform whether our programs and other opportunities in science and technology reach and benefit a diverse set of researchers. 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.

[ALL, REQUIRED]

Q30.* Are you a veteran?

A veteran is a person who served in the active military, naval, or air service and who was discharged or released under conditions other than dishonorable.

- ☐ Yes, I am a veteran. **1**
- ☐ No, I am not a veteran. **2**
- ☐ I do not wish to provide this information. **3**

[ALL, REQUIRED]

Q31.* What sex were you assigned at birth on your original birth certificate?

- ☐ Male **1**
- ☐ Female **2**
- ☐ I do not wish to provide this information **3**

[ALL, REQUIRED]

Q32.* What gender do you identify with?

- ☐ Male **1**
- ☐ Female **2**
- ☐ Other (please specify _____) **3**
- ☐ I do not wish to provide this information **4**

[ALL, REQUIRED]

Q33.* Are you Hispanic or Latino?

A Hispanic or Latino is a person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race.

- ☐ Yes, I am Hispanic or Latino. **1**
- ☐ No, I am not Hispanic or Latino. **2**
- ☐ I do not wish to provide this information. **3**

[ALL, REQUIRED]

Q34.* What is your race?

Please select all that apply.

Click here for race category definitions.

[RACE CATEGORY DEFINITIONS WILL APPEAR IN POP-UP INFORMATION WINDOW] ⓘ

American Indian or Alaska Native - A person having origins in any of the original peoples of North and South America (including Central America) and who maintains tribal affiliation or community attachment

Asian - A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam

Black or African American - A person having origins in any of the Black racial groups of Africa

Native Hawaiian or Other Pacific Islander - A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands

White - A person having origins in any of the original peoples of Europe, the Middle East, or North Africa

- ☐ 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) 6 [text box]
- ☐ I do not wish to provide this information 7 [exclusive answer]

[ALL, REQUIRED]

Q35.* 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 serious disability related to a physical, mental, or emotional condition
- ☐ Yes (1)
 - ☐ No 2
 - ☐ I do not wish to provide this information 3

Q36. Please enter any additional comments you may have about NSF's merit review process in the space below: _____ [Essay text box]

[ALL, NOT REQUIRED]

You have reached the end of the survey.

Thank you for participation!

A.3 2021 Survey Data Collection Methods

The 2021 Merit Review Survey was conducted as a census of applicants and reviewers who had participated in NSF's merit review process. Survey invitations were sent to approximately 66,000 individuals included in the final universe file who have submitted a proposal and/or served as a reviewer between fiscal year (FY) 2019 and FY 2020. Data collection began September 9, 2021, and lasted 11 weeks. Each universe member was sent up to five emails:

- **Prenotification email from NSF:** Prior to the survey's distribution, NSF sent a prenotification email from its servers to the intended survey recipients as identified by the universe file. The email introduced Team RIVA and the survey effort and notified recipients they should expect an invitation to participate in the survey within the week.
- **Survey invitation:** Universe members received an email invitation from Qualtrics containing a unique link to the web survey.
- **Nonresponse follow-up email:** A follow-up email was sent to nonrespondents approximately 1 week after the initial survey invitation. Emails were only sent to individuals who had not yet begun the survey (nonrespondents) and individuals who began the survey but had not finished (breakoffs).
- **NSF reminder email:** NSF sent an email using a mail merge process to all individuals who were invited to participate in the survey, reminding those who have not yet finished the survey to do so and thanking those who already completed it.
- **Last-chance email:** A final reminder email was sent to nonrespondents 1 week before the end of the data collection period.

The data collection period included an adaptive design and an experiment to test the effectiveness of shortening the outreach emails. To do this, the total universe of respondents was randomly split into two waves. Wave 1 contained a subset of about 10,000 universe members randomly assigned to an experiment or control group. The subset of members who received the longer version of the initial and reminder emails (control) consisted of 4,980 individuals. The initial invitation email had 344 words and included examples of past study results, and the follow-up email had 212 words and included example survey topics and named contractors. The treatment group, which consisted of 4,978 individuals, received shorter emails. Both the initial and follow-up emails had 140 words and omitted past study results, examples of survey topics, and contractor information. All other variables (email personalization, inclusion of survey hyperlink, and timing) were held constant.

An analysis conducted after the first 4 weeks of data collection found that completion rates were significantly higher for those who had received the treatment email (18 percent) than those who received the control email (16 percent) after the first reminder. The 2 percent significant difference was sustained at the end of data collection after the remaining nonrespondent emails were sent to wave 1 but was significant at the 0.075 rather than the 0.05 level. The results of the experiment were used to recommend that the shorter (140-word) versions be used for all wave 2 outreach.

Secondary emails were also found for email addresses that were undeliverable or did not yield responses. Invitations were sent to these emails through the nonresponse follow-up process.

The survey closed November 29, 2021, for wave 1 and wave 2.

See table A.3.1 for the full outreach schedule.

Table A.3.1. Data collection outreach timeline

Outreach type	Recruitment email	Date
NSF listserv	Wave 1 pre-invitation email	9/8/21
Survey launch	--	9/9/21
Wave 1 outreach	Wave 1 initial survey invitation (experiment and control groups)	9/9/21
Wave 1 outreach	Wave 1 first nonresponse follow-up	9/22/21
NSF listserv	Wave 1 NSF reminder email	10/6/21
Wave 1 outreach	Wave 1 last chance email	10/13/21
NSF listserv	Wave 2 pre-invitation email	10/18/21
Wave 2 outreach	Wave 2 initial survey invitation	10/19/21
Wave 2 outreach	Secondary email survey invitation	10/26/21
Wave 2 outreach	Wave 2 first nonresponse follow-up	11/2/21
NSF listserv	Wave 2 NSF reminder email	11/16/21
Wave 2 outreach	Wave 2 last chance email	11/22/21
Survey close	--	11/29/21

Appendix B. Supplemental Information on Analytic Approach

B.1 2021 Survey Data Cleaning and Variable Construction

Prior to conducting analyses of the survey data, Team RIVA performed data cleaning activities, including designation of partial survey completion, identification of ineligible respondents, and analytic variable construction.

After closing the 2021 Merit Review Survey, the team cleaned the survey datafile and prepared data for analyses. Data processing procedures to ensure high-quality data included checking ranges to verify that each field contained only allowable codes, checking skip patterns, and checking numeric sums to verify that numeric responses made up the sum of component data items.

The team used data produced from the survey software Qualtrics to determine a final survey disposition status for each individual. Any records the software identified as “Finished Survey”^{B-1} the team considered to be completed surveys. From those completed surveys, respondents who indicated in the survey that they had served as either a reviewer or an applicant (or both) during FY 2019 or FY 2020 were considered to be eligible, and those who indicated they had served as neither a reviewer nor an applicant in the same timeframe were considered ineligible. Those with a designation of “Partially Completed Survey” who also indicated they were eligible for the survey and who completed all of the survey items that composed the satisfaction indices were considered eligible partial completes. The team considered those who had a status of “Email Sent” to be nonrespondents and those with a status of “Opted Out” to be refusals. Lastly, those with statuses of “Email Bounced” or “Email Failed” were considered no contacts.

Addressing Small Cell Sizes

Categorical survey items used to identify reviewer and applicant demographic information and other characteristics were translated into categorical variables. The team retained as many of the original classes within each independent variable as possible, though the following classes were collapsed together to ensure adequate cell sizes for regression modelling:

- The independent variable for race and ethnicity contained a class labeled “other,” which included survey respondents who identified as American Indian or Alaska Native, Native Hawaiian or Pacific Islander, and those who selected “other” in the survey. This variable also contained a class labeled “unknown,” which consisted of individuals who did not respond to the race and ethnicity survey items and those who responded “do not wish to provide.”
- The variable for institutional affiliation contained a class labeled “other,” which grouped respondents who work for institutions of higher education with unknown MSI status (both R1 and non-R1 institutions), federally funded research and development centers, primary or secondary academic institutions, and “other” types of institutions.
- Within the variable for primary directorate, the Office of Integrative Activities (OIA) and the Office of International Science and Engineering (OISE) were grouped to form the Office of the Director (OD) class.
- Within the variable for gender, respondents who selected “other” and “do not wish to provide” and individuals with unknown gender were grouped together.

^{B-1} “Finished Survey” is assigned when the respondent hits the “next” button on the final question available to them in the survey.

Variable Construction

The variable for applicant and reviewer institutional affiliation was constructed from a combination of survey and administrative data. In the cleaning of the administrative data, the team used DUNS numbers, institutions names, and institution IDs, in conjunction with the Integrated Postsecondary Education Data System data and Carnegie classifications, to assign all institutions of higher education to one of six groups:

- R1/MSI
- R1/non-MSI
- R1/unknown MSI status
- Non-R1/MSI
- Non-R1/non-MSI
- Non-R1/unknown MSI status

R1 refers to doctoral universities with very high research activity, while non-R1 refers to any institutions of higher education that do not meet this description. MSI is an acronym for “minority-serving institution” and refers to historically Black colleges and universities, tribal colleges and universities, or other colleges or universities in which a specific minority group constitutes 25 percent of the total undergraduate enrollment or in which minority students as a whole constitute at least 50 percent of the total undergraduate enrollment.

If survey respondents indicated they worked for an institution of higher education, the team concatenated “institution of high education” with their respective institutional affiliation variable from the administrative data. If survey respondents indicated they did not work for an institution of higher education, the team did not concatenate with administrative data but used solely the institution type respondents indicated in the survey (federally funded research and development center, nonprofit organization, for-profit organization, government, K–12 education institution, or other).

Per NSF guidance, the team constructed the variable for early career status by establishing the threshold for early career status as anyone who received their highest degree within the past 10 years (meaning, anyone whose highest degree was conferred in 2012 or later). However, NSF’s own guidance and definition regarding early career status have evolved. In the 2019 survey, the benchmark for early career was 7 years (meaning 2013 or later). The 2019 early career status indicator was not recalculated to account for this change to keep the construction of this variable true to its definition at the time of the 2019 survey. This is one limitation in the year-over-year models that must be considered when interpreting these results.

The NSF experience variable took on two or three total values depending on the model the variable is used in. The variable had three values overall: experience serving as a reviewer only, experience serving as an applicant only, and experience serving as both a reviewer and an applicant. However, most models in the analysis only applied to either those with applicant experience or those with reviewer experience because survey skip logic dictated that only those with relevant experience were asked certain questions about their experience. As a result, in many models, this variable either only took on values for “applicant only” and “both applicant and reviewer” or “reviewer only” and “both applicant and reviewer.”

Satisfaction indices were constructed to more holistically capture applicant and reviewer satisfaction with the merit review process. First, a factor analysis was performed to identify the

key survey items that constituted this construct. Separate factor analyses were performed for applicants and reviewers. Next, the index scores were calculated using the following formula:

$$S_i = \frac{100 \left(\sum_{q=1}^Q f_q X_{qi} / m_q \right)}{\sum_{q=1}^Q f_q}$$

—where S_i is the satisfaction index score for respondent i ; Q is the number of items in the satisfaction index; f_q is the factor loading for question q ; X_{qi} is the value of question q for respondent i ; and m_q is the maximum response value for question q .

Factor loadings are included in the formula as a weighting mechanism to ensure survey items with stronger correlations to satisfaction contribute more to the indices than those with weaker correlations. The indices were also scaled to fall within a range of 0–100 for ease of interpretation.

Lastly, the team also performed a log transformation on survey items Q4, Q7, Q12, and Q13. Q4 and Q12 correspond to the number of reviews or proposals, respectively, submitted to NSF since October 1, 2018, while Q7 and Q13 correspond to the amount of time (in hours) required to compose the most recent review or proposal, respectively. The team performed a log transformation on these variables prior to specifying them as dependent variables in models because each of the distributions of these variables was highly skewed to the right. The log transformations normalized these distributions.

B.2 Nonresponse Bias Analysis

Nonresponse bias was analyzed in three ways. First, the team quantified bias by performing an analysis of covariance between survey response and item values using key outcome variables. To do this, data from regression models were used to predict values of survey item y for nonrespondents to the Merit Review Survey. The covariance between y (a random variable representing the value of the survey item for individuals in the population) and p (the probability of response) was estimated. Covariances far from zero may indicate nonresponse bias because they indicate strong relationships between a survey item response and the probability of response. This analysis enabled the estimation of nonresponse bias for each key outcome variable.

Second, the team compared response patterns for sample members responding at three different stages of the 2021 data collection process:

- early responders, who responded during approximately the first 2 weeks of fielding;
- respondents captured during nonresponse follow-up, approximately a month-long window beginning after the early responder window ended; and
- late responders, who responded after nonresponse follow-up ended, approximately six weeks after fielding began.

Early responders, those captured during nonresponse follow-up, and late responders may have different response patterns, with late responders potentially resembling nonrespondents more

than early responders. The goal of this analysis was to determine whether survey estimates were associated with response stage (i.e., if estimates from late responders were different than those from early responders). If so, this would suggest the possible presence of nonresponse bias. The team calculated confidence intervals for respondents in each of the three response stages for several key survey variables.

Third, relationships of survey items to population characteristics^{B-2} and of response propensity to population characteristics were examined using a regression analysis. If survey responses and response rates differ by individuals' characteristics, a potential for nonresponse bias exists. The goal of the regression analysis was twofold: to look for signs of nonresponse bias and to determine which characteristics to use in nonresponse weighting adjustment to mitigate the potential for nonresponse bias.

The covariance analysis produced estimated biases that were predominantly 1 percent or lower of the mean survey item value across role types and survey items. Table B.2.1 shows the estimated biases as a percentage of the mean survey item value for each key survey outcome. The analysis indicated the potential for nonresponse bias was generally low.

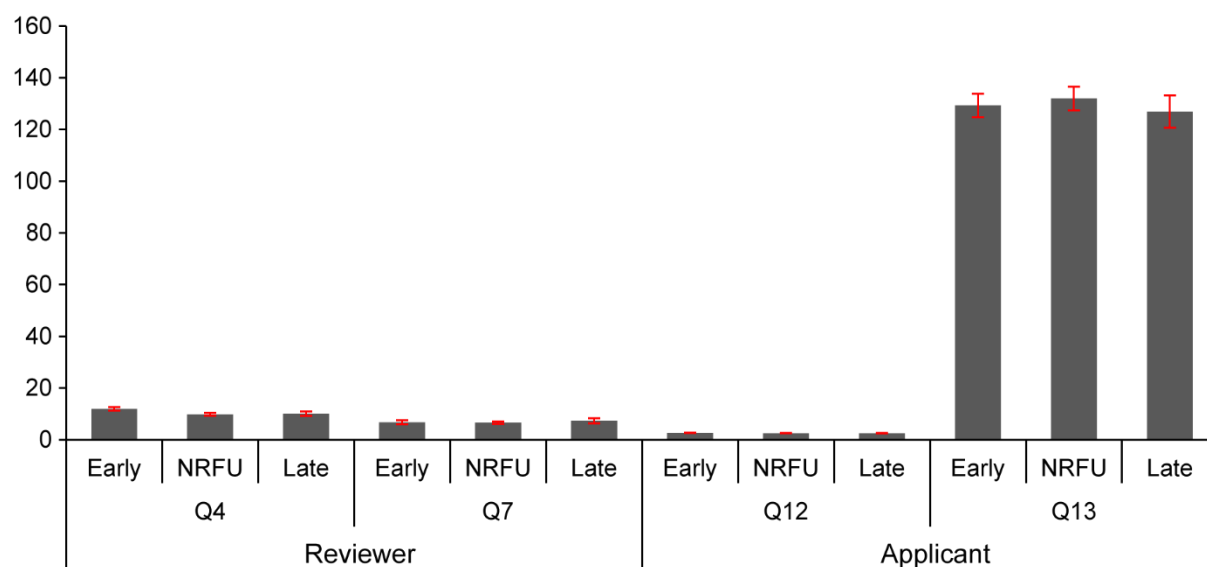
Table B.2.1. Covariance of key survey outcomes

Survey item	Estimated bias
Q4	1.0%
Q7	-1.3%
Q10A	0.3%
Q10B	0.3%
Q12	-0.4%
Q13	-1.2%
Q20A	0.8%
Q20B	0.7%
Q20C	0.8%
Q20D	0.8%
Q21A	0.7%
Q21B	0.9%
Q21E	0.9%

The results of the stage-of-response analysis showed little statistical evidence of variation in response by stage of response. Figure B.2.1 shows the mean and confidence intervals for survey questions with continuous responses by response stage. Mean item values by stage of response were statistically different for a minority of the survey items. Confidence intervals for the mean item values generally overlapped across the three stages.

^{B-2} These population characteristic variables were constructed using NSF administrative data. See appendix A.1 for additional details.

Figure B.2.1. Mean and confidence interval for continuous survey items, by response stage



The population characteristics that were statistically significant in the regressions were also inconsistent across items. Table B.2.2 shows the distribution of respondent and nonrespondent characteristics, and table B.2.3 shows which respondent characteristics were significantly associated with key survey item values. A number of relationships between survey items and population characteristics were statistically significant in the survey item regressions; fewer were significant in the response propensity analysis, but the results shown in tables B.2.3 and B.2.4 helped guide the selection of variables to use for weighting. A similar result was found in the regressions of response propensity on population characteristics when examining all three analyses (covariance, stage-of-response, and regressions). Population characteristics were rarely associated with response propensity among reviewers.

Table B.2.2. Distribution of respondent and nonrespondent characteristics based on NSF administrative data

Universe characteristics (n = 66,585)	Percentage of respondents (n = 17330)	Percentage of nonrespondents (n = 49255)
Race and ethnicity		
Hispanic, any race	6.0%	4.8%
Asian, non-Hispanic	18.0%	18.0%
Black or African American, non-Hispanic	3.3%	3.0%
White, non-Hispanic	58.2%	52.3%
Multiple races, non-Hispanic	1.2%	1.2%
Other, NHPI, AIAN	0.6%	0.5%
Native Hawaiian or Other Pacific Islander, non-Hispanic (NHPI)	0.0%	0.1%
American Indian or Alaska Native, non-Hispanic (AIAN)	0.2%	0.2%
Other race, non-Hispanic (other)	0.4%	0.3%
Unknown or do not wish to provide	12.7%	20.2%

Universe characteristics (n = 66,585)	Percentage of respondents (n = 17330)	Percentage of nonrespondents (n = 49255)
Gender		
Male	56.3%	55.1%
Female	29.7%	25.4%
Other, do not wish to provide, or unknown	14.0%	19.5%
Disability status		
Yes—hearing, mobility, vision, or other disability	1.7%	1.4%
No	89.8%	83.8%
Do not wish to provide	8.5%	14.8%
NSF directorate		
BIO	11.5%	10.6%
CISE	10.7%	13.5%
EHR	13.8%	11.0%
ENG	20.0%	22.4%
GEO	11.6%	10.1%
MPS	20.4%	18.5%
SBE	10.7%	12.3%
OD (OIA and OISE)	1.5%	1.6%
Institution type		
R1/MSI	5.4%	5.4%
R1/non-MSI	38.9%	40.3%
R1/unknown	0.3%	0.3%
Non-R1/MSI	7.5%	6.8%
Non-R1/non-MSI	43.8%	44.0%
Non-R1/unknown	4.1%	3.2%
Early career (degree less than 10 years ago)		
Yes	9.6%	12.2%
No	90.4%	87.8%
Experience		
Applicant only	60.4%	67.7%
Reviewer only	22.1%	19.7%
Both applicant and reviewer	17.5%	12.7%

Table B.2.3. Respondent characteristics significantly associated with survey item values

Characteristic	Q4	Q7	Q10A	Q10B	Q12	Q13	Q20A	Q20B	Q20C	Q20D	Q21A	Q21B	Q21E
Institution type	A/R	A	A	A	A/R	A/R	R	R	R	A/R	R	A/R	A/R
Race and ethnicity	R	A/R	A/R	A/R	A	A/R					A/R	A	A/R
Disability status	A/R	A	A	A/R	A		A	A	A	A	A	A	A/R
Early career status	A	A/R	A	A		A	A	A	A	A	A	A	A
Gender	R	A	A	A/R	A	A	A	A	A	A	A/R	A/R	A/R
Primary directorate	A/R	A/R	A/R	A/R	A/R	A/R	A/R	A/R	A/R	A/R	A/R	A/R	A/R

A = Characteristic significantly predicts survey item value among applicants, $p < 0.05$.

R = Characteristic significantly predicts survey item value among reviewers, $p < 0.05$.

A/R = Characteristic significantly predicts survey item value among applicants and reviewers, $p < 0.05$.

Table B.2.4. Population characteristics significantly associated with response propensity

Characteristic	Q4	Q7	Q10A	Q10B	Q12	Q13	Q20A	Q20B	Q20C	Q20D	Q21A	Q21B	Q21E
Institution type			A	A									
Race and ethnicity											A/R	A/R	A/R
Disability status											A	A	A
Early career status	A		A	A		A	A	A	A	A	A	A	A
Gender		A	A	A	A	A	A	A	A	A	A/R	A/R	A/R
Primary directorate	A	A	A	A	A	A	A	A	A	A	A/R	A/R	A/R

A = characteristic significantly predicts response propensity among applicants, $p < 0.05$.

A/R = characteristic significantly predicts response propensity among applicants and reviewers, $p < 0.05$.

Overall, these results indicated that the potential for nonresponse bias was low. Nonetheless, these analyses provided direction in choosing the characteristics used in the nonresponse weighting adjustment to help mitigate the potential for nonresponse bias in the 2021 survey. The population characteristics that had a greater tendency to yield statistically significant results in the regressions (directorate, gender, and race and ethnicity) were the ones chosen for the nonresponse weighting adjustment. The population characteristics that had a lesser tendency to yield statistically significant results (institution type, early career, and disability) were not used in the nonresponse weighting adjustment.

Table B.2.5 indicates which variables were used in each of the nonresponse bias analyses.

Table B.2.5. Variables used in nonresponse bias analysis

Survey item	Variable description	Used in covariance analysis	Used in response stage analysis	Used in regression analysis
Q4	Number of reviews written for NSF since October 1, 2018	•	•	•
Q7	Hours spent on most recent review	•	•	•
Q10A	Proposals reviewed have been of high quality	•		
Q10B	Individuals submitting proposals are treated fairly	•		•
Q12	Number of proposals submitted to directorate since October 1, 2018	•	•	•
Q13	Hours spent preparing most recent proposal	•	•	•
Q20A	Written reviews are thorough	•		•
Q20B	Written reviews are technically sound	•		•
Q20C	Written reviews are of high quality	•		•
Q20D	Panel summaries are of high quality	•		•
Q21A	NSF's merit review process is fair	•		•
Q21B	NSF's merit review process is effective	•		•
Q21E	Satisfaction with NSF's merit review process	•		•

B.3 Correcting for Nonresponse Bias With Cell Weighting

Using the results of the nonresponse bias analyses described in appendix section B.2, the team formulated an appropriate nonresponse weighting approach to mitigate the potential for bias in the 2021 Merit Review Survey. Regression analyses were used to determine which variables were statistically significant predictors of response. Using those findings as a guideline, directorate/office, gender, and race and ethnicity were selected as characteristics to use when creating weighting cells. To choose the weighting cells within each of these characteristics, the team explored several options based on cell sample sizes and similarity of responses from individuals across cells. The team chose a weighting scheme that provided for the calculation of a stable set of weights that summed to the population size and enabled respondents to represent nonrespondents based on characteristics related to response propensity and variation in survey item responses. The selected weighting scheme employed the following decisions to create subgroups (weighting classes) based on the value of each population characteristic:

- Each directorate was a separate weighting class for respondents with experience as applicants only or experience as both applicants and reviewers. For respondents with experience as reviewers only, the EHR and OD directorates were combined because of small sample size.
- Males were a weighting class separate from females and those with unknown gender.

- Four weighting classes were created based on race and ethnicity groups: White, Asian, unknown or do not wish to provide, and all other race and ethnicity groups.

B.4 Multiple Regression Equations

The regression models used in the 2021 survey data analyses examined contemporaneous group differences by analyzing the association between the dependent variables from each research dimension (burden, quality, fairness, and satisfaction) and the seven independent variables: race and ethnicity, gender, disability status, NSF directorate/office, institution type, early career status,^{B-3} and NSF experience.^{B-4} Table B.4.1 lists the survey items (dependent variables) and regression model types for each research dimension. Each regression model is then described in greater detail, including equations.

Table B.4.1. Dependent variables and regression model types, by research dimension

2021 survey item number	Description of survey item/dependent variable	Type of variable	Regression model type
Satisfaction dimension			
Q10A-B and Q20A-D	Reviewer satisfaction index	Continuous	Linear
Q18A-C, Q19A-H, Q20A-C, and Q20E	Applicant satisfaction index	Continuous	Linear
Q21E	Overall satisfaction with merit review process	Ordinal	Ordinal logistic
Burden dimension			
Q4	Number of reviews conducted for NSF	Continuous	Linear
Q5 A-C	If reviewer declined to serve as an ad hoc reviewer, face-to-face panel, or remote panelist for NSF	Binary	Logistic
Q6B	Extent to which lack of time influenced decision to decline most recent request to review	Ordinal	Ordinal logistic
Q6D	Extent to which too many NSF review requests influenced decision to decline most recent request to review	Ordinal	Ordinal logistic
Q6E	Extent to which competing professional pressures influenced decision to decline most recent request to review	Ordinal	Ordinal logistic
Q6G	Extent to which commitments as a reviewer to other federal agencies influenced decision to decline most recent request to review	Ordinal	Ordinal logistic
Q6H	Extent to which inability to travel to a face-to-face panel influenced decision to decline most recent request to review	Ordinal	Ordinal logistic
Q7	Amount of time to read the proposal, write, and submit the most recent review	Continuous	Linear
Q8	Time of day reviewers typically read proposals and write reviews	Nominal	Multinomial logistic
Q9	Employer's view of participation as a reviewer	Nominal	Multinomial logistic
Q12	Number of proposals submitted to NSF	Continuous	Linear

^{B-3} Early career is defined as less than 10 years since highest degree earned.

^{B-4} Experience is an indicator for individuals who are both applicants and reviewers.

2021 survey item number	Description of survey item/dependent variable	Type of variable	Regression model type
Q13	Time spent preparing the most recent proposal	Continuous	Linear
Q15	Amount of effort to write and complete a proposal compared with other agencies	Ordinal	Ordinal logistic
Quality dimension			
Q10A	Proposals reviewed have been of high quality	Ordinal	Ordinal logistic
Q18A	Extent to which reviews improved understanding of proposal process	Ordinal	Ordinal logistic
Q18B	Extent to which reviews provided useful information for improving next proposal	Ordinal	Ordinal logistic
Q18C	Extent to which reviews influenced decision to submit to another funding agency	Ordinal	Ordinal logistic
Q18D	Extent to which reviews provided discouraged from revising and submitting your proposals to NSF	Ordinal	Ordinal logistic
Q19A	Satisfaction with quality of information provided by NSF	Ordinal	Ordinal logistic
Q20A	Extent to which written reviews are thorough	Ordinal	Ordinal logistic
Q20 B	Extent to which written reviews are technically sound	Ordinal	Ordinal logistic
Q20 C	Extent to which written reviews are of high quality overall	Ordinal	Ordinal logistic
Q20 D	Extent to which panel summary or summaries are of high quality	Ordinal	Ordinal logistic
Q20 E	Extent to which information provided regarding the outcomes of the competition is of high quality	Ordinal	Ordinal logistic
Q22	Factors for improvement	Nominal	Multinomial logistic
Fairness dimension			
Q10B	Individuals submitting proposals are treated fairly	Ordinal	Ordinal logistic
Q20H	Individuals submitting proposals are treated fairly	Ordinal	Ordinal logistic
Q21A	NSF's merit review process is fair overall	Ordinal	Ordinal logistic

B.4.1. Linear regression (for continuous random variables)

The multiple linear regression model for a continuous dependent variable (e.g., number of reviews) on the seven independent variables was given by the following equation:

$$E(y_i) = \beta_0 + \beta_1(\text{race and ethnicity}_i) + \beta_2(\text{gender}_i) + \beta_3(\text{disability status}_i) + \beta_4(\text{NSF directorate/office}_i) + \beta_5(\text{institution type}_i) + \beta_6(\text{early career}_i) + \beta_7(\text{experience}_i)$$

—where

- $E(y_i)$ denotes the mean of the dependent continuous variable for individual i ;
- the subscript i indexes individuals; and
- y_i is the response of individual i .

B.4.2 Logistic regression

The multiple logistic regression model of a dichotomous dependent variable (e.g., if a reviewer declined to serve as a reviewer for NSF) on the seven independent variables was given by the following equation:

$$\ln \left[\frac{P(y_i = 1)}{1 - P(y_i = 1)} \right] = \beta_0 + \beta_1(\text{race and ethnicity}_i) + \beta_2(\text{gender}_i) + \beta_3(\text{disability status}_i) + \beta_4(\text{NSF directorate/office}_i) + \beta_5(\text{institution type}_i) + \beta_6(\text{early career}_i) + \beta_7(\text{experience}_i)$$

—where

- $y_i=1$ if individual i has the dichotomous attribute (e.g., is a respondent to the survey); and
- 0 if individual i does not have the attribute (e.g., is a nonrespondent to the survey).

This equation is the form of the logistic regression model for any dichotomous dependent variable, including response status.

B.4.3 Ordinal logistic regression

The logistic regression models for an ordinal dependent variable where the dependent variable (e.g., overall satisfaction) can take values 1, 2, ..., m , were as follows:

$$\ln \left[\frac{P(y_i \leq 1)}{1 - P(y_i \leq 1)} \right] = \beta_{10} + \beta_1(\text{race and ethnicity}_i) + \beta_2(\text{gender}_i) + \beta_3(\text{disability status}_i) + \beta_4(\text{NSF directorate/office}_i) + \beta_5(\text{institution type}_i) + \beta_6(\text{early career}_i) + \beta_7(\text{experience}_i)$$

$$\ln \left[\frac{P(y_i \leq 2)}{1 - P(y_i \leq 2)} \right] = \beta_{20} + \beta_1(\text{race and ethnicity}_i) + \beta_2(\text{gender}_i) + \beta_3(\text{disability status}_i) + \beta_4(\text{NSF } \frac{\text{directorates}}{\text{office}_i}) + \beta_5(\text{institution type}_i) + \beta_6(\text{early career}_i) + \beta_7(\text{experience}_i)$$

$$\ln \left[\frac{P(y_i \leq m)}{1 - P(y_i \leq m)} \right] = \beta_{m0} + \beta_1(\text{race and ethnicity}_i) + \beta_2(\text{gender}_i) + \beta_3(\text{disability status}_i) + \beta_4(\text{NSF directorate/office}_i) + \beta_5(\text{institution type}_i) + \beta_6(\text{early career}_i) + \beta_7(\text{experience}_i)$$

B.4.4 Multinomial logistic regression

The logistic regression models for a multinomial dependent variable (e.g., factors for improvement), where the dependent variable can take values 0, 1, 2, ..., m , were as follows:

$$\begin{aligned}
\ln \left[\frac{P(y_i = 1)}{P(y_i = 0)} \right] &= \beta_{10} + \beta_{11}(\text{race and ethnicity}_i) + \beta_{12}(\text{gender}_i) + \beta_{13}(\text{disability status}_i) \\
&\quad + \beta_{14}(\text{NSF directorate/office}_i) + \beta_{15}(\text{institution type}_i) + \beta_{16}(\text{early career}_i) \\
&\quad + \beta_{17}(\text{experience}_i) \\
\\
\ln \left[\frac{P(y_i = 2)}{P(y_i = 0)} \right] &= \beta_{20} + \beta_{21}(\text{race and ethnicity}_i) + \beta_{22}(\text{gender}_i) + \beta_{23}(\text{disability status}_i) \\
&\quad + \beta_{24}(\text{NSF directorate/office}_i) + \beta_{25}(\text{institution type}_i) + \beta_{26}(\text{early career}_i) \\
&\quad + \beta_{27}(\text{experience}_i) \\
\\
&\dots \\
\\
\ln \left[\frac{P(y_i = m)}{P(y_i = 0)} \right] &= \beta_{m0} + \beta_{m1}(\text{race and ethnicity}_i) + \beta_{m2}(\text{gender}_i) + \beta_{m3}(\text{disability status}_i) \\
&\quad + \beta_{m4}(\text{NSF directorate/office}_i) + \beta_{m5}(\text{institution type}_i) + \beta_{m6}(\text{early career}_i) \\
&\quad + \beta_{m7}(\text{experience}_i)
\end{aligned}$$

The 2019–2021 comparisons examined whether applicant and reviewer experiences had changed between years and whether any differences were conditioned by race and ethnicity, gender, institution type, or early career status. These analyses used six of the seven population characteristics (race and ethnicity, gender, NSF directorate/office, institution type, early career status, and experience as an applicant and reviewer)^{B-5} and the following five additional terms: Year (e.g., 2021 or 2019); Year * Race and ethnicity; Year * Gender; Year * Institution type; and Year * Early career status.

^{B-5} Disability status was not available in the 2019 data and, therefore, was not used in the year-to-year analysis.

Appendix C. Supplemental Findings

C.1 Descriptive Statistics

The descriptive statistics presented in table C.1.1 show frequencies for the seven population characteristics: race and ethnicity, gender, disability status, National Science Foundation (NSF) directorate, institution type, early career status, and NSF merit review experience (applicant, reviewer, or both).

Table C.1.1. Profile of 2021 Merit Review Survey frequencies by population characteristic

Population characteristics (n = 17,330)	Count	Percentage
Race and ethnicity		
Asian, non-Hispanic	3,104	17.91
Black or African American, non-Hispanic	616	3.55
Hispanic, any race	1,095	6.32
White, non-Hispanic	10,661	61.25
Multiple races, non-Hispanic	279	1.61
Other, NHPI, AIAN, HSP	275	1.59
Native Hawaiian or Other Pacific Islander, non-Hispanic (NHPI)	8	0.05
American Indian or Alaska Native, non-Hispanic (AIAN)	29	0.17
Other race, non-Hispanic (other)	238	1.37
Unknown or do not wish to provide	1,300	7.50
Gender		
Male	10,629	61.33
Female	5,885	33.96
Other, do not wish to provide, or unknown	816	4.71
Disability status		
Yes—hearing, mobility, vision, or other disability	532	3.07
No	16,128	93.06
Do not wish to provide	670	3.87
NSF directorate		
BIO	2,478	14.30
CISE	2,073	11.96
EHR	1,859	10.73
ENG	2,830	16.33
GEO	1,845	10.65
MPS (ref.)	3,723	21.48
SBE	1,867	10.77
OD (OIA and OISE)	121	0.70
Other	534	3.08
Institution type		
R1/MSI	923	5.33
R1/non-MSI (ref.)	6,591	38.03
Non-R1/MSI	1,272	7.34

Population characteristics (n = 17,330)	Count	Percentage
Non-R1/non-MSI	5,832	33.65
For-profit organization	1,084	6.26
Government (local, state, federal, or tribal)	181	1.04
Nonprofit organization	491	2.83
Other, primary and secondary institution, R1/unknown, non-R1/unknown, federally funded research and development center	956	5.52
Early career (degree less than 10 years ago)		
Yes	4,255	24.55
No	13,075	75.45
Experience		
Applicant only	5,554	32.05
Reviewer only	3,494	20.16
Both applicant and reviewer	8,282	47.79

Note: BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; OD = Office of the Director; OIA = Office of Integrative Activities; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

Tables C.1.2, C.1.3, and C.1.4 present descriptive statistics for all outcome variables organized by research dimension (satisfaction, burden, quality, and fairness) for all respondents, applicants, and reviewers, respectively.

Table C.1.2. Descriptive statistics for all applicant and reviewer outcomes

Respondent outcomes	Count	Percentage
Number of respondents (N)	17,330	100.0
Satisfaction		
Applicant satisfaction index (N = 13,836)		
Minimum	0	NA
Median	68.4	NA
Mean	65.7	NA
Maximum	100	NA
Reviewer satisfaction index (N = 11,776)		
Minimum	0	NA
Median	80	NA
Mean	74.8	NA
Maximum	100	NA
Overall satisfaction with merit review process (Q21E) (N = 17,286)		
Strongly agree	3,117	18.4
Agree	6,983	40.4
Neither agree nor disagree	3,665	21.2
Disagree	2,415	14.0
Strongly disagree	1,046	6.0

Respondent outcomes	Count	Percentage
Burden		
Number of proposal reviews (Q4) (N = 11,728)		
Minimum	0	NA
Median	7	NA
Mean	10.4	NA
Maximum	100	NA
Declined NSF request to conduct a review (Q5A–Q5C composite) (N = 6,996)		
Yes	4,742	67.8
No	2,254	32.2
Reason for declining: lack of time (Q6B) (N = 4,558)		
To a great extent	2,848	62.5
To a moderate extent	678	14.9
To a small extent	360	7.9
Not at all	672	14.7
Reason for declining: too many NSF review requests (Q6D) (N = 4,404)		
To a great extent	202	4.6
To a moderate extent	452	10.3
To a small extent	958	21.8
Not at all	2,792	63.4
Reason for declining: competing professional pressures (Q6E) (N = 4,561)		
To a great extent	2,471	54.2
To a moderate extent	920	20.2
To a small extent	380	8.3
Not at all	790	17.3
Reason for declining: reviewer commitments to other federal agencies (Q6G) (N = 4,545)		
To a great extent	220	4.8
To a moderate extent	418	9.2
To a small extent	686	15.1
Not at all	3,221	70.9
Reason for declining: inability to travel to face-to-face panel (Q6H) (N = 2,120)		
To a great extent	474	22.4
To a moderate extent	346	16.3
To a small extent	279	13.2
Not at all	1,021	48.2
Hours spent conducting the most recent review (Q7) (N = 11,696)		
Minimum	0	NA
Median	4	NA
Mean	6.7	NA
Maximum	501	NA

Respondent outcomes	Count	Percentage
Time of day for conducting reviews (Q8) (N = 11,731)		
Within normal working hours	1,043	8.9
Outside normal working hours	3,238	27.6
Both outside and within normal working hours	7,450	63.5
Employers view of participating as a reviewer (Q9) (N = 11,739)		
Within the scope of work duties	8,311	70.8
Outside the scope of work duties	1,542	13.1
Don't know	1,886	16.1
Number of proposals submitted to NSF (Q12) (N = 13,836)		
Minimum	0	NA
Median	2	NA
Mean	2.7	NA
Maximum	38	NA
Hours spent preparing the most recent proposal (Q13) (N = 13,731)		
Minimum	0	NA
Median	80	NA
Mean	125.9	NA
Maximum	1,000	NA
Amount of effort to write and complete a proposal compared with other agencies (Q15) (N = 8,493)		
More effort	2,628	30.9
Nearly the same effort	4,941	58.2
Less effort	924	10.9
Quality		
Proposals reviewed have been of high quality (Q10A) (N = 11,732)		
Strongly agree	1,889	16.1
Agree	6,762	57.6
Neither agree nor disagree	2,117	18.0
Disagree	833	7.1
Strongly disagree	131	1.1
Extent to which reviews improved understanding of proposal process (Q18A) (N = 12,386)		
To a great extent	1,006	8.1
To a moderate extent	3,771	30.5
To a small extent	4,185	33.8
Not at all	3,424	27.6
Extent to which reviews provided useful information for improving next proposal (Q18B) (N = 12,384)		
To a great extent	2,162	17.5
To a moderate extent	4,763	38.5
To a small extent	3,810	30.8
Not at all	1,649	13.2

Respondent outcomes	Count	Percentage
Extent to which reviews influenced decision to submit to another funding agency (Q18C) (N = 12,364)		
To a great extent	1,491	12.1
To a moderate extent	2,429	19.7
To a small extent	2,755	22.3
Not at all	5,689	46.0
Extent to which reviews provided discouraged from revising and submitting your proposals to NSF (Q18D) (N = 12,384)		
To a great extent	2,470	20.0
To a moderate extent	2,643	21.3
To a small extent	3,033	24.5
Not at all	4,238	34.2
Satisfaction with quality of information provided by NSF (Q19A) (N = 12,404)		
Very dissatisfied	529	4.3
Somewhat dissatisfied	1,083	8.7
Neither dissatisfied nor satisfied	2,027	16.3
Somewhat satisfied	3,808	30.7
Very satisfied	4,852	39.1
Not applicable	105	0.9
Extent to which written reviews are thorough (Q20A) (N = 13,803)		
Strongly agree	933	6.8
Agree	5,465	39.6
Neither agree nor disagree	3,386	24.5
Disagree	2,537	18.4
Strongly disagree	1,293	9.4
Not applicable	189	1.4
Extent to which written reviews are technically sound (Q20B) (N = 13,792)		
Strongly agree	854	6.2
Agree	5,572	40.4
Neither agree nor disagree	3,784	27.4
Disagree	2,147	15.6
Strongly disagree	1,223	8.9
Not applicable	212	1.5
Extent to which written reviews are of high quality overall (Q20C) (N = 13,782)		
Strongly agree	879	6.4
Agree	4,659	33.8
Neither agree nor disagree	4,445	32.3
Disagree	2,375	17.2
Strongly disagree	1,232	8.9
Not applicable	192	1.4

Respondent outcomes	Count	Percentage
Extent to which panel summary or summaries are of high quality (Q20D) (N = 13,790)		
Strongly agree	1,558	11.3
Agree	5,491	39.8
Neither agree nor disagree	3,459	25.1
Disagree	1,799	13.1
Strongly disagree	1,061	7.7
Not applicable	422	3.1
Extent to which information provided regarding the outcomes of the competition is of high quality (Q20E) (N = 13,787)		
Strongly agree	934	6.8
Agree	3,872	28.1
Neither agree nor disagree	4,331	31.4
Disagree	2,565	18.6
Strongly disagree	1,460	10.6
Not applicable	625	4.5
Factors to improve the NSF merit review process (Q22) (N = 17,047)		
Timeliness of decisions about, and responsiveness to, proposals by NSF staff	3,798	22.3
Quality of feedback to PIs in the form of comments in written reviews	4,753	27.9
Quality of feedback to PIs in the form of comments in panel summaries	2,622	15.4
Quality of PI conversations with, and written comments from, program officers	2,367	13.9
Quality of information available during proposal submission	1,279	7.5
Quality of the review process from the perspective of a reviewer	2,228	13.1
Fairness		
Individuals submitting proposals are treated fairly (Q10B) (N = 11,688)		
Strongly agree	3,409	29.2
Agree	5,550	47.5
Neither agree nor disagree	1,768	15.1
Disagree	719	6.2
Strongly disagree	242	2.1
Individuals submitting proposals are treated fairly (Q20H) (N = 13,790)		
Strongly agree	2,602	18.9
Agree	5,224	37.9
Neither agree nor disagree	3,240	23.5
Disagree	1,266	9.2
Strongly disagree	1,097	8.0
Not applicable	361	2.6

Respondent outcomes	Count	Percentage
NSF's merit review process is fair overall (Q21A) (N = 17,285)		
Strongly agree	3,361	19.4
Agree	7,982	46.2
Neither agree nor disagree	3,403	19.7
Disagree	1,828	10.6
Strongly disagree	711	4.1

Note: PI = principal investigator

Table C.1.3. Descriptive statistics for all applicant outcomes

Applicant outcomes	Count	Percentage
Number of applicants (N)	13,836	100.0
Satisfaction		
Applicant satisfaction index (N = 13,836)		
Minimum	0	NA
Median	68.4	NA
Mean	65.7	NA
Maximum	100	NA
Overall satisfaction with merit review process (Q21E) (N = 13,802)		
Strongly agree	2,247	16.3
Agree	5,206	37.7
Neither agree nor disagree	3,117	22.6
Disagree	2,221	16.1
Strongly disagree	1,011	7.3
Burden		
Number of proposals submitted to NSF (Q12) (N = 13,836)		
Minimum	0	NA
Median	2	NA
Mean	2.7	NA
Maximum	38	NA
Hours spent preparing the most recent proposal (Q13) (N = 13,731)		
Minimum	0	NA
Median	80	NA
Mean	125.9	NA
Maximum	1,000	NA
Amount of effort to write and complete a proposal compared with other agencies (Q15) (N = 8,493)		
More effort	2,628	30.9
Nearly the same effort	4,941	58.2
Less effort	924	10.9

Applicant outcomes	Count	Percentage
Quality		
Extent to which reviews improved understanding of proposal process (Q18A) (N = 12,386)		
To a great extent	1,006	8.1
To a moderate extent	3,771	30.5
To a small extent	4,185	33.8
Not at all	3,424	27.6
Extent to which reviews provided useful information for improving next proposal (Q18B) (N = 12,384)		
To a great extent	2,162	17.5
To a moderate extent	4,763	38.5
To a small extent	3,810	30.8
Not at all	1,649	13.3
Extent to which reviews influenced decision to submit to another funding agency (Q18C) (N = 12,364)		
To a great extent	1,491	12.1
To a moderate extent	2,429	19.7
To a small extent	2,777	22.3
Not at all	5,689	46.0
Extent to which reviews provided discouraged you from revising and submitting your proposals to NSF (Q18D) (N = 12,384)		
To a great extent	2,470	20.0
To a moderate extent	2,643	21.3
To a small extent	3,033	24.5
Not at all	4,238	34.2
Satisfaction with quality of information provided by NSF (Q19A) (N = 12,404)		
Very dissatisfied	529	4.3
Somewhat dissatisfied	1,083	8.7
Neither dissatisfied nor satisfied	2,027	16.3
Somewhat satisfied	3,808	30.7
Very satisfied	4,852	39.1
Not applicable	104	0.9
Extent to which written reviews are thorough (Q20A) (N = 13,803)		
Strongly agree	933	6.8
Agree	5,465	39.6
Neither agree nor disagree	3,386	24.5
Disagree	2,537	18.4
Strongly disagree	1,293	9.4
Not applicable	189	1.4

Applicant outcomes	Count	Percentage
Extent to which written reviews are technically sound (Q20B) (N = 13,792)		
Strongly agree	854	6.2
Agree	5,572	40.4
Neither agree nor disagree	3,784	27.4
Disagree	2,147	15.6
Strongly disagree	1,223	8.9
Not applicable	212	1.5
Extent to which written reviews are of high quality overall (Q20C) (N = 13,782)		
Strongly agree	865	6.4
Agree	4,659	33.8
Neither agree nor disagree	4,445	32.3
Disagree	2,375	17.2
Strongly disagree	1,232	8.9
Not applicable	192	1.4
Extent to which panel summary or summaries are of high quality (Q20D) (N = 13,790)		
Strongly agree	1,558	11.3
Agree	5,491	39.8
Neither agree nor disagree	3,459	25.1
Disagree	1,799	13.1
Strongly disagree	1,061	7.7
Not applicable	422	3.1
Extent to which information provided regarding the outcomes of the competition is of high quality (Q20E) (N = 13,787)		
Strongly agree	934	6.8
Agree	3,872	28.1
Neither agree nor disagree	4,331	31.4
Disagree	2,565	18.6
Strongly disagree	1,460	10.6
Not applicable	625	4.5
Factors to improve the NSF merit review process (Q22) (N = 13,655)		
Timeliness of decisions about, and responsiveness to, proposals by NSF staff	3,309	24.2
Quality of feedback to PIs in the form of comments in written reviews	3,954	29.0
Quality of feedback to PIs in the form of comments in panel summaries	1,936	14.2
Quality of PI conversations with, and written comments from, program officers	1,834	13.4
Quality of information available during proposal submission	891	6.5
Quality of the review process from the perspective of a reviewer	1,731	12.7

Applicant outcomes	Count	Percentage
Fairness		
Individuals submitting proposals are treated fairly (Q20H) (N = 13,790)		
Strongly agree	2,602	18.9
Agree	5,224	37.9
Neither agree nor disagree	3,240	23.5
Disagree	1,266	9.2
Strongly disagree	1,097	8.0
Not applicable	361	2.6
NSF's merit review process is fair overall (Q21A) (N = 13,799)		
Strongly agree	2,336	16.9
Agree	6,126	44.4
Neither agree nor disagree	2,961	21.5
Disagree	1,689	12.2
Strongly disagree	687	5.0

Note: PI = principal investigator

Table C.1.4. Descriptive statistics for all reviewer outcomes

Reviewer outcomes	Count	Percentage
Number of reviewers (N)	11,776	100.0
Satisfaction		
Reviewer satisfaction index (N = 11,776)		
Minimum	0	NA
Median	80	NA
Mean	74.8	NA
Maximum	100	NA
Overall satisfaction with merit review process (Q21E) (N = 11,747)		
Strongly agree	2,517	21.4
Agree	5,255	44.7
Neither agree nor disagree	2,260	19.2
Disagree	1,286	11.0
Strongly disagree	429	3.7
Burden		
Number of proposal reviews (Q4) (N = 11,728)		
Minimum	0	NA
Median	7	NA
Mean	10.4	NA
Maximum	100	NA
Hours spent conducting the most recent review (Q7) (N = 11,696)		
Minimum	0	NA
Median	4	NA
Mean	6.7	NA

Reviewer outcomes	Count	Percentage
Maximum	501	NA
Time of day for conducting reviews (Q8) (N = 11,731)		
Within normal working hours	1,043	8.9
Outside normal working hours	3,238	27.6
Both outside and within normal working hours	7,450	63.5
Employers view of participating as a reviewer (Q9) (N = 11,739)		
Within the scope of work duties	8,311	70.8
Outside the scope of work duties	1,542	13.1
Don't know	1,886	16.1
Declined NSF request to conduct a review (Q5A–Q5C composite) (N = 6,996)		
Yes	4,742	67.8
No	2,254	32.2
Reason for declining: lack of time (Q6B) (N = 4,558)		
To a great extent	2,848	62.5
To a moderate extent	678	14.9
To a small extent	360	7.9
Not at all	672	14.7
Reason for declining: too many NSF review requests (Q6D) (N = 4,404)		
To a great extent	202	4.6
To a moderate extent	452	10.3
To a small extent	958	21.8
Not at all	2,792	63.4
Reason for declining: competing professional pressures (Q6E) (N = 4,561)		
To a great extent	2,471	54.2
To a moderate extent	920	20.2
To a small extent	380	8.3
Not at all	790	17.3
Reason for declining: reviewer commitments to other federal agencies (Q6G) (N = 4,545)		
To a great extent	220	4.8
To a moderate extent	418	9.2
To a small extent	686	15.1
Not at all	3,221	70.9
Reason for declining: inability to travel to face-to-face panel (Q6H) (N = 2,120)		
To a great extent	474	22.4
To a moderate extent	346	16.3
To a small extent	279	13.2
Not at all	1,021	48.2

Reviewer outcomes	Count	Percentage
Quality		
Proposals reviewed have been of high quality (Q10A) (N = 11,732)		
Strongly agree	1,889	16.1
Agree	6,762	57.6
Neither agree nor disagree	2,117	18.0
Disagree	833	7.1
Strongly disagree	131	1.1
Factors to improve the NSF merit review process (Q22) (N = 11,578)		
Timeliness of decisions about, and responsiveness to, proposals by NSF staff	2,544	22.0
Quality of feedback to PIs in the form of comments in written reviews	3,063	26.5
Quality of feedback to PIs in the form of comments in panel summaries	1,994	17.2
Quality of PI conversations with, and written comments from, program officers	1,683	14.5
Quality of information available during proposal submission	829	7.2
Quality of the review process from the perspective of a reviewer	1,465	12.7
Fairness		
Individuals submitting proposals are treated fairly (Q10B) (N = 11,688)		
Strongly agree	3,409	29.2
Agree	5,550	47.5
Neither agree nor disagree	1,768	15.1
Disagree	719	6.2
Strongly disagree	242	2.1
NSF's merit review process is fair overall (Q21A) (N = 11,748)		
Strongly agree	2,735	23.3
Agree	5,874	50.0
Neither agree nor disagree	1,916	16.3
Disagree	941	8.0
Strongly disagree	282	2.4

Note: PI = principal investigator

C.2 Satisfaction Dimension: Regression Tables

Regression coefficients for figures 4.2.1 to 4.2.12 in chapter 4.2 (*Satisfaction With the Merit Review Process*) are presented in tables C.2.1 to C.2.6.

Table C.2.1. Applicant and reviewer agreement with being satisfied with the NSF merit review process

Characteristics	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Race and ethnicity					
Asian, non-Hispanic	-0.01*	-0.03***	-0.01	0.01	0.04***
Black or African American, non-Hispanic	0.01	0.03	0.03	-0.08***	0.01

Characteristics	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Hispanic	0.00	0.01	0.01	-0.05**	0.02
White, non-Hispanic (ref.)	-	-	-	-	-
Multiple races, non-Hispanic	-0.00	0.01	0.06	-0.03	-0.03
Other, NHPI, AIAN	0.05**	0.04	0.06	-0.14***	-0.01
Do not wish to provide or unknown	0.01	0.04**	0.05*	-0.06***	-0.03**
Gender					
Female	-0.01**	-0.01	0.02*	-0.00	0.00
Male (ref.)	-	-	-	-	-
Other or do not wish to provide or unknown	0.02*	0.04*	0.02	-0.07**	-0.01
Disability status					
Yes—hearing, mobility, vision, or other disability	0.05***	0.06**	0.01	-0.08***	-0.04**
No—no disability (ref.)	-	-	-	-	-
Do not wish to provide	0.03**	0.04*	0.05	-0.07*	-0.05**
Directorate					
BIO	-0.02**	-0.01	0.00	0.01	0.01
CISE	-0.01	0.01	0.02	-0.03*	0.01
EHR	-0.03***	-0.03***	-0.01	0.03*	0.04**
ENG	-0.00	0.02*	0.02*	-0.03**	-0.01
GEO	-0.02***	-0.00	0.01	0.01	0.00
MPS (ref.)	-	-	-	-	-
SBE	-0.01	0.01	-0.01	0.00	0.01
OD (OIA and OISE)	-0.04***	0.05	0.04	0.01	-0.05*
Other	0.02	0.03	0.01	-0.09***	0.03
Institution type					
R1/MSI	0.01	-0.01	0.01	0.02	-0.02
R1/non-MSI (ref.)	-	-	-	-	-
Non-R1/MSI	-0.00	-0.03**	-0.03*	0.01	0.05***
Non-R1/non-MSI	-0.00	-0.02*	-0.00	0.00	0.02*
For-profit organization	0.05***	0.02	-0.07***	-0.06**	0.06***
Government (local, state, federal, or tribal)	0.00	-0.06*	0.01	0.04	0.01
Nonprofit organization	-0.00	-0.01	-0.00	0.05	-0.03*
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.01	-0.03*	-0.01	-0.01	0.04**
Early career (degree ≤10 years ago)					
Early career	-0.01***	-0.02***	-0.00	0.04***	0.00
Not early career (ref.)	-	-	-	-	-

Characteristics	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Role in the merit review process					
Applicant only	0.05***	0.08***	0.06***	-0.11***	-0.08***
Reviewer only	-0.04***	-0.08***	-0.05***	0.10***	0.06***
Both applicant and reviewer (ref.)	-	-	-	-	-

Note: $N = 17,286$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented in coefficients.

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences
A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.2.2. Applicant satisfaction index

Characteristics	Coefficient
Race and ethnicity	
Asian, non-Hispanic	1.50***
Black or African American, non-Hispanic	-0.63
Hispanic	-0.22
White, non-Hispanic (ref.)	-
Multiple races, non-Hispanic	-1.17
Other, NHPI, AIAN	-7.85***
Do not wish to provide or unknown	-3.97***
Gender	
Female	1.72***
Male (ref.)	-
Other or do not wish to provide or unknown	-3.36**
Disability status	
Yes—hearing, mobility, vision, or other disability	-4.83***
No—no disability (ref.)	-
Do not wish to provide	-5.56***
Directorate	
BIO	2.86***
CISE	-0.50
EHR	4.33***
ENG	-1.41**
GEO	3.06***
MPS (ref.)	-
SBE	2.41***
OD (OIA and OISE)	0.12

Characteristics	Coefficient
Other	-3.14*
Institution type	
R1/MSI	-0.93
R1/non-MSI (ref.)	-
Non-R1/MSI	2.72***
Non-R1/non-MSI	1.22**
For-profit organization	-2.14*
Government (local, state, federal, or tribal)	2.93
Nonprofit organization	0.06
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	-0.42
Early career (degree ≤10 years ago)	
Yes	3.74***
No (ref.)	-
Is respondent both investigator and reviewer	
Yes	5.81***
No (ref.)	-

Note: $N = 13,836$

Linear regression

Findings are presented in coefficients.

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences
A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.2.3. Reviewer satisfaction index

Characteristics	Coefficient
Race and ethnicity	
Asian, non-Hispanic	1.64***
Black or African American, non-Hispanic	-2.07
Hispanic	-1.15
White, non-Hispanic (ref.)	-
Multiple races, non-Hispanic	-1.11
Other, NHPI, AIAN	-7.48***
Do not wish to provide or unknown	-3.79***
Gender	
Female	-0.19
Male (ref.)	-
Other or do not wish to provide or unknown	-5.13***
Disability status	

Characteristics	Coefficient
Yes—hearing, mobility, vision, or other disability	-5.18***
No—no disability (ref.)	-
Do not wish to provide	-6.73***
Directorate	
BIO	0.69
CISE	0.03
EHR	2.33***
ENG	-1.84**
GEO	-0.31
MPS (ref.)	-
SBE	-1.08
OD (OIA and OISE)	0.10
Other	-0.42
Institution type	
R1/MSI	-0.81
R1/non-MSI (ref.)	-
Non-R1/MSI	2.45**
Non-R1/non-MSI	0.73
For-profit organization	2.51
Government (local, state, federal, or tribal)	1.72
Nonprofit organization	0.49
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	2.27**
Early career (degree ≤10 years ago)	
Yes	0.07
No (ref.)	-
Is respondent both investigator and reviewer	
Yes	-6.72***
No (ref.)	-

Note: $N = 11,776$

Linear regression

Findings are presented in coefficients.

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences
A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.2.4.A. Applicant and reviewer agreement to being satisfied with the NSF merit review process—2021 and 2019 differences

Characteristics	Level 1	Level 2	Level 3	Level 4
Year				
2021	0.33***	0.41***	0.43***	0.63***
2019 (ref.)	-	-	-	-
Race and ethnicity				
Year * Asian, non-Hispanic	1.04	0.97	0.98	0.94
Year * Black or African American, non-Hispanic	0.62	0.73	0.84	1.04
Year * Hispanic	1.29	1.05	1.11	1.12
Year * White, non-Hispanic (ref.)	-	-	-	-
Year * Multiple races, non-Hispanic	1.39	1.12	0.80	0.88
Year * Other, NHPI, AIAN	0.65	0.48	0.45**	0.75
Year * Do not wish to provide or unknown	1.15	1.11	0.97	0.95
Gender				
Year * Female	0.85	0.94	0.97	1.09
Year * Male (ref.)	-	-	-	-
Year * Other or do not wish to provide or unknown	0.84	0.92	0.94	1.06
Institution type				
Year * Academic (ref.)	-	-	-	-
Year * For-profit organization	0.50***	0.60***	0.71***	0.82
Year * Government (local, state, federal, or tribal)	1.06	1.42	1.66*	1.26
Year * Federally funded R&D	0.63	1.28	1.13	1.73*
Year * Nonprofit organization	1.39	1.27	1.06	0.87
Year * Other, primary institution, secondary institution	1.71	1.31	1.18	1.05
Early career (degree ≤10 years ago)				
Year * Early career	0.75	0.78**	0.86*	0.95
Year * Not early career (ref.)	-	-	-	-

Note: N = 42,230

Generalized Ordinal Logistic Regression (Gologit)

Variable has five ordinal outcomes: 1 = strongly disagree; 2 = disagree; 3 = neither agree nor disagree; 4 = agree; 5 = strongly agree.

Coefficients presented in odds ratios and across four levels indicating to a greater extent.

Level 1: odds ratios of “strongly disagree” versus “disagree,” “neither agree nor disagree,” “agree,” and “Strongly agree”

Level 2: odds ratios of “strongly disagree” and “disagree” versus “neither agree nor disagree,” “agree,” and “strongly agree”

Level 3: odds ratios of “strongly disagree,” “disagree,” and “neither agree nor disagree” versus “agree” and “strongly agree”

Level 4: odds ratios of “strongly disagree,” “disagree,” “neither agree nor disagree,” and “agree” versus “strongly agree”

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

Findings are weighted to adjust for nonresponse bias.

See appendix B.4 for the full list of independent variables included in the model.

AIAN = American Indian or Alaska Native; HSP = Hispanic; NHPI = Native Hawaiian or Other Pacific Islander; R&D = research and development

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ A cell with a dash (-) equals no data available

Table C.2.4.B. Applicant and reviewer agreement to being satisfied with the NSF merit review process—2021 and 2019 differences (average marginal effects for 2021)

Characteristics	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Year					
2021	0.04***	0.08***	0.08***	-0.14***	-0.07***
2019 (ref.)	-	-	-	-	-

Note: $N = 42,230$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME) for the year effect of “2021” modeled in table C.2.4.A.

Findings are weighted to adjust for nonresponse bias.

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

See appendix B.4 for the full list of independent variables included in the model.

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.2.5. Applicant satisfaction index—2021 and 2019 differences

Characteristics	Coefficient
Year	
2021	-0.05***
2019 (ref.)	-
Race and ethnicity	
Year * Asian, non-Hispanic	0.00
Year * Black or African American, non-Hispanic	0.00
Year * Hispanic	0.02*
Year * White, non-Hispanic (ref.)	-
Year * Multiple races, non-Hispanic	-0.01
Year * Other, NHPI, AIAN, HSP	-0.09***
Year * Do not wish to provide or unknown	0.00
Gender	
Year * Female	0.01
Year * Male (ref.)	-
Year * Other or do not wish to provide or unknown	-0.02
Institution type	
Year * Academic (ref.)	-
Year * For-profit organization	-0.05***
Year * Government (local, state, federal, or tribal)	-0.00
Year * Federally funded R&D	-0.01
Year * Nonprofit organization	0.01
Year * Other, primary institution, secondary institution	0.01

Characteristics	Coefficient
Early career (degree ≤10 years ago)	
Year * Early career	-0.01
Year * Not early career (ref.)	-

Note: $N = 32,612$

Linear regression

Findings are presented in coefficients.

Findings are weighted to adjust for nonresponse bias.

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

See appendix B.4 for the full list of independent variables included in the model.

AIAN = American Indian or Alaska Native; HSP = Hispanic; NHPI = Native Hawaiian or Other Pacific Islander; R&D = research and development

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.2.6. Reviewer satisfaction index—2021 and 2019 differences

Characteristics	Coefficient
Year	
2021	-0.05***
2019 (ref.)	-
Race and ethnicity	
Year * Asian, non-Hispanic	-0.01
Year * Black or African American, non-Hispanic	-0.01
Year * Hispanic	-0.00
Year * White, non-Hispanic (ref.)	-
Year * Multiple races, non-Hispanic	-0.01
Year * Other, NHPI, AIAN, HSP	-0.06
Year * Do not wish to provide or unknown	-0.02
Gender	
Year * Female	0.01
Year * Male (ref.)	-
Year * Other or do not wish to provide or unknown	-0.04*
Institution type	
Year * Academic (ref.)	-
Year * For-profit organization	-0.01
Year * Government (local, state, federal, or tribal)	0.03*
Year * Federally funded R&D	0.04*
Year * Nonprofit organization	0.02
Year * Other, primary institution, secondary institution	0.02

Characteristics	Coefficient
Early career (degree ≤10 years ago)	
Year * Early career	-0.01
Year * Not early career (ref.)	-

Note: N = 30,108

Linear regression

Findings are presented in coefficients.

Findings are weighted to adjust for nonresponse bias.

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

See appendix B.4 for the full list of independent variables included in the model.

AIAN = American Indian or Alaska Native; HSP = Hispanic; NHPI = Native Hawaiian or Other Pacific Islander; R&D = research and development

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C.3 Burden: Regression Tables

Regression coefficients for figures 4.3.1 to 4.3.5 in chapter 4.2 (*Burden in the Merit Review Process*) are presented in tables C.3.1 to C.3.9.

Table C.3.1. Number of proposals submitted and time spent preparing a proposal

Characteristics	Number of proposals submitted	Hours to prepare proposal
Race and ethnicity		
Asian, non-Hispanic	0.11***	0.31***
Black or African American, non-Hispanic	0.00	0.16**
Hispanic	0.00	0.08*
White, non-Hispanic (ref.)	-	-
Multiple races, non-Hispanic	0.07*	0.05
Other, NHPI, AIAN	0.07*	0.24***
Do not wish to provide or unknown	0.11***	0.24***
Gender		
Female	-0.05***	0.09***
Male (ref.)	-	-
Other or do not wish to provide or unknown	-0.04	0.00
Disability status		
Yes—hearing, mobility, vision, or other disability	0.04	0.05
No—no disability (ref.)	-	-
Do not wish to provide	-0.01	0.08
Directorate		
BIO	-0.06***	0.29***
CISE	0.20***	0.14***
EHR	0.07***	0.09**
ENG	0.14***	0.23***
GEO	0.01	0.14***
MPS (ref.)	-	-

Characteristics	Number of proposals submitted	Hours to prepare proposal
SBE	-0.11***	0.02
OD (OIA and OISE)	-0.24***	0.21*
Other	-0.05*	0.13**
Institution type		
R1/MSI	0.05*	0.04
R1/non-MSI (ref.)	-	-
Non-R1/MSI	-0.09***	0.07*
Non-R1/non-MSI	-0.07***	0.01
For-profit organization	-0.28***	0.12***
Government (local, state, federal, or tribal)	-0.17**	0.11
Nonprofit organization	-0.04	-0.08
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	-0.12***	0.08
Early career (degree ≤10 years ago)		
Early career	0.05***	0.12***
Not early career (ref.)	-	-
Is respondent both investigator and reviewer		
Yes	0.14***	-0.05***
No (ref.)	-	-

Note: number of proposals submitted $N = 13,836$; hours to prepare recent proposal $N = 13,731$

Linear regression

Dependent variables are log transformed.

Findings presented are using coefficients.

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3.2. Effort to submit to NSF compared with other federal agencies' proposal systems

Characteristics	Response options - Less effort (AME)	Response options - Nearly the same effort (AME)	Response options - More effort (AME)
Race and ethnicity			
Asian, non-Hispanic	0.02*	-0.01	-0.00
Black or African American, non-Hispanic	0.05*	-0.07*	0.02
Hispanic	0.00	0.01	-0.02
White, non-Hispanic (ref.)	-	-	-
Multiple races, non-Hispanic	0.00	0.00	-0.00
Other, NHPI, AIAN	-0.00	-0.04	0.04
Do not wish to provide or unknown	-0.02	-0.02	0.04

Characteristics	Response options - Less effort (AME)	Response options - Nearly the same effort (AME)	Response options - More effort (AME)
Gender			
Female	-0.01	0.02	-0.01
Male (ref.)	-	-	-
Other or do not wish to provide or unknown	0.05	-0.09*	0.04
Disability status			
Yes—hearing, mobility, vision, or other disability	0.04	-0.11**	0.07*
No—no disability (ref.)	-	-	-
Do not wish to provide	-0.05***	0.01	0.04
Directorate			
BIO	0.01	0.01	-0.03
CISE	0.08***	-0.11***	0.04*
EHR	0.05**	-0.05*	0.01
ENG	0.03***	-0.06***	0.02
GEO	0.02	-0.07***	0.05**
MPS (ref.)	-	-	-
SBE	0.09***	-0.07**	-0.02
OD (OIA and OISE)	0.07	-0.19**	0.12
Other	0.04	-0.06	0.02
Institution type			
R1/MSI	0.01	0.03	-0.04
R1/non-MSI (ref.)	-	-	-
Non-R1/MSI	0.00	0.02	-0.02
Non-R1/non-MSI	0.01	0.00	-0.02
For-profit organization	0.00	-0.01	0.00
Government (local, state, federal, or tribal)	-0.06	-0.14	0.20*
Nonprofit organization	-0.04*	-0.10**	0.14***
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.00	0.00	-0.01
Early career (degree ≤10 years ago)			
Early career	0.00	-0.04**	0.03**
Not early career (ref.)	-	-	-
Is respondent both investigator and reviewer			
Yes	0.00	0.08***	-0.08***
No (ref.)	-	-	-

Note: N = 8,493

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of

Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences
A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3.3. Number of reviews conducted and the amount of time to conduct a review

Characteristics	Number of reviews conducted	Hours to conduct a review
Race and ethnicity		
Asian, non-Hispanic	-0.06*	0.28***
Black or African American, non-Hispanic	0.09	0.17***
Hispanic	0.09*	0.06*
White, non-Hispanic (ref.)	-	-
Multiple races, non-Hispanic	0.11	0.04
Other, NHPI, AIAN	-0.01	0.14*
Do not wish to provide or unknown	-0.02	0.20***
Gender		
Female	0.17***	-0.03*
Male (ref.)	-	-
Other or do not wish to provide or unknown	0.03	-0.01
Disability status		
Yes—hearing, mobility, vision, or other disability	0.04	-0.03
No—no disability (ref.)	-	-
Do not wish to provide	0.01	-0.03
Directorate		
BIO	-0.14***	0.04
CISE	0.28***	-0.11***
EHR	0.19***	-0.15***
ENG	0.19***	-0.11***
GEO	-0.36***	0.17***
MPS (ref.)	-	-
SBE	-0.46***	-0.05*
OD (OIA and OISE)	-0.17	-0.09
Other	-0.20**	0.09
Institution type		
R1/MSI	-0.01	0.01
R1/non-MSI (ref.)	-	-
Non-R1/MSI	-0.13***	0.10**
Non-R1/non-MSI	-0.04	0.04**
For-profit organization	-0.24***	-0.15**
Government (local, state, federal, or tribal)	-0.24***	0.19*
Nonprofit organization	-0.06	0.01

Characteristics	Number of reviews conducted	Hours to conduct a review
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	-0.19***	0.08**
Early career (degree ≤10 years ago)		
Early career	-0.18***	0.05**
Not early career (ref.)	-	-
Is respondent both investigator and reviewer		
Yes	0.04	0.00
No (ref.)	-	-

Note: number of reviews conducted $N = 11,728$; hours to conduct recent review $N = 11,696$

Linear regression

Dependent variables are log transformed.

Findings are presented using coefficients.

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences
A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3.4. Time-of-day respondents conducted reviews and employer's view of participating as a reviewer

Characteristics	Time of day to conduct reviews response options - During normal working hours (AME)	Time of day to conduct reviews response options - Outside of normal working hours (AME)	Time of day to conduct reviews response options - Both during and outside normal working hours (AME)	Employer's view of conducting reviews response options - Within scope of normal work duties (AME)	Employer's view of conducting reviews response options - Outside scope of normal work duties (AME)	Employer's view of conducting reviews response options - I don't know (AME)
Race and ethnicity						
Asian, non-Hispanic	-0.02**	-0.03*	0.05***	-0.13***	0.04***	0.09***
Black or African American, non-Hispanic	-0.05***	0.06*	-0.00	-0.11***	0.06**	0.05*
Hispanic	-0.02*	-0.03	0.05**	-0.04*	0.03*	0.02
White, non-Hispanic (ref.)	-	-	-	-	-	-
Multiple races, non-Hispanic	-0.02	-0.03	0.05	-0.06	0.04	0.02
Other, NHPI, AIAN	-0.02	0.03	-0.01	-0.14***	0.10**	0.04
Do not wish to provide or unknown	-0.01	0.02	-0.01	-0.11***	-0.00	0.11***
Gender						
Female	-0.02***	0.08***	-0.06***	-0.06***	0.03***	0.04***
Male (ref.)	-	-	-	-	-	-
Other or do not wish to provide or unknown	-0.01	0.02	-0.01	-0.04	0.06*	-0.02
Disability status						
Yes—hearing, mobility, vision, or other disability (ref.)	0.01	-0.05*	0.04	-0.01	-0.00	0.02
No—no disability (ref.)	-	-	-	-	-	-
Do not wish to provide	-0.01	0.09**	-0.08*	-0.06	0.01	0.05
Directorate						
BIO	0.04***	-0.05***	0.02	0.05***	-0.01	-0.04***
CISE	-0.01	0.01	-0.00	0.04*	0.01	-0.05***
EHR	-0.01	0.05**	-0.04*	-0.07***	0.07***	-0.00
ENG	-0.02**	0.05***	-0.03	0.02	0.01	-0.04**
GEO	0.07***	-0.06***	-0.01	0.05**	-0.03**	-0.02
MPS (ref.)	-	-	-	-	-	-
SBE	0.10***	-0.06***	-0.04*	0.01	0.02*	-0.04**
OD (OIA and OISE)	-0.00	0.01	-0.01	0.01	0.04	-0.05
Other	0.02	0.04	-0.06	-0.03	0.04*	-0.02

Characteristics	Time of day to conduct reviews response options - During normal working hours (AME)	Time of day to conduct reviews response options - Outside of normal working hours (AME)	Time of day to conduct reviews response options - Both during and outside normal working hours (AME)	Employer's view of conducting reviews response options - Within scope of normal work duties (AME)	Employer's view of conducting reviews response options - Outside scope of normal work duties (AME)	Employer's view of conducting reviews response options - I don't know (AME)
Institution type						
R1/MSI	-0.01	-0.00	0.01	0.02	0.00	-0.02
R1/non-MSI (ref.)	-	-	-	-	-	-
Non-R1/MSI	-0.01	0.04*	-0.02	-0.11***	0.09***	0.02
Non-R1/non-MSI	-0.01	0.02*	-0.02	-0.05***	0.05***	-0.00
For-profit organization	-0.03*	0.15***	-0.12***	-0.32***	0.31***	0.01
Government (local, state, federal, or tribal)	0.10**	-0.04	-0.06	-0.06	0.06*	-0.00
Nonprofit organization	-0.01	0.04	-0.03	-0.11***	0.14***	-0.03
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.01	0.02	-0.03	-0.09***	0.06***	0.03
Early career (degree ≤10 years ago)						
Early career	0.05***	-0.03**	-0.03*	-0.08***	0.02*	0.06***
Not early career (ref.)	-	-	-	-	-	-
Is respondent both investigator and reviewer						
Yes	0.00	-0.01	0.01	0.03**	-0.03***	0.01
No (ref.)	-	-	-	-	-	-

Note: time of day $N = 11,731$; employer's view $N = 11,739$

Multinomial logistic regression

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3.5. Declined to serve as a reviewer for NSF

Characteristics	Declined to serve as a reviewer for NSF (AME)
Race and ethnicity	
Asian, non-Hispanic	-0.11***
Black or African American, non-Hispanic	0.04
Hispanic	0.01
White, non-Hispanic (ref.)	-
Multiple races, non-Hispanic	0.11**
Other, NHPI, AIAN	-0.07
Do not wish to provide or unknown	-0.11***
Gender	
Female	0.13***
Male (ref.)	-
Other or do not wish to provide or unknown	0.09*
Disability status	
Yes—hearing, mobility, vision, or other disability	0.03
No—no disability (ref.)	-
Do not wish to provide	0.05
Directorate	
BIO	0.06**
CISE	0.12***
EHR	0.10***
ENG	0.07***
GEO	0.10***
MPS (ref.)	-
SBE	0.05*
OD (OIA and OISE)	-0.07
Other	0.06
Institution type	
R1/MSI	-0.04
R1/non-MSI (ref.)	-
Non-R1/MSI	-0.07**
Non-R1/non-MSI	-0.04**
For-profit organization	-0.11**
Government (local, state, federal, or tribal)	-0.11*
Nonprofit organization	0.03
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	-0.07**

Characteristics	Declined to serve as a reviewer for NSF (AME)
Early career (degree ≤10 years ago)	
Early career	-0.07***
Not early career (ref.)	-
Is respondent both investigator and reviewer	
Yes	-0.03*
No (ref.)	-

Note: $N = 6,996$

Logistic regression

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences
A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3.6.A. Reasons for declining to review an NSF proposal: Lack of time and too many NSF review requests

Characteristics	Lack of time response options - Not at all (AME)	Lack of time response options - To a small extent (AME)	Lack of time response options - To a moderate extent (AME)	To a great extent (AME)	Too many NSF review requests Response - Options Not at all (AME)	Too many NSF review requests Response - To a small extent (AME)	Too many NSF review requests Response - To a moderate extent (AME)	Too many NSF review requests Response - To a great extent (AME)
Race and ethnicity								
Asian, non-Hispanic	0.08***	0.01	0.01	-0.11***	0.01	-0.04	0.02	0.00
Black or African American, non-Hispanic	0.06	-0.04*	-0.02	-0.00	0.05	-0.08*	0.01	0.01
Hispanic	0.00	-0.01	-0.02	0.03	0.02	-0.05	0.03	-0.00
White, non-Hispanic (ref.)	-	-	-	-	-	-	-	-
Multiple races, non-Hispanic	-0.01	-0.01	-0.10**	0.11*	0.06	-0.08	0.04	-0.03**
Other, NHPI, AIAN	0.04	0.03	-0.00	-0.07	0.03	-0.03	0.00	0.01
Do not wish to provide or unknown	0.05	0.05	0.03	-0.13**	0.06	-0.04	-0.02	0.00
Gender								
Female	-0.05***	-0.02*	-0.03*	0.09***	-0.02	-0.01	0.01	0.02***
Male (ref.)	-	-	-	-	-	-	-	-
Other or do not wish to provide or unknown	-0.06*	0.01	0.00	0.05	-0.06	-0.03	0.08*	0.00
Disability status								
Yes—hearing, mobility, vision, or other disability	-0.01	0.02	0.00	-0.01	0.03	-0.01	-0.00	-0.01
No—no disability (ref.)	-	-	-	-	-	-	-	-
Do not wish to provide	0.01	-0.05*	0.06	-0.02	0.01	-0.03	0.01	0.01
Directorate								
BIO	-0.05**	0.02	-0.01	0.05	0.05	-0.01	-0.01	-0.02*
CISE	-0.06**	0.00	0.02	0.04	-0.09**	0.04	0.02	0.02
EHR	-0.06**	0.02	-0.01	0.05	0.07**	-0.04	-0.02	-0.01
ENG	0.00	0.02	0.01	-0.03	0.04	-0.04	0.01	-0.01
GEO	-0.01	-0.01	-0.03	0.05	-0.03	0.01	0.01	0.01
MPS (ref.)	-	-	-	-	-	-	-	-
SBE	-0.03	-0.00	-0.00	0.03	0.09**	-0.02	-0.04*	-0.03**
OD (OIA and OISE)	-0.08	-0.01	-0.07	0.16	0.01	-0.10	0.08	0.01

Characteristics	Lack of time response options - Not at all (AME)	Lack of time response options - To a small extent (AME)	Lack of time response options - To a moderate extent (AME)	To a great extent (AME)	Too many NSF review requests Response - Options Not at all (AME)	Too many NSF review requests Response - To a small extent (AME)	Too many NSF review requests Response - To a moderate extent (AME)	Too many NSF review requests Response - To a great extent (AME)
Other	0.12*	-0.01	-0.06	-0.05	0.17***	-	-0.07**	0.03
Institution type								
R1/MSI	0.04	0.00	-0.03	-0.01	-0.01	-0.02	0.01	0.02
R1/non-MSI (ref.)	-	-	-	-	-	-	-	-
Non-R1/MSI	0.01	-0.01	-0.05**	0.05	0.10**	-0.05	-0.04*	-0.01
Non-R1/non-MSI	0.02	-0.01	0.01	-0.02	0.04*	-0.02	-0.01	-0.00
For-profit organization	0.09*	-0.02	0.05	-0.12*	0.18***	-0.10**	-0.09***	-0.00
Government (local, state, federal, or tribal)	0.06	0.03	-0.01	-0.08	0.05	0.03	-0.06*	-0.02
Nonprofit organization	0.02	0.02	0.03	-0.07	0.08	-0.07	0.00	-0.01
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.02	0.02	0.03	-0.07*	0.03	0.02	-0.03	-0.01
Early career (degree ≤10 years ago)								
Early career	0.06***	-0.00	-0.04**	-0.03	0.10***	-0.05**	-0.04***	-0.02**
Not early career (ref.)	-	-	-	-	-	-	-	-
Is respondent both investigator and reviewer								
Yes	0.03**	-0.00	-0.03*	-0.01	-0.05**	0.03*	0.01	0.02**
No (ref.)	-	-	-	-	-	-	-	-

Note: Lack of time $N = 4,558$; too many NSF requests $N = 4,404$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

Question was asked only of reviewers who indicated they had declined to review since October 1, 2018

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3.6.B. Reasons for declining to review an NSF proposal: Competing professional pressures and commitments to other federal agencies

Characteristics	Competing professional pressures response options - Not at all (AME)	Competing professional pressures response options - To a small extent (AME)	Competing professional pressures response options - To a moderate extent (AME)	Competing professional pressures response options - To a great extent (AME)	Commitments to other federal agencies response options - Not at all (AME)	Commitments to other federal agencies response options - To a small extent (AME)	Commitments to other federal agencies response options - To a moderate extent (AME)	Commitments to other federal agencies response options - To a great extent (AME)
Race and ethnicity								
Asian, non-Hispanic	0.08***	0.03*	0.03	-0.14***	-0.04	0.02	0.02	0.01
Black or African American, non-Hispanic	0.01	-0.04*	0.03	-0.00	0.05	-0.05*	-0.04	0.04
Hispanic	0.02	-0.01	-0.05	0.04	-0.03	-0.00	0.00	0.03
White, non-Hispanic (ref.)	-	-	-	-	-	-	-	-
Multiple races, non-Hispanic	-0.05	-0.01	0.04	0.02	0.04	-0.01	-0.01	-0.02
Other, NHPI, AIAN	0.06	-0.07**	-0.01	0.03	0.02	-0.00	0.01	-0.02
Do not wish to provide or unknown	0.08*	0.04	-0.05	-0.06	0.03	-0.04	0.01	-0.01
Gender								
Female	-0.05***	-0.02*	-0.04**	0.12***	0.01	-0.01	-0.00	-0.00
Male (ref.)	-	-	-	-	-	-	-	-
Other or do not wish to provide or unknown	-0.07*	0.01	-0.00	0.06	-0.06	0.01	-0.04	0.09
Disability status								
Yes—hearing, mobility, vision, or other disability	-0.00	-0.00	-0.00	0.01	0.03	-0.03	-0.01	0.01
No—no disability (ref.)	-	-	-	-	-	-	-	-
Do not wish to provide	-0.01	-0.03	0.04	0.00	-0.06	0.08	-0.01	-0.02
Directorate								
BIO	-0.06**	-0.01	-0.01	0.08**	-0.08**	0.05**	0.03*	-0.01
CISE	-0.05*	0.03	-0.03	0.04	0.02	0.04*	-0.04**	-0.02
EHR	-0.04	0.00	0.03	-0.00	0.13***	-0.03	-0.06***	-0.04***
ENG	0.03	0.02	-0.01	-0.04	0.01	0.01	0.01	-0.03*
GEO	0.01	0.01	-0.07**	0.04	-0.10***	0.03	0.07***	0.00
MPS (ref.)	-	-	-	-	-	-	-	-
SBE	-0.01	-0.00	0.01	0.00	0.04	0.01	-0.02	-0.03*
OD (OIA and OISE)	-0.01	-0.05	0.13	-0.07	-0.14	0.09	0.07	-0.02

Characteristics	Competing professional pressures response options - Not at all (AME)	Competing professional pressures response options - To a small extent (AME)	Competing professional pressures response options - To a moderate extent (AME)	Competing professional pressures response options - To a great extent (AME)	Commitments to other federal agencies response options - Not at all (AME)	Commitments to other federal agencies response options - To a small extent (AME)	Commitments to other federal agencies response options - To a moderate extent (AME)	Commitments to other federal agencies response options - To a great extent (AME)
Other	0.03	0.03	0.02	-0.08	0.10*	-0.06*	0.00	-0.04*
Institution type								
R1/MSI	0.03	0.03	-0.02	-0.05	0.05	-0.04	0.00	-0.02
R1/non-MSI (ref.)	-	-	-	-	-	-	-	-
Non-R1/MSI	0.07**	0.02	-0.01	-0.08*	0.05	-0.01	-0.05**	0.01
Non-R1/non-MSI	0.02	0.02*	0.00	-0.04*	0.06***	-0.01	-0.04***	-0.01
For-profit organization	0.16***	0.06	-0.02	-0.20***	0.16***	-0.06	-0.08***	-0.02
Government (local, state, federal, or tribal)	0.11	0.05	0.12	-0.28***	0.14**	-0.05	-0.09***	0.00
Nonprofit organization	0.11*	0.04	-0.02	-0.13**	0.08*	-0.02	-0.05*	-0.00
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.06*	0.08***	0.00	-0.14***	-0.01	0.03	-0.01	-0.01
Early career (degree ≤10 years ago)								
Early career	0.10***	0.02	-0.02	-0.10***	0.10***	-0.04*	-0.03**	-0.03***
Not early career (ref.)	-	-	-	-	-	-	-	-
Is respondent both investigator and reviewer								
Yes	0.03*	-0.00	0.01	-0.04*	0.03	-0.01	-0.01	-0.02*
No (ref.)	-	-	-	-	-	-	-	-

Note: competing professional pressure $N = 4,561$; commitments to other federal agencies $N = 4,545$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

Question was asked only of reviewers who indicated they had declined to review since October 1, 2018,

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3.6.C. Reasons for declining to review an NSF Proposal: Inability to travel

Characteristics	Inability to travel response options - Not at all (AME)	Inability to travel response options - To a small extent (AME)	Inability to travel response options - To a moderate extent (AME)	Inability to travel response options - To a great extent (AME)
Race and ethnicity				
Asian, non-Hispanic	-0.03	0.04	0.00	-0.01
Black or African American, non-Hispanic	0.20***	-0.10***	-0.00	-0.10*
Hispanic	0.02	0.02	-0.04	0.00
White, non-Hispanic (ref.)	-	-	-	-
Multiple races, non-Hispanic	-0.05	0.04	0.03	-0.02
Other, NHPI, AIAN	-0.09	0.09	0.06	-0.06
Do not wish to provide or unknown	0.01	-0.04	-0.01	0.03
Gender				
Female	-0.03	-0.00	0.02	0.01
Male (ref.)	-	-	-	-
Other or do not wish to provide or unknown	-0.09	-0.04	0.17*	-0.04
Disability status				
Yes—hearing, mobility, vision, or other disability	-0.06	-0.05	-0.03	0.14*
No—no disability (ref.)	-	-	-	-
Do not wish to provide	0.01	0.00	-0.05	0.03
Directorate				
BIO	0.04	0.05	0.01	-0.09**
CISE	0.01	0.02	0.03	-0.06
EHR	0.09*	0.01	0.01	-0.11**
ENG	0.04	-0.00	0.01	-0.05
GEO	0.03	-0.07**	0.06	-0.02
MPS (ref.)	-	-	-	-
SBE	0.08	0.02	-0.03	-0.08
OD (OIA and OISE)	-0.08	0.20	0.08	-0.19**
Other	0.18*	-0.03	0.03	-0.19***
Institution type				
R1/MSI	0.00	-0.05	0.00	0.05
R1/non-MSI (ref.)	-	-	-	-
Non-R1/MSI	-0.05	0.03	-0.02	0.03
Non-R1/non-MSI	0.00	0.01	-0.01	-0.00
For-profit organization	-0.06	0.10	0.03	-0.07
Government (local, state, federal, or tribal)	0.23*	-0.01	-0.05	-0.17**
Nonprofit organization	0.09	-0.10**	-0.05	0.06
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	-0.01	0.04	-0.02	-0.02
Early career (degree ≤10 years ago)				
Early career	0.01	-0.01	-0.04	0.03

Characteristics	Inability to travel response options - Not at all (AME)	Inability to travel response options - To a small extent (AME)	Inability to travel response options - To a moderate extent (AME)	Inability to travel response options - To a great extent (AME)
Not early career (ref.)	-	-	-	-
Is respondent both investigator and reviewer				
Yes	0.04	0.01	-0.03	-0.02
No (ref.)	-	-	-	-

Note: $N = 2,120$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences
A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3.7. Hours conducting most recent review—2021 and 2019 differences

Characteristics	Coefficient
Year	
2021	-0.04***
2019 (ref.)	-
Race and ethnicity	
Year * Asian, non-Hispanic	0.03
Year * Black or African American, non-Hispanic	0.19**
Year * Hispanic	-0.12**
Year * White, non-Hispanic (ref.)	-
Year * Multiple races, non-Hispanic	0.15*
Year * Other, NHPI, AIAN	0.21
Year * Do not wish to provide or unknown	-0.01
Gender	
Year * Female	-0.04*
Year * Male (ref.)	-
Year * Other or do not wish to provide or unknown	0.02
Institution type	
Year * Academic (ref.)	-
Year * For-profit organization	-0.06
Year * Government (local, state, federal, or tribal)	-0.01
Year * Federally funded R&D	-0.08
Year * Nonprofit organization	-0.05
Year * Other, primary institution, secondary institution	0.13*

Characteristics	Coefficient
Early career (degree ≤10 years ago)	
Year * Early career	-0.01
Year * Not early career (ref.)	-

Note: $N = 29,895$

Linear regression

Findings are presented in coefficients.

Findings are weighted to adjust for nonresponse bias.

Dependent variable is log transformed.

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

See appendix B.4 for the full list of independent variables included in the model.

AIAN = American Indian or Alaska Native; HSP = Hispanic; NHPI = Native Hawaiian or Other Pacific Islander; R&D = research and development

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3.8.A. Time-of-day respondents conducted reviews and employer's view of participating as a reviewer—2021 and 2019 differences

Characteristics	Time of Day - During normal hours odds ratios	Time of Day - Both during and outside normal hours odds ratios	Employer's View - Outside scope odds ratios	Employer's View - I don't know odds ratios
Year				
2021	0.71***	0.96	1.94***	0.74***
2019 (ref.)	-	-	-	-
Race and ethnicity				
Year * Asian, non-Hispanic	0.75*	1.07	1.15	1.39***
Year * Black or African American, non-Hispanic	0.80	0.77	1.11	1.22
Year * Hispanic	0.72	0.82	1.30	0.95
Year * White, non-Hispanic (ref.)	-	-	-	-
Year * Multiple races, non-Hispanic	0.58	0.56*	0.98	0.96
Year * Other, NHPI, AIAN	1.48	0.65	2.20	1.27
Year * Do not wish to provide or unknown	0.91	1.21	0.74	1.65***
Gender				
Year * Female	1.34**	1.08	0.92	1.00
Year * Male (ref.)	-	-	-	-
Year * Other or do not wish to provide or unknown	1.04	1.15	1.09	0.68
Institution type				
Year * Academic (ref.)	-	-	-	-
Year * For-profit organization	1.24	0.86	0.64*	1.20
Year * Government (local, state, federal, or tribal)	1.49	0.84	0.81	1.10
Year * Federally funded R&D	1.55	1.43	0.52	1.12
Year * Nonprofit organization	0.82	0.87	0.85	0.88
Year * Other, primary institution, secondary institution	1.13	1.08	0.72	0.94

Characteristics	Time of Day - During normal hours odds ratios	Time of Day - Both during and outside normal hours odds ratios	Employer's View - Outside scope odds ratios	Employer's View - I don't know odds ratios
Early career (degree ≤10 years ago)				
Year * Early career	1.37**	1.07	0.97	1.15
Year * Not early career (ref.)	-	-	-	-

Note: time of day $N = 30,019$; employer's view $N = 29,990$

Multinomial logistic regression

Time of day has three nominal outcomes: 1 = during normal working hours; 2 = outside normal working hours (reference); 3 = both during and outside normal working hours.

Employers' view has three nominal outcomes: 1 = within the scope of my normal working duties (reference); 2 = outside the scope of my normal working duties; 3 = I don't know.

Findings are presented in relative risks ratios.

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

Findings are weighted to adjust for nonresponse bias.

See appendix B.4 for the full list of independent variables included in the model.

AIAN = American Indian or Alaska Native; HSP = Hispanic; NHPI = Native Hawaiian or Other Pacific Islander; R&D = research and development

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3.8.B. Time-of-day respondents conducted reviews and employer's view of participating as a reviewer – 2021 and 2019 differences (average marginal effects for 2021)

Characteristics	Time of day to conduct reviews response options - During normal working hours (AME)	Time of day to conduct reviews response options - Outside normal working hours (AME)	Time of day to conduct reviews response options - Both during and outside normal working hours (AME)	Employer's view of conducting reviews response options - Within scope of normal work duties (AME)	Employer's view of conducting reviews response options - Outside scope of normal work duties (AME)	Employer's view of conducting reviews response options - I don't know (AME)
Year						
2021	-0.02***	0.01	0.02**	-0.02**	0.05***	-0.04***
2019 (ref.)	-	-	-	-	-	-

Note: time of day $N = 30,019$; employer's view $N = 29,990$

Multinomial logistic regression

Findings presented using Average Marginal Effects (AME) for the year effect of "2021" modeled in table C.3.8.A.

Findings are weighted to adjust for nonresponse bias.

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

See appendix B.4 for the full list of independent variables included in the model.

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3.9. Declined to serve as a reviewer for NSF—2021 and 2019 differences

Characteristics	Odds ratio
Year	
2021	1.57***
2021 AME = 0.10***	-
2019 (ref.)	-
Race and ethnicity	
Year * Asian, non-Hispanic	0.82**
Year * Black or African American, non-Hispanic	1.21
Year * Hispanic	0.99
Year * White, non-Hispanic (ref.)	-
Year * Multiple races, non-Hispanic	1.64*
Year * Other, NHPI, AIAN	0.36**
Year * Do not wish to provide or unknown	0.72**
Gender	
Year * Female	1.19**
Year * Male (ref.)	-
Year * Other or do not wish to provide or unknown	1.33
Institution type	
Year * Academic (ref.)	-
Year * For-profit organization	0.81
Year * Government (local, state, federal, or tribal)	1.02
Year * Federally funded R&D	0.92
Year * Nonprofit organization	1.22
Year * Other, primary institution, secondary institution	1.27
Early career (degree ≤10 years ago)	
Year * Early career	0.92
Year * Not early career (ref.)	-

Note: $N = 30,108$

Logistic regression

Findings are presented using odds ratios.

Findings are weighted to adjust for nonresponse bias.

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

See appendix B.4 for the full list of independent variables included in the model.

AIAN = American Indian or Alaska Native; AME = Average Marginal Effects; HSP = Hispanic; NHPI = Native Hawaiian or Other Pacific Islander; R&D = research and development

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C.4 Quality: Regression Tables

Regression coefficients for figures 4.4.1 to 4.4.10 in chapter 4.3 (*Proposal and Review Quality in the Merit Review Process*) are presented in tables C.4.1 to C.4.6.

Table C.4.1 Applicant satisfaction with the quality of information provided by NSF

Characteristics	Very dissatisfied (AME)	Somewhat dissatisfied (AME)	Neither dissatisfied nor satisfied (AME)	Somewhat satisfied (AME)	Very satisfied (AME)
Race and ethnicity					
Asian, non-Hispanic	0.00	-0.02*	-0.04***	-0.04***	0.09***
Black or African American, non-Hispanic	0.02	-0.01	-0.00	-0.05*	0.04
Hispanic	0.01	0.01	0.01	-0.02	-0.01
White, non-Hispanic (ref.)	-	-	-	-	-
Multiple races, non-Hispanic	0.01	0.02	-0.02	0.02	-0.03
Other, NHPI, AIAN	0.03	0.05	0.05	-0.08*	-0.05
Do not wish to provide or unknown	0.00	0.02	0.00	-0.03	0.01
Gender					
Female	-0.00	0.00	0.01	0.01	-0.02
Male (ref.)	-	-	-	-	-
Other or do not wish to provide or unknown	0.03	0.02	0.04	-0.02	-0.07*
Disability status					
Yes—hearing, mobility, vision, or other disability	0.03*	0.01	0.03	0.03	-0.09***
No—no disability (ref.)	-	-	-	-	-
Do not wish to provide	0.01	0.01	0.04	0.01	-0.08**
Directorate					
BIO	-0.01	0.01	-0.01	0.00	0.00
CISE	-0.00	0.00	-0.00	0.01	-0.01
EHR	-0.01	-0.00	-0.02	0.03	0.01
ENG	0.00	-0.01	0.00	-0.01	0.02
GEO	-0.01	-0.00	-0.02	0.00	0.03
MPS (ref.)	-	-	-	-	-

Characteristics	Very dissatisfied (AME)	Somewhat dissatisfied (AME)	Neither dissatisfied nor satisfied (AME)	Somewhat satisfied (AME)	Very satisfied (AME)
SBE	0.00	0.04**	-0.01	0.01	-0.03
OD (OIA and OISE)	-0.01	0.00	0.05	-0.00	-0.04
Other	0.01	0.02	-0.01	0.03	-0.05
Institution type					
R1/MSI	0.02*	-0.00	-0.06***	0.04	-0.00
R1/non-MSI (ref.)	-	-	-	-	-
Non-R1/MSI	-0.01	-0.02	-0.04**	0.01	0.06**
Non-R1/non-MSI	0.00	-0.00	-0.02*	0.03**	-0.01
For-profit organization	0.03**	0.05***	-0.06***	-0.02	-0.00
Government (local, state, federal, or tribal)	-0.01	0.04	-0.03	0.04	-0.04
Nonprofit organization	0.01	0.02	-0.01	0.03	-0.05
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.02	0.01	-0.02	0.00	-0.01
Early career (degree ≤10 years ago)					
Early career	-0.00	0.02***	0.00	0.02*	-0.05***
Not early career (ref.)	-	-	-	-	-
Is respondent both investigator and reviewer					
Yes	-0.02***	-0.03***	-0.02*	-0.01	0.08***
No (ref.)	-	-	-	-	-

Note: N = 12,299

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.4.2.A. Applicant agreement that written reviews are thorough and technically sound

Characteristics	Written reviews are thorough - Strongly disagree (AME)	Written reviews are thorough - Disagree (AME)	Written reviews are thorough - Neither agree nor disagree (AME)	Written reviews are thorough - Agree (AME)	Written reviews are thorough - Strongly agree (AME)	Written reviews are technically sound - Strongly disagree (AME)	Written reviews are technically sound - Disagree (AME)	Written reviews are technically sound - Neither agree nor disagree (AME)	Written reviews are technically sound - Agree (AME)	Written reviews are technically sound - Strongly agree (AME)
Race and ethnicity										
Asian, non-Hispanic	0.00	-0.02*	-0.03**	0.03*	0.02*	0.01	0.01	-0.04***	-0.00	0.02***
Black or African American, non-Hispanic	-0.00	0.02	-0.02	-0.02	0.03*	0.02	-0.01	-0.02	-0.02	0.02
Hispanic	-0.01	0.02	0.00	-0.02	0.01	-0.01	0.02	0.00	-0.02	0.01
White, non-Hispanic (ref.)	-	-	-	-	-	-	-	-	-	-
Multiple races, non-Hispanic	-0.01	-0.00	-0.00	0.03	-0.01	-0.00	0.01	0.05	-0.04	-0.02
Other, NHPI, AIAN	0.10***	0.06	0.03	-0.18***	-0.00	0.09***	0.05	0.04	-0.18***	0.00
Do not wish to provide or unknown	0.03*	0.03	0.01	-0.06**	-0.01	0.03*	0.05**	0.05*	-0.12***	-0.00
Gender										
Female	-0.02***	-0.02*	-0.00	0.03**	0.02**	-0.02***	-0.02**	0.00	0.03**	0.01
Male (ref.)	-	-	-	-	-	-	-	-	-	-
Other or do not wish to provide or unknown	0.02	0.04	-0.01	-0.06	0.01	0.02	0.04	-0.02	-0.02	-0.01
Disability status										
Yes—hearing, mobility, vision, or other disability	0.04*	-0.01	-0.01	-0.01	-0.01	0.06**	0.02	0.00	-0.06*	-0.02
No—no disability (ref.)	-	-	-	-	-	-	-	-	-	-
Do not wish to provide	0.07***	0.01	0.01	-0.07*	-0.02*	0.05**	0.04	0.02	-0.09**	-0.02
Directorate										
BIO	-0.03***	-0.04**	-0.04**	0.09***	0.02*	-0.03**	-0.00	-0.03	0.05**	0.01
CISE	0.01	0.01	-0.03	0.00	0.00	0.00	0.01	-0.00	0.00	-0.01
EHR	-0.05***	-0.04**	-0.04*	0.10***	0.02*	-0.05***	-0.01	-0.01	0.06***	0.01
ENG	0.01	0.01	-0.01	-0.01	-0.01	0.01	0.02*	0.01	-0.03*	-0.01*
GEO	-0.04***	-0.06***	0.02	0.08***	0.00	-0.04***	-0.03*	0.04*	0.04**	-0.00
MPS (ref.)	-	-	-	-	-	-	-	-	-	-

Characteristics	Written reviews are thorough - Strongly disagree (AME)	Written reviews are thorough - Disagree (AME)	Written reviews are thorough - Neither agree nor disagree (AME)	Written reviews are thorough - Agree (AME)	Written reviews are thorough - Strongly agree (AME)	Written reviews are technically sound - Strongly disagree (AME)	Written reviews are technically sound - Disagree (AME)	Written reviews are technically sound - Neither agree nor disagree (AME)	Written reviews are technically sound - Agree (AME)	Written reviews are technically sound - Strongly agree (AME)
SBE	-0.01	-0.05***	-0.05**	0.08***	0.03**	-0.02*	-0.03*	-0.02	0.05**	0.02
OD (OIA and OISE)	-0.03	0.08	-0.06	-0.02	0.04	-0.05*	0.04	0.04	-0.04	0.01
Other	0.05*	-0.01	-0.00	-0.03	0.00	0.05**	-0.01	0.03	-0.07**	0.00
Institution type										
R1/MSI	0.00	0.01	0.01	-0.01	-0.01	0.03*	-0.01	-0.02	0.01	-0.01
R1/non-MSI (ref.)	-	-	-	-	-	-	-	-	-	-
Non-R1/MSI	-0.02	-0.02	-0.05***	0.06***	0.03**	-0.02	-0.00	-0.02	0.01	0.03***
Non-R1/non-MSI	-0.00	-0.01	-0.03***	0.03*	0.02**	-0.00	-0.02*	-0.04***	0.04***	0.01*
For-profit organization	0.05***	0.00	-0.06***	-0.00	0.02	0.05***	0.03	-0.09***	-0.01	0.02
Government (local, state, federal, or tribal)	0.03	0.01	-0.07	-0.01	0.04	-0.01	-0.05	-0.03	0.01	0.08
Nonprofit organization	-0.01	-0.01	-0.02	0.06*	-0.01	-0.00	-0.02	0.02	0.02	-0.01
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.04*	-0.04*	-0.02	-0.02	0.04*	0.04*	-0.01	-0.04	-0.02	0.02
Early career (degree ≤10 years ago)										
Yes	-0.03***	-0.02*	-0.05***	0.07***	0.03***	-0.03***	0.03***	-0.03***	0.07***	0.03***
No (ref.)	-	-	-	-	-	-	-	-	-	-

Characteristics	Written reviews are thorough - Strongly disagree (AME)	Written reviews are thorough - Disagree (AME)	Written reviews are thorough - Neither agree nor disagree (AME)	Written reviews are thorough - Agree (AME)	Written reviews are thorough - Strongly agree (AME)	Written reviews are technically sound - Strongly disagree (AME)	Written reviews are technically sound - Disagree (AME)	Written reviews are technically sound - Neither agree nor disagree (AME)	Written reviews are technically sound - Agree (AME)	Written reviews are technically sound - Strongly agree (AME)
Experience as applicant and reviewer										
Yes	-0.06***	-0.03***	0.03***	0.07***	-0.01**	-0.05***	0.05***	0.02*	0.08***	-0.00
No (ref.)	-	-	-	-	-	-	-	-	-	-

Note: Written reviews are thorough $N = 13,614$; written reviews are technically sound $N = 13,580$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.4.2.B. Applicant agreement that written reviews are of high quality and panel summaries are of high quality

Characteristics	Written reviews are of high quality - Strongly disagree (AME)	Written reviews are of high quality - Disagree (AME)	Written reviews are of high quality - Neither agree nor disagree (AME)	Written reviews are of high quality - Agree (AME)	Written reviews are of high quality - Strongly agree (AME)	Panel summaries are of high quality - Strongly disagree (AME)	Panel summaries are of high quality - Disagree (AME)	Panel summaries are of high quality - Neither agree nor disagree (AME)	Panel summaries are of high quality - Agree (AME)	Panel summaries are of high quality - Strongly agree (AME)
Race and ethnicity										
Asian, non-Hispanic	0.01	-0.01	-0.02	-0.00	0.02***	0.01	-0.00	-0.02	-0.01	0.02**
Black or African American, non-Hispanic	0.00	0.01	-0.02	-0.02	0.03*	0.00	-0.02	0.02	-0.04	0.03
Hispanic	-0.01	0.01	0.01	-0.03	0.01	-0.01	0.00	0.01	-0.02	0.02
White, non-Hispanic (ref.)	-	-	-	-	-	-	-	-	-	-
Multiple races, non-Hispanic	-0.01	-0.03	0.05	-0.00	-0.01	0.00	-0.01	-0.02	0.05	-0.02
Other, NHPI, AIAN	0.08***	0.08**	0.00	-	-0.00	0.07**	0.10**	0.01	-0.19***	0.00

Characteristics	Written reviews are of high quality - Strongly disagree (AME)	Written reviews are of high quality - Disagree (AME)	Written reviews are of high quality - Neither agree nor disagree (AME)	Written reviews are of high quality - Agree (AME)	Written reviews are of high quality - Strongly agree (AME)	Panel summaries are of high quality - Strongly disagree (AME)	Panel summaries are of high quality - Disagree (AME)	Panel summaries are of high quality - Neither agree nor disagree (AME)	Panel summaries are of high quality - Agree (AME)	Panel summaries are of high quality - Strongly agree (AME)
Do not wish to provide or unknown	0.03*	0.07***	-0.01	0.07***	-0.01	0.01	0.06***	0.04*	-0.09***	-0.02
Gender										
Female	0.03***	0.03***	0.01	0.03***	0.01*	-0.03***	-0.02***	-0.01	0.04***	0.02**
Male (ref.)	-	-	-	-	-	-	-	-	-	-
Other or do not wish to provide or unknown	0.01	0.03	0.01	-0.04	-0.01	0.03*	0.01	0.00	-0.01	-0.03*
Disability status										
Yes—hearing, mobility, vision, or other disability	0.07***	-0.00	-0.01	-0.04	-0.01	0.05**	0.05*	-0.03	-0.07**	0.01
No—no disability (ref.)	-	-	-	-	-	-	-	-	-	-
Do not wish to provide	0.05**	0.02	0.02	-0.07**	-0.02	0.04*	0.03	0.03	-0.09**	-0.02
Directorate										
BIO	0.03***	-0.01	-0.02	0.05***	0.01	-0.02**	-0.01	-0.06***	0.06***	0.03***
CISE	-0.00	0.02	-0.01	-0.01	-0.01	-0.01	0.01	-0.02	0.00	0.02*
EHR	0.03***	-0.04**	-0.02	0.07***	0.02*	-0.04***	-0.04***	-0.07***	0.10***	0.05***
ENG	0.01	0.03*	0.00	-0.03*	-0.01*	0.00	0.02*	-0.00	-0.01	-0.01
GEO	0.04***	0.05***	0.04*	0.06***	-0.01	-0.03***	-0.01	-0.04**	0.07***	0.02
MPS (ref.)	-	-	-	-	-	-	-	-	-	-
SBE	-0.02*	-0.03*	-0.05**	0.07***	0.03**	-0.01	-0.01	-0.03	0.02	0.02*
OD (OIA and OISE)	-0.05*	-0.01	0.02	0.01	0.03	-0.03	-0.01	0.07	-0.00	-0.03
Other	0.04*	0.02	-0.02	-0.04	0.00	0.03	0.00	-0.01	-0.02	0.00
Institution type										
R1/MSI	0.02	-0.01	0.01	-0.01	-0.01	0.00	0.02	0.00	-0.01	-0.02
R1/non-MSI (ref.)	-	-	-	-	-	-	-	-	-	-
Non-R1/MSI	-0.01	-0.02	-0.05**	0.05**	0.02**	-0.01	-0.02	-0.03*	0.03	0.04***
Non-R1/non-MSI	-0.00	-0.02*	-0.02*	0.03**	0.01**	-0.00	-0.00	-0.03**	0.01	0.02***

Characteristics	Written reviews are of high quality - Strongly disagree (AME)	Written reviews are of high quality - Disagree (AME)	Written reviews are of high quality - Neither agree nor disagree (AME)	Written reviews are of high quality - Agree (AME)	Written reviews are of high quality - Strongly agree (AME)	Panel summaries are of high quality - Strongly disagree (AME)	Panel summaries are of high quality - Disagree (AME)	Panel summaries are of high quality - Neither agree nor disagree (AME)	Panel summaries are of high quality - Agree (AME)	Panel summaries are of high quality - Strongly agree (AME)
For-profit organization	0.05***	0.00	-	0.00	0.02*	0.05***	0.05***	-0.07***	-0.04*	0.01
Government (local, state, federal, or tribal)	0.02	-0.02	-0.04	-0.04	0.08	0.03	0.06	-0.16**	0.01	0.06
Nonprofit organization	-0.00	-0.01	0.00	0.02	-0.01	-0.00	-0.01	-0.00	0.02	-0.00
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.03*	-0.03	-0.04	0.01	0.03*	0.04*	-0.01	-0.03	-0.00	0.01
Early career (degree ≤10 years ago)										
Yes	0.03***	0.03***	0.03***	0.07***	0.02***	-0.03***	-0.04***	-0.05***	0.07***	0.05***
No (ref.)	-	-	-	-	-	-	-	-	-	-
Experience as applicant and reviewer										
Yes	0.05***	0.06***	0.03**	0.08***	-0.00	-0.04***	-0.04***	-0.03**	0.10***	0.02**
No (ref.)	-	-	-	-	-	-	-	-	-	-

Note: Written reviews are of high quality $N = 13,590$; panel summaries are of high quality: $N = 13,368$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.4.2.C. Applicant agreement that information from NSF on the outcomes of the competition is of high quality

Characteristics	Information on the outcomes is of high quality - Strongly disagree (AME)	Information on the outcomes is of high quality - Disagree (AME)	Information on the outcomes is of high quality - Neither agree nor disagree (AME)	Information on the outcomes is of high quality - Agree (AME)	Information on the outcomes is of high quality - Strongly agree (AME)
Race and ethnicity					
Asian, non-Hispanic	-0.00	-0.04***	-0.03*	0.04***	0.03***
Black or African American, non-Hispanic	0.00	-0.02	0.00	-0.01	0.03*
Hispanic	0.00	0.01	0.01	-0.04*	0.02
White, non-Hispanic (ref.)	-	-	-	-	-
Multiple races, non-Hispanic	0.02	0.00	0.01	-0.02	-0.01
Other, NHPI, AIAN	0.08**	0.01	0.00	-0.11***	0.01
Do not wish to provide or unknown	0.02	0.02	0.01	-0.05**	-0.01
Gender					
Female	-0.02**	-0.00	0.00	0.01	0.01*
Male (ref.)	-	-	-	-	-
Other or do not wish to provide or unknown	0.02	0.04	-0.02	-0.05	0.00
Disability status					
Yes—hearing, mobility, vision, or other disability	0.04*	0.06*	-0.03	-0.06**	-0.00
No—no disability (ref.)	-	-	-	-	-
Do not wish to provide	0.08***	-0.01	-0.00	-0.06*	-0.01
Directorate					
BIO	-0.02	0.00	-0.02	0.02	0.01
CISE	0.00	0.04**	-0.01	-0.02	-0.01
EHR	-0.02*	-0.01	-0.03	0.04**	0.02
ENG	0.02	0.03**	0.01	-0.05***	-0.01
GEO	-0.05***	0.02	-0.00	0.03	0.00
MPS (ref.)	-	-	-	-	-
SBE	-0.00	0.01	-0.05**	0.03	0.01
OD (OIA and OISE)	0.01	-0.08*	0.07	0.02	-0.03
Other	0.05*	0.01	-0.01	-0.04	-0.00
Institution type					
R1/MSI	0.02	-0.01	-0.01	0.01	-0.02*
R1/non-MSI (ref.)	-	-	-	-	-
Non-R1/MSI	-0.03*	-0.02	-0.05**	0.06***	0.03**
Non-R1/non-MSI	-0.00	-0.01	-0.01	0.01	0.01
For-profit organization	0.06***	-0.01	-0.06**	-0.00	0.01
Government (local, state, federal, or tribal)	0.03	-0.02	-0.10	0.01	0.09
Nonprofit organization	-0.02	-0.06**	0.03	0.06*	-0.01

Characteristics	Information on the outcomes is of high quality - Strongly disagree (AME)	Information on the outcomes is of high quality - Disagree (AME)	Information on the outcomes is of high quality - Neither agree nor disagree (AME)	Information on the outcomes is of high quality - Agree (AME)	Information on the outcomes is of high quality - Strongly agree (AME)
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.01	-0.02	-0.04	0.04	0.01
Early career (degree ≤10 years ago)					
Yes	-0.03***	-0.03***	-0.01	0.04***	0.02***
No (ref.)	-	-	-	-	-
Experience as applicant and reviewer					
Yes	-0.04***	-0.03***	0.00	0.06***	0.01
No (ref.)	-	-	-	-	-

Note: $N = 13,162$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.4.3.A. Applicant views for the extent to which receiving a declined review improved understanding of the proposal process and provided useful information for improving a future proposal

Characteristics	Improved understanding of the proposal process - Not at all (AME)	Improved understanding of the proposal process - To a small extent (AME)	Improved understanding of the proposal process - To a moderate extent (AME)	Improved understanding of the proposal process - To a great extent (AME)	Provided useful information for improving future proposals - Not at all (AME)	Provided useful information for improving future proposals - To a small extent (AME)	Provided useful information for improving future proposals - To a moderate extent (AME)	Provided useful information for improving future proposals - To a great extent (AME)
Race and ethnicity								
Asian, non-Hispanic	-0.10***	-0.05***	0.07***	0.07***	-0.01	-0.04***	0.02	0.03**
Black or African American, non-Hispanic	-0.03	-0.07**	0.04	0.06***	0.04	-0.04	-0.04	0.05*
Hispanic	-0.04*	-0.02	0.03	0.04**	0.00	-0.02	-0.02	0.04*
White, non-Hispanic (ref.)	-	-	-	-	-	-	-	-
Multiple races, non-Hispanic	0.01	0.01	-0.05	0.03	-0.01	0.06	-0.06	0.00
Other, NHPI, AIAN	0.08*	-0.06	-0.03	0.01	0.11***	0.04	-0.08*	-0.06**
Do not wish to provide or unknown	-0.01	-0.00	0.01	0.00	0.06***	0.02	-0.07**	-0.02
Gender								
Female	0.02*	-0.01	-0.01	0.01	-0.00	-0.02*	-0.00	0.03***
Male (ref.)	-	-	-	-	-	-	-	-
Other or do not wish to provide or unknown	0.08**	0.01	-0.07**	-0.01	0.02	0.02	0.01	-0.05*
Disability status								
Yes—hearing, mobility, vision, or other disability	0.05	0.05	-0.08***	-0.02	0.04	0.06*	-0.04	-0.05**
No—no disability (ref.)	-	-	-	-	-	-	-	-
Do not wish to provide	0.10***	0.00	-0.09***	-0.01	0.07**	0.07*	-0.09**	-0.05**
Directorate								
BIO	-0.04**	-0.04*	0.07***	0.01	-0.07***	-0.06***	0.07***	0.05***
CISE	-0.01	0.01	0.01	-0.00	-0.02	0.01	0.02	-0.01
EHR	-0.03	-0.01	0.04*	0.01	-0.09***	-0.03	0.06***	0.06***
ENG	0.00	-0.01	0.01	-0.00	-0.03**	0.02	0.01	-0.00
GEO	0.01	-0.02	0.02	-0.01	-0.08***	-0.03*	0.09***	0.02

Characteristics	Improved understanding of the proposal process - Not at all (AME)	Improved understanding of the proposal process - To a small extent (AME)	Improved understanding of the proposal process - To a moderate extent (AME)	Improved understanding of the proposal process - To a great extent (AME)	Provided useful information for improving future proposals - Not at all (AME)	Provided useful information for improving future proposals - To a small extent (AME)	Provided useful information for improving future proposals - To a moderate extent (AME)	Provided useful information for improving future proposals - To a great extent (AME)
MPS (ref.)	-	-	-	-	-	-	-	-
SBE	-0.02	-0.06**	0.05**	0.03*	-0.04**	-0.08***	0.03	0.08***
OD (OIA and OISE)	0.02	0.08	-0.05	-0.05***	-0.01	0.01	0.07	-0.08**
Other	0.03	-0.01	-0.01	-0.01	0.00	0.03	-0.04	0.01
Institution type								
R1/MSI	0.01	-0.03	0.02	0.01	0.00	-0.01	0.01	0.00
R1/non-MSI (ref.)	-	-	-	-	-	-	-	-
Non-R1/MSI	-0.04*	-0.06**	0.05**	0.04***	-0.03**	-0.04*	0.01	0.05***
Non-R1/non-MSI	-0.02*	-0.00	0.02	0.01	-0.01	-0.04***	0.01	0.03***
For-profit organization	0.02	-0.06**	0.00	0.04*	0.10***	-0.03	-0.09***	0.03
Government (local, state, federal, or tribal)	0.09	-0.14	0.02	0.03	0.06	-0.13	-0.04	0.12
Nonprofit organization	-0.03	0.07*	-0.03	-0.01	0.00	-0.01	0.02	-0.02
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	-0.01	-0.02	-0.00	0.03	0.06**	-0.06*	-0.03	0.02
Early career (degree ≤10 years ago)								
Yes	-0.08***	-0.01	0.06***	0.03***	-0.05***	-0.03**	0.02	0.06***
No (ref.)	-	-	-	-	-	-	-	-

Characteristics	Improved understanding of the proposal process - Not at all (AME)	Improved understanding of the proposal process - To a small extent (AME)	Improved understanding of the proposal process - To a moderate extent (AME)	Improved understanding of the proposal process - To a great extent (AME)	Provided useful information for improving future proposals - Not at all (AME)	Provided useful information for improving future proposals - To a small extent (AME)	Provided useful information for improving future proposals - To a moderate extent (AME)	Provided useful information for improving future proposals - To a great extent (AME)
Experience as applicant and reviewer								
Yes	-0.04***	0.00	0.03**	0.01	-0.05***	-0.01	0.05***	0.01
No (ref.)	-	-	-	-	-	-	-	-

Note: improved understanding of proposal process $N = 12,386$; provided useful information for future proposals $N = 12,384$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.4.3.B. Applicant views for the extent to which receiving a declined review influenced the decision to submit to a non-NSF funding agency and discouraged revising and submitting a proposal to NSF

Characteristics	Influenced decision to submit to non-NSF funding - Not at all (AME)	Influenced decision to submit to non-NSF funding - To a small extent (AME)	Influenced decision to submit to non-NSF funding - To a moderate extent (AME)	Influenced decision to submit to non-NSF funding - To a great extent (AME)	Discouraged revising and submitting to NSF - Not at all (AME)	Discouraged revising and submitting to NSF - To a small extent (AME)	Discouraged revising and submitting to NSF - To a moderate extent (AME)	Discouraged revising and submitting to NSF - To a great extent (AME)
Race and ethnicity								
Asian, non-Hispanic	-0.11***	0.06***	0.04***	0.00	0.02	0.04**	-0.04***	-0.01
Black or African American, non-Hispanic	-0.13***	0.01	0.02	0.09***	-0.03	0.00	-0.02	0.05*
Hispanic	-0.04	-0.03*	0.04*	0.03*	-0.01	-0.02	0.00	0.03
White, non-Hispanic (ref.)	-	-	-	-	-	-	-	-
Multiple races, non-Hispanic	-0.04	0.00	0.04	-0.00	0.01	-0.07*	0.01	0.06
Other, NHPI, AIAN	-0.06	-0.00	0.00	0.06*	-0.04	-0.05	-0.04	0.13***
Do not wish to provide or unknown	-0.12***	0.05*	0.04	0.04*	-0.06**	0.02	-0.01	0.06**

Characteristics	Influenced decision to submit to non-NSF funding - Not at all (AME)	Influenced decision to submit to non-NSF funding - To a small extent (AME)	Influenced decision to submit to non-NSF funding - To a moderate extent (AME)	Influenced decision to submit to non-NSF funding - To a great extent (AME)	Discouraged revising and submitting to NSF - Not at all (AME)	Discouraged revising and submitting to NSF - To a small extent (AME)	Discouraged revising and submitting to NSF - To a moderate extent (AME)	Discouraged revising and submitting to NSF - To a great extent (AME)
Gender								
Female	0.02	-0.01	-0.01	0.00	-0.00	-0.01	0.00	0.01
Male (ref.)	-	-	-	-	-	-	-	-
Other or do not wish to provide or unknown	0.02	-0.04	0.00	0.02	-0.06*	-0.01	0.04	0.02
Disability status								
Yes—hearing, mobility, vision, or other disability	-0.01	-0.02	0.01	0.01	-0.07**	-0.01	0.02	0.06*
No—no disability (ref.)	-	-	-	-	-	-	-	-
Do not wish to provide	-0.02	-0.01	-0.00	0.03	-0.07*	-0.06*	0.06*	0.07**
Directorate								
BIO	0.00	0.00	-0.01	0.01	0.02	0.00	-0.01	-0.01
CISE	-0.00	0.01	0.00	-0.01	-0.04*	0.02	0.01	0.01
EHR	0.06***	-0.02	-0.02	-0.03**	0.04*	0.03	-0.02	-0.05***
ENG	-0.10***	0.03*	0.06***	0.01	-0.07***	0.00	0.05***	0.02
GEO	0.07***	-0.02	-0.02	-0.03**	0.02	-0.01	0.02	-0.03*
MPS (ref.)	-	-	-	-	-	-	-	-
SBE	0.05**	-0.02	-0.01	-0.01	-0.01	0.01	-0.02	0.02
OD (OIA and OISE)	0.06	0.05	-0.07*	-0.04	-0.04	-0.01	0.13*	-0.09*
Other	0.03	-0.03	-0.01	0.01	0.00	-0.00	-0.04	0.04
Institution type								
R1/MSI	-0.03	0.00	0.02	0.00	0.04	-0.03	-0.01	0.01
R1/non-MSI (ref.)	-	-	-	-	-	-	-	-
Non-R1/MSI	-0.01	0.01	-0.01	-0.00	0.02	0.01	-0.02	-0.01
Non-R1/non-MSI	0.00	0.00	-0.00	0.00	0.00	-0.01	0.00	0.00
For-profit organization	-0.00	-0.01	-0.02	0.03*	-0.04*	-0.09***	-0.02	0.15***
Government (local, state, federal, or tribal)	0.11	-0.01	-0.08	-0.01	0.03	-0.02	-0.09	0.08
Nonprofit organization	0.01	-0.01	0.02	-0.01	-0.03	0.03	0.00	0.00
Other, primary and secondary institution,	-0.01	0.01	-0.02	0.01	0.00	-0.06**	0.01	0.05

Characteristics	Influenced decision to submit to non-NSF funding - Not at all (AME)	Influenced decision to submit to non-NSF funding - To a small extent (AME)	Influenced decision to submit to non-NSF funding - To a moderate extent (AME)	Influenced decision to submit to non-NSF funding - To a great extent (AME)	Discouraged revising and submitting to NSF - Not at all (AME)	Discouraged revising and submitting to NSF - To a small extent (AME)	Discouraged revising and submitting to NSF - To a moderate extent (AME)	Discouraged revising and submitting to NSF - To a great extent (AME)
federally funded R&D, R1/unknown, non-R1/unknown								
Early career (degree ≤10 years ago)								
Yes	-0.02*	0.02	0.01	-0.01	0.04***	0.02*	-0.01	-0.05***
No (ref.)	-	-	-	-	-	-	-	-
Experience as applicant and reviewer								
Yes	-0.00	0.06***	-0.02*	-0.04***	0.10***	0.06***	-0.04***	-0.12***
No (ref.)	-	-	-	-	-	-	-	-

Note: influenced decision to submit to non-NSF $N = 12,364$; discouraged from revising and resubmitting $N = 12,384$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.4.4. Reviewer agreement that proposals are of high quality

Characteristics	Strongly disagree (AME)	Disagree (AME)	Neither agree nor disagree (AME)	Agree (AME)	Strongly agree (AME)
Race and ethnicity					
Asian, non-Hispanic	0.00	0.02*	0.04***	-0.02	-0.04***
Black or African American, non-Hispanic	0.02	-0.01	-0.03	0.03	-0.00
Hispanic	0.00	0.01	-0.02	-0.01	0.01
White, non-Hispanic (ref.)	-	-	-	-	-
Multiple races, non-Hispanic	-0.00	-0.02	0.06*	-0.04	-0.01
Other, NHPI, AIAN	0.01	0.07*	0.02	-0.07	-0.03
Do not wish to provide or unknown	-0.00	0.04**	0.08***	-0.06*	-0.06***
Gender					
Female	-0.00	-0.00	-0.02*	-0.01	0.03***
Male (ref.)	-	-	-	-	-
Other or do not wish to provide or unknown	0.01	0.03	0.02	-0.07*	0.01
Disability status					
Yes—hearing, mobility, vision, or other disability	0.01	0.01	0.03	-0.07**	0.02
No—no disability (ref.)	-	-	-	-	-
Do not wish to provide	0.00	-0.00	0.01	-0.03	0.03
Directorate					
BIO	-0.00	-0.01	0.03*	0.05***	-0.07***
CISE	0.00	0.05***	0.09***	-0.02	-0.13***
EHR	0.00	0.07***	0.11***	-0.03	-0.15***
ENG	0.00	0.05***	0.11***	-0.02	-0.14***
GEO	-0.00	-0.00	0.00	0.07***	-0.07***
MPS (ref.)	-	-	-	-	-
SBE	-0.00	0.02*	0.04**	0.03	-0.09***
OD (OIA and OISE)	-0.01***	0.03	0.09	-0.01	-0.11*
Other	-0.00	0.03	0.05	0.03	-0.10***
Institution type					
R1/MSI	0.00	0.00	-0.01	-0.00	0.01
R1/non-MSI (ref.)	-	-	-	-	-
Non-R1/MSI	0.00	-0.02**	-0.02	0.04	0.01
Non-R1/non-MSI	0.00	-0.01*	-0.03***	0.03**	0.01
For-profit organization	0.01	0.04*	0.09***	-0.10***	-0.03

Characteristics	Strongly disagree (AME)	Disagree (AME)	Neither agree nor disagree (AME)	Agree (AME)	Strongly agree (AME)
Government (local, state, federal, or tribal)	0.01	-0.00	-0.05	0.04	0.00
Nonprofit organization	0.01	-0.01	0.01	-0.01	0.01
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.01	0.00	0.00	-0.02	0.01
Early career (degree ≤10 years ago)					
Yes	0.00	-0.01*	-0.00	0.01	-0.00
No (ref.)	-	-	-	-	-
Experience as applicant and reviewer					
Yes	-0.00	0.01**	0.05***	-0.01	-0.05***
No (ref.)	-	-	-	-	-

Note: $N = 11,732$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.4.5. Factors for improvement cited by applicants and reviewers

Characteristics	Factors					
	Timeliness of decisions (AME)	Written reviews (AME)	Panel summaries (AME)	PI conversations	Information available during proposal submission (AME)	Review process from the perspective of a reviewer (AME)
Race and ethnicity						
Asian, non-Hispanic	-0.02	0.01	-0.00	-0.01	-0.03***	0.04***
Black or African American, non-Hispanic	-0.03	-0.02	-0.03*	0.03	0.01	0.05**
Hispanic	-0.01	-0.02	-0.02	0.01	0.01	0.03*
White, non-Hispanic (ref.)	-	-	-	-	-	-
Multiple races, non-Hispanic	0.01	-0.06*	-0.04*	0.05*	0.02	0.02
Other, NHPI, AIAN	-0.03	-0.06*	-0.03	0.03	0.00	0.08**
Do not wish to provide or unknown	-0.02	0.01	-0.02	-0.01	0.01	0.03*
Gender						
Female	-0.00	-0.03***	-0.00	0.01	0.02***	0.00
Male (ref.)	-	-	-	-	-	-
Other or do not wish to provide or unknown	-0.05*	-0.01	0.01	0.03	0.01	0.01
Disability status						
Yes—hearing, mobility, vision, or other disability	-0.03	0.04	-0.03	0.00	0.02	-0.01
No—no disability (ref.)	-	-	-	-	-	-
Do not wish to provide	-0.00	0.02	-0.01	0.00	-0.02*	0.02
Directorate						
BIO	0.02	-0.03*	0.03**	0.01	0.01	-0.03***
CISE	0.01	-0.01	0.00	0.02	0.01	-0.02*
EHR	0.02	-0.04**	-0.02	0.04***	0.04***	-0.04***
ENG	0.00	-0.02	-0.01	0.03***	-0.00	-0.00
GEO	0.01	-0.01	0.04***	0.01	0.00	-0.04***
MPS (ref.)	-	-	-	-	-	-

Characteristics	Factors					
	Timeliness of decisions (AME)	Written reviews (AME)	Panel summaries (AME)	PI conversations	Information available during proposal submission (AME)	Review process from the perspective of a reviewer (AME)
SBE	-0.02	0.04*	0.01	-0.00	0.03***	-0.05***
OD (OIA and OISE)	-0.05	-0.07	-0.02	0.06	0.07*	0.01
Other	-0.01	-0.05*	-0.02	0.04*	0.04**	-0.00
Institution type						
R1/MSI	-0.02	-0.00	0.01	0.02	-0.00	0.00
R1/non-MSI (ref.)	-	-	-	-	-	-
Non-R1/MSI	-0.01	-0.01	0.00	-0.01	0.02*	0.01
Non-R1/non-MSI	-0.01	-0.01	-0.00	0.01	0.01	0.01
For-profit organization	0.11***	-0.08***	-0.06***	0.00	0.02*	0.01
Government (local, state, federal, or tribal)	0.05	-0.04	-0.03	-0.01	0.03	0.01
Nonprofit organization	-0.01	-0.01	-0.02	-0.02	0.02	0.04*
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	-0.01	-0.03	-0.01	0.01	0.03**	0.01
Early career (degree ≤10 years ago)						
Yes	0.09***	-0.02*	-0.02***	-0.03***	0.01*	-0.02***
No (ref.)	-	-	-	-	-	-

Characteristics	Factors					
	Timeliness of decisions (AME)	Written reviews (AME)	Panel summaries (AME)	PI conversations	Information available during proposal submission (AME)	Review process from the perspective of a reviewer (AME)
Experience as applicant and reviewer						
Applicant only	-0.04***	0.05***	-0.03***	-0.01*	0.02***	0.02**
Reviewer only	-0.11***	-0.03**	0.04***	0.01	0.04***	0.03***
Applicant and reviewer (ref.)	-	-	-	-	-	-

Note: $N = 17,047$

Multinomial logistic regression

Factors for improvement have six nominal outcomes: 1 = timeliness of decisions about, and responsiveness to, proposals by NSF staff; 2 = quality of feedback to PIs in the form of comments in written reviews (reference); 3 = quality of feedback to PIs in the form of comments in panel summaries; 4 = quality of PI conversations with, and written comments from, program officers; 5 = quality of information available during proposal submission; 6 = quality of the review process from the perspective of a reviewer.

Findings are presented in Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; PI = principal investigator; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.4.6.A. Applicant satisfaction with the quality of information provided by NSF—2021 and 2019 differences

Characteristics	Level 1	Level 2	Level 3	Level 4
Year				
2021	0.61***	0.67***	0.80***	1.25***
2019 (ref.)	-	-	-	-
Race and ethnicity				
Year * Asian, non-Hispanic	0.75	0.83	0.95	1.01
Year * Black or African American, non-Hispanic	0.76	0.76	0.76	0.77
Year * Hispanic	1.43	1.03	0.89	0.88
Year * White, non-Hispanic (ref.)	-	-	-	-
Year * Multiple races, non-Hispanic	1.30	1.01	1.02	0.90
Year * Other, NHPI, AIAN, HSP	1.24	0.66	0.53	0.81
Year * Do not wish to provide or unknown	1.60	1.11	1.06	1.02
Gender				
Year * Female	1.08	1.22*	1.01	0.91
Year * Male (ref.)	-	-	-	-
Year * Other or do not wish to provide or unknown	0.81	0.92	0.81	0.85
Institution type				
Year * Academic (ref.)	-	-	-	-
Year * For-profit organization	1.00	1.04	1.06	1.05
Year * Government (local, state, federal, or tribal)	1.82	1.49	1.57	1.25
Year * Federally funded R&D	0.34	1.01	0.82	0.71
Year * Nonprofit organization	1.00	1.02	0.99	0.91
Year * Other, primary institution, secondary institution	1.79	0.99	0.99	0.96
Early career (degree ≤10 years ago)				
Year * Early career	0.81	0.94	0.95	1.08
Year * Not early career (ref.)	-	-	-	-

Note: N = 29,431

Generalized Ordinal Logistic Regression (Gologit)

Variable has five ordinal outcomes: 1 = very dissatisfied; 2 = dissatisfied; 3 = neither dissatisfied nor satisfied; 4 = satisfied; 5 = very satisfied.

Coefficients presented in odds ratios and across four levels indicating to a greater extent.

Level 1: odds ratios of “very dissatisfied” versus “dissatisfied,” “neither dissatisfied nor satisfied,” “satisfied,” and “very satisfied”

Level 2: odds ratios of “very dissatisfied,” “dissatisfied” versus “neither dissatisfied nor satisfied,” “satisfied,” and “very satisfied”

Level 3: odds ratios of “very dissatisfied,” “dissatisfied,” and “neither dissatisfied nor satisfied” versus “satisfied” and “very satisfied”

Level 4: odds ratios of “very dissatisfied,” “dissatisfied,” “neither dissatisfied nor satisfied,” and “satisfied” versus “very satisfied”

Findings are presented using odds ratios.

Findings are weighted to adjust for nonresponse bias.

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

See appendix B.4 for the full list of independent variables included in the model.

AIAN = American Indian or Alaska Native; HSP = Hispanic; NHPI = Native Hawaiian or Other Pacific Islander; R&D = research and development

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.4.6.B. Applicant satisfaction with the quality of information provided by NSF—2021 and 2019 differences (average marginal effects for 2021)

Characteristics	Very dissatisfied (AME)	Somewhat dissatisfied (AME)	Neither dissatisfied nor satisfied (AME)	Somewhat satisfied (AME)	Very satisfied (AME)
Year					
2021	0.02***	0.02***	0.01**	-0.09***	0.04***
2019 (ref.)	-	-	-	-	-

Note: N = 29, 431

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME) for the year effect of “2021” modeled in table C.4.6.A.

Findings are weighted to adjust for nonresponse bias.

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

See appendix B.4 for the full list of independent variables included in the model.

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C.5 Fairness: Regression Tables

Regression coefficients for figures 4.5.1 to 4.5.5 in chapter 4.4 (*Fairness in the Merit Review Process*) are presented in tables C.5.1 to C.5.4.

Table C.5.1. Applicants and reviewers agreement that the NSF merit review process is fair

Characteristics	Strongly disagree (AME)	Disagree (AME)	Neither agree nor disagree (AME)	Agree (AME)	Strongly agree (AME)
Race and ethnicity					
Asian, non-Hispanic	-0.00	-0.01	-0.02**	0.01	0.03**
Black or African American, non-Hispanic	0.03**	0.04*	0.04*	-0.10***	-0.01
Hispanic	0.00	0.03*	0.02	-0.09***	0.03**
White, non-Hispanic (ref.)	-	-	-	-	-
Multiple races, non-Hispanic	0.01	0.01	0.02	-0.02	-0.02
Other, NHPI, AIAN	0.04**	0.07**	0.07*	-0.17***	-0.00
Do not wish to provide or unknown	0.01	0.04**	0.05**	-0.06**	-0.03**
Gender					
Female	-0.01*	0.01	0.03***	-0.01	-0.02***
Male (ref.)	-	-	-	-	-
Other or do not wish to provide or unknown	0.01	0.05**	0.01	-0.05	-0.02
Disability status					
Yes—hearing, mobility, vision, or other disability	0.04***	0.06***	0.02	-0.08***	-0.05***
No—no disability (ref.)	-	-	-	-	-
Do not wish to provide	0.04**	0.01	0.07**	-0.07*	-0.06***

Characteristics	Strongly disagree (AME)	Disagree (AME)	Neither agree nor disagree (AME)	Agree (AME)	Strongly agree (AME)
Directorate					
BIO	-0.01*	-0.01	-0.00	0.01	0.02*
CISE	-0.01	-0.01	0.01	-0.03*	0.03**
EHR	-0.02***	-0.02**	-0.01	0.01	0.05***
ENG	-0.01	0.01	0.02*	-0.03*	0.00
GEO	-0.02***	-0.01	0.02	0.00	0.00
MPS (ref.)	-	-	-	-	-
SBE	-0.01	-0.01	0.03	-0.02	0.01
OD (OIA and OISE)	-0.01	0.02	0.01	-0.04	0.03
Other	0.01	0.02	0.07**	-0.12***	0.03
Institution type					
R1/MSI	0.01	0.00	-0.02	0.03	-0.02*
R1/non-MSI (ref.)	-	-	-	-	-
Non-R1/MSI	-0.00	-0.00	-0.00	-0.02	0.03*
Non-R1/non-MSI	0.00	0.00	-0.00	-0.01	0.01
For-profit organization	0.03***	0.01	-0.04*	-0.06***	0.06***
Government (local, state, federal, or tribal)	0.01	-0.06*	-0.03	0.06	0.02
Nonprofit organization	-0.00	-0.03*	0.03	0.05	-0.04**
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.01	-0.00	-0.03	-0.00	0.02
Early career (degree ≤10 years ago)					
Yes	-0.00	-0.02**	-0.00	0.03**	-0.00
No (ref.)	-	-	-	-	-
Role in the Merit Review Process					
Applicant only	0.03***	0.07***	0.10***	-0.10***	-0.10***
Reviewer only	-0.02***	-0.06***	-0.05***	0.06***	0.08***
Both applicant and reviewer (ref.)	-	-	-	-	-

Note: N = 17,285

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; PI = principal investigator; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.5.2. Applicant agreement that individuals submitting proposals are treated fairly

Characteristics	Strongly disagree (AME)	Disagree (AME)	Neither agree nor disagree (AME)	Agree (AME)	Strongly agree (AME)
Race and ethnicity					
Asian, non-Hispanic	0.01	-0.00	0.00	-0.00	-0.01
Black or African American, non-Hispanic	0.04*	0.02	0.07**	-0.09***	-0.03
Hispanic	0.01	0.02	-0.01	-0.02	0.01
White, non-Hispanic (ref.)	-	-	-	-	-
Multiple races, non-Hispanic	0.02	0.03	0.03	-0.02	-0.05*
Other, NHPI, AIAN	0.10***	0.07**	0.02	-0.15***	-0.03
Do not wish to provide or unknown	0.04**	0.01	0.07***	-0.08***	-0.03
Gender					
Female	-0.01	0.01*	0.03**	-0.02	-0.02*
Male (ref.)	-	-	-	-	-
Other or do not wish to provide or unknown	0.03	0.02	0.03	-0.05	-0.02
Disability status					
Yes—hearing, mobility, vision, or other disability	0.08***	0.02	-0.01	-0.04	-0.06***
No—no disability (ref.)	-	-	-	-	-
Do not wish to provide	0.05**	0.07***	0.01	-0.07**	-0.05*
Directorate					
BIO	-0.02*	-0.01	-0.02	0.01	0.03*
CISE	-0.01	0.01	0.01	-0.02	0.01
EHR	-0.04***	-0.01	-0.02	0.02	0.05***
ENG	-0.00	0.02*	0.01	-0.03*	-0.00
GEO	-0.03***	-0.00	0.01	-0.01	0.04**
MPS (ref.)	-	-	-	-	-
SBE	-0.03**	-0.02	0.02	-0.02	0.04**
OD (OIA and OISE)	0.01	-0.00	-0.04	0.08	-0.05
Other	0.01	0.01	0.04	-0.05	-0.00
Institution type					
R1/MSI	0.01	-0.01	0.02	0.00	-0.03
R1/non-MSI (ref.)	-	-	-	-	-
Non-R1/MSI	-0.01	0.00	-0.00	-0.01	0.02
Non-R1/non-MSI	0.01*	-0.01*	0.00	-0.01	0.00
For-profit organization	0.03**	-0.02	-0.02	-0.04*	0.05**
Government (local, state, federal, or tribal)	0.03	-0.01	-0.08	-0.08	0.14
Nonprofit organization	-0.00	-0.04**	0.06*	0.00	-0.02
Other, primary and secondary institution, federally funded	0.04*	-0.02	-0.03	-0.01	0.02

Characteristics	Strongly disagree (AME)	Disagree (AME)	Neither agree nor disagree (AME)	Agree (AME)	Strongly agree (AME)
R&D, R1/unknown, non-R1/unknown					
Early career (degree ≤10 years ago)					
Yes	-0.01	-0.01	-0.01	0.02*	0.01
No (ref.)	-	-	-	-	-
Experience as applicant and reviewer					
Yes	-0.05***	-0.02***	-0.09***	0.11***	0.06***
No (ref.)	-	-	-	-	-

Note: N = 13,429

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences. A cell with a dash (-) equals no data available. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table C.5.3. Reviewer agreement that individuals submitting proposals are treated fairly

Characteristics	Strongly disagree (AME)	Disagree (AME)	Neither agree nor disagree (AME)	Agree (AME)	Strongly agree (AME)
Race and ethnicity					
Asian, non-Hispanic	0.00	-0.00	-0.02	0.02	0.00
Black or African American, non-Hispanic	0.02*	0.05*	0.07*	-0.06*	-0.07**
Hispanic	0.02*	0.02	-0.00	-0.06**	0.03
White, non-Hispanic (ref.)	-	-	-	-	-
Multiple races, non-Hispanic	-0.01	-0.01	0.02	0.01	-0.01
Other, NHPI, AIAN	0.04	0.03	0.06	-0.12***	-0.01
Do not wish to provide or unknown	0.00	0.02	0.02*	0.00	-0.04
Gender					
Female	-0.00	0.01*	0.03**	0.00	-0.04***
Male (ref.)	-	-	-	-	-
Other or do not wish to provide or unknown	0.02	0.04**	0.03	-0.06*	-0.03
Disability status					
Yes—hearing, mobility, vision, or other disability	0.03*	0.05**	0.00	-0.07	-0.01
No—no disability (ref.)	-	-	-	-	-
Do not wish to provide	0.02	0.02	0.07*	-0.07	-0.04

Characteristics	Strongly disagree (AME)	Disagree (AME)	Neither agree nor disagree (AME)	Agree (AME)	Strongly agree (AME)
Directorate					
BIO	-0.01	-0.01	0.03	-0.00	-0.01
CISE	0.00	0.00	-0.02*	-0.03	0.05**
EHR	-0.01**	-0.01*	-0.05***	-0.02	0.10***
ENG	0.00	0.02	0.01	-0.05*	0.02
GEO	-0.01	-0.01	0.06***	-0.02	-0.04
MPS (ref.)	-	-	-	-	-
SBE	-0.01	-0.01	0.04*	-0.03	0.01
OD (OIA and OISE)	0.02	-0.01	-0.00	0.01	-0.02
Other	0.00	-0.01	0.04	-0.08	0.04
Institution type					
R1/MSI	-0.00	0.00	-0.01	0.02	-0.02
R1/non-MSI (ref.)	-	-	-	-	-
Non-R1/MSI	-0.01*	-0.01	0.02	-0.00	0.00
Non-R1/non-MSI	0.00	-0.00	0.00	-0.00	-0.00
For-profit organization	0.01	-0.02*	-0.04*	-0.08*	0.14***
Government (local, state, federal, or tribal)	-0.00	-0.03	0.00	0.05	-0.02
Nonprofit organization	0.00	-0.03*	0.03	0.01	-0.02
Other, primary and secondary institution, federally funded R&D, R1/unknown, non-R1/unknown	0.00	-0.02*	0.01	-0.02	0.03
Early career (degree ≤10 years ago)					
Yes	0.00	-0.00	0.02*	0.04	-0.06***
No (ref.)	-	-	-	-	-
Experience as applicant and reviewer					
Yes	0.01*	0.03***	0.03**	-0.02	-0.05***
No (ref.)	-	-	-	-	-

Note: N = 11,688

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME).

Findings are weighted to adjust for nonresponse bias.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences
A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.5.4.A. Applicants and reviewers agreement that the NSF merit review process is fair—2021 and 2019 differences

Characteristics	Level 1	Level 2	Level 3	Level 4
Year				
2021	0.49***	0.53***	0.52***	0.59***
2019 (ref.)	-	-	-	-
Race and ethnicity				
Year * Asian, non-Hispanic	0.84	0.87	0.98	0.98
Year * Black or African American, non-Hispanic	0.66	0.72	0.83	0.98
Year * Hispanic	1.11	1.06	1.10	1.25*
Year * White, non-Hispanic (ref.)	-	-	-	-
Year * Multiple races, non-Hispanic	0.88	1.26	1.00	0.89
Year * Other, NHPI, AIAN, HSP	0.89	0.75	0.45**	0.69
Year * Do not wish to provide or unknown	1.06	1.10	0.98	0.99
Gender				
Year * Female	1.09	0.94	0.99	1.06
Year * Male (ref.)	-	-	-	-
Year * Other or do not wish to provide or unknown	0.70	0.85	0.99	0.95
Institution type				
Year * Academic (ref.)	-	-	-	-
Year * For-profit organization	0.54**	0.63***	0.70***	0.74**
Year * Government (local, state, federal, or tribal)	0.88	1.74	2.13**	1.23
Year * Federally funded R&D	1.19	1.34	1.25	1.73*
Year * Nonprofit organization	1.28	1.74**	1.07	0.85
Year * Other, primary institution, secondary institution	1.64	0.93	1.16	0.91
Early career (degree ≤10 years ago)				
Year * Early career	0.67*	0.78**	0.87*	0.96
Year * Not early career (ref.)	-	-	-	-

Note: N = 42,187

Generalized Ordinal Logistic Regression (Gologit)

Variable has five ordinal outcomes: 1 = strongly disagree; 2 = disagree; 3 = neither agree nor disagree; 4 = agree; 5 = strongly agree.

Coefficients presented in odds ratios and across four levels indicating to a greater extent.

Level 1: odds ratios of “strongly disagree” versus “disagree,” “neither agree nor disagree,” “agree,” and “strongly agree”

Level 2: odds ratios of “strongly disagree” and “disagree” versus “neither agree nor disagree,” “agree,” and “strongly agree”

Level 3: odds ratios of “strongly disagree,” “disagree,” and “neither agree nor disagree” versus “agree” and “strongly agree”

Level 4: odds ratios of “strongly disagree,” “disagree,” “neither agree nor disagree,” and “agree” versus “strongly agree”

Findings are presented in odds ratios.

Findings are weighted to adjust for nonresponse bias.

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

See appendix B.4 for the full list of independent variables included in the model.

AIAN = American Indian or Alaska Native; BIO = Biological Sciences; CISE = Computer and Information Science and Engineering; EHR = Education and Human Resources; ENG = Engineering; GEO = Geosciences; HSP = Hispanic; ISE = International Science and Engineering; MPS = Mathematical and Physical Sciences; MSI = minority-serving organization; NHPI = Native Hawaiian or Other Pacific Islander; OD = Office of the Director; OIA = Office of Integrative Activities; R&D = research and development; R1 = doctoral universities with very high research activities (Carnegie Classification of Institutions of Higher Education N.d.); SBE = Social, Behavioral, and Economic Sciences

An cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.5.4.B. Applicants and reviewers agreement that the NSF merit review process is fair—2021 and 2019 differences (average marginal effects for 2021)

Characteristics	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Year					
2021	0.02***	0.05***	0.08***	-0.07***	-0.08***
2019 (ref.)	-	-	-	-	-

Note: $N = 42,187$

Generalized Ordinal Logistic Regression (Gologit)

Findings are presented using Average Marginal Effects (AME) for the year effect of “2021” modeled in Table C.5.4.A.

Findings are weighted to adjust for nonresponse bias.

Model controls for race and ethnicity, gender, institution, early career, directorate/office, and role in the merit review process.

See appendix B.4 for the full list of independent variables included in the model.

A cell with a dash (-) equals no data available.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$