United States Government Accountability Office

Report to the Ranking Member, Committee on Commerce, Science, and Transportation, U.S. Senate

July 2022

RESEARCH RELIABILITY

Federal Actions Needed to Promote Stronger Research Practices

GAO-22-104411

July 2022

RESEARCH RELIABILITY

Federal Actions Needed to Promote Stronger Research Practices

GAO Highlights

Highlights of GAO-22-104411, a report to the Ranking Member, Committee on Commerce, Science, and Transportation, U.S. Senate

Why GAO Did This Study

In 2019, the U.S. government funded more than $42 billion in basic scientific research across a wide range of scientific disciplines. Unsuccessful attempts to reproduce and replicate research results have been documented across many scientific disciplines, including those funded by NASA, NIH, and NSF. The scientific community has expressed concern over the difficulty of replicating prior research results.

GAO was asked to review strategies to improve the reliability of federally funded research. Among other things, this report (1) examines what actions, according to experts, federal agencies could take to foster rigor and transparency in the research they fund; and (2) assesses the extent to which selected federal science funding agencies have taken actions to improve rigor and transparency. GAO conducted a literature review; reviewed NIH, NSF, and NASA documents; and conducted four roundtable discussions with 22 experts. GAO also interviewed agency officials as well as stakeholders from academia, professional societies, publishing, and other parts of the scientific community.

What GAO Recommends

GAO is making six recommendations, two each to NIH, NSF, and NASA to evaluate research using indicators of rigor and transparency, and to use this information to inform further actions. NIH and NSF concurred with the recommendations. NASA did not concur with our first recommendation and partially concurred with our second. GAO continues to believe the recommendations are valid.

View GAO-22-104411. For more information, contact Candice Wright at (202) 512-6888 or wrightc@gao.gov.

What GAO Found

The National Institutes of Health (NIH), the National Science Foundation (NSF), and the National Aeronautics and Space Administration (NASA) are the three largest federal funders of basic scientific research in the United States. According to leading experts GAO interviewed, these agencies could do more to increase the rigor and transparency of the research they fund by taking actions to better align awards and recognition for researchers with more rigorous and transparent research practices. Experts suggested, for example, that agencies could incentivize or mandate that researchers preregister their studies as a means to share their research plans before the research is conducted. Doing so would enable other researchers to comment on and strengthen the methodology and analysis plans. Experts further suggested that agencies help improve standards for data repositories where research data are stored publicly, encourage the publication of null research results, and support training in statistical analysis and study design. Although the scientific community has developed many such practices to enhance research reliability, GAO found that they are not widely adopted because of researcher misconceptions and misaligned incentives in funding and publishing, among other things.

Role of Rigor and Transparency in Research Reliability

Rigor
Soundness and precision of study design, execution, data collection, and analysis.

Transparency
Ensuring that information about study design, execution, data collection and analysis, and conclusions are clearly documented and shared freely.

- Rigor and transparency increase the likelihood of reliable research results. They help ensure that results are valid and can be understood, and that other researchers have confidence in interpreting the results and building on them.
- When researchers provide a clear, specific, and complete accounting of the materials and methods they used, the results they found, and the uncertainty associated with the methods and results, other researchers will know how to interpret the findings.

Source: GAO analysis of information from the National Academies of Sciences, Engineering, and Medicine. | GAO-22-104411

NIH, NSF, and NASA have taken steps to promote and support additional rigor and transparency in research, such as establishing requirements for researchers to disclose research results and associated data publicly. However, these agencies largely rely on grant application reviews and the prepublication peer review process to help ensure research rigor. GAO found that these agencies do not evaluate the rigor and transparency of the research they fund to help identify strategies for improvement. Specifically, they do not collect indicators of rigorous study design and transparency of research results such as study sample size, adherence to research plans, or the extent to which research data are findable, accessible, and usable. As a result, the agencies lack information to support changes to the grant making process and research funding priorities. Federal guidance and *Standards for Internal Control in the Federal Government* call for agencies to prioritize making federally funded research more rigorous and transparent and to use quality information to achieve agency objectives. Without this information on the research they fund, agencies are limited in their ability to take effective actions to improve research reliability, like those the experts described to GAO.

Contents

Abbreviations

FAIR	findable, accessible, interoperable, and reusable
NASA	National Aeronautics and Space Administration
NIH	National Institutes of Health
NSF	National Science Foundation
OSTP	Office of Science and Technology Policy
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2

U.S. GOVERNMENT ACCOUNTABILITY OFFICE
441 G St. N.W.
Washington, DC 20548

July 28, 2022

The Honorable Roger F. Wicker
Ranking Member
Committee on Commerce, Science, and Transportation
United States Senate

Dear Mr. Wicker:

Historically, the federal government has been the single largest funding source of the nation's basic research. In 2019, the U.S. government funded more than $42 billion in basic scientific research aimed at generating new knowledge across a wide range of scientific disciplines, including life sciences, physical sciences, and social sciences.[1] This funding helps spur technological breakthroughs that benefit our economy, strengthen our national security, and improve the overall health and well-being of our society.

The importance of promoting rigor and transparency in research has become more prominent over the past decade or more, as evidenced by a series of reports and workshops from the National Academies of Sciences, Engineering, and Medicine (National Academies).[2] Scientific discoveries can be serendipitous. However, the 2019 National Academies report notes that a body of reliable knowledge tends to be the cumulative product of investigations by successive research employing careful design, testing, corrections, and confirmation over a period of years. Policymakers; researchers; federal and nonfederal funders of research; academic, corporate, and independent research institutions; professional societies; publishers; journalists; and many others depend on reliable research.

[1]Data on U.S. research expenditures (in current dollars) are from the National Science Foundation, National Center for Science and Engineering Statistics, *National Patterns of R&D Resources: 2019–20 Data Update*, NSF 22-320, Table 7, Feb. 22, 2022, the most recently available data at the time of our review.

[2]See National Academies reports including: *Reproducibility and Replicability in Science* (Washington, D.C.: The National Academies Press, 2019); *Open Science by Design: Realizing a Vision for 21st Century* (Washington, D.C.: The National Academies Press, 2018); and *Advancing Open Science Practices: Stakeholder Perspectives on Incentives and Disincentives, Proceedings of a Workshop in Brief* (Washington, D.C.: The National Academies Press, 2020).

The 2019 National Academies report also highlights that reliability of research is dependent, in large part, on methodological rigor. In addition, transparency in reporting the data and analytic methods employed is important to enable other researchers to assess and understand the results. Without the latter, it can be difficult to test prior work and learn from it. A number of recent studies attempted and found it difficult to replicate peer-reviewed research results from a wide spectrum of disciplines. This has prompted many in the scientific community to call for improvements to protocols and practices among researchers and funders to better assure both rigor and transparency for existing and planned research. The difficulty of replicating some of the research has been due, in whole or in part, to shortcomings in methodological rigor or a lack of transparency that provides sufficient information about the research to allow another researcher to replicate and confirm the findings. Not every type of scientific research lends itself to replication, nor does research need to be entirely replicable to be useful or informative.

You asked us to review strategies to improve the reliability of federally funded research, and to examine the processes by which federal funders promote or support rigor and transparency in the research they fund.

In this report, we (1) outline strategies that, according to stakeholders in the scientific community, are available to promote research rigor and transparency, and factors they say currently discourage wider adoption in the research community; (2) examine what actions, according to experts, federal agencies could take to foster rigor and transparency in the research they fund; and (3) assess the extent to which selected federal science funding agencies have taken actions to improve rigor and transparency.

To address the first objective, we conducted a review of relevant literature spanning 15 years on topics related to the reproducibility and replicability of research and interviewed a non-generalizable sample of stakeholders from across the scientific community.[3] These stakeholders included representatives from research institutions, non-profit and for-profit publishers, libraries, professional societies, private funders, and others. Those we interviewed were not a representative sample of all stakeholders with expertise on research reliability, but they each

[3]See the selected bibliography following the appendixes.

demonstrated extensive involvement on these topics and offered a range of perspectives.

To address the second objective, we carried out a series of interviews and held four roundtable discussions with leading experts from academia and nonprofit research organizations who have focused on this issue and have backgrounds in various scientific disciplines such as geophysics, biochemistry, psychology, and economics. We identified experts through what is termed a "snowball sample" by selecting a core group of experts and surveying them for the names of other experts they would recommend. We continued in this manner, iteratively, until we compiled a list of more than 700 experts. Our four discussion groups consisted of 22 experts, chosen primarily based on the number of times that each expert had been recommended by their peers and our assessment of their publications and experience. The roundtable discussions focused on federal actions that could be taken to address challenges to research reproducibility and replicability. To select viewpoints for inclusion in the body of this report, we considered the extent to which a particular topic was discussed, the degree to which other experts agreed or disagreed with one another, and whether the experts provided sufficient support for a particular discussion topic, among other factors. Our analysis of the results of our interviews and roundtable discussions sought to characterize the range of factors that met this threshold. Our approach was designed to capture the range of views, and the associated rationale for these of views, rather than to quantify the prevalence of support among experts for any particular view.

To address the third objective, we interviewed officials from the three largest civilian federal funders of basic research in the United States: the National Institutes of Health (NIH), the National Science Foundation (NSF), and the National Aeronautics and Space Administration (NASA). We also spoke with officials from the Office of Science and Technology Policy (OSTP) and the National Institute of Standards and Technology who have a role in developing related guidance and standards, respectively. We reviewed selected agencies' policies, procedures, and guidance, applicable laws and regulations, and federal standards for internal control. For more information on our objectives, scope, and methodology, see appendix I.

We conducted this performance audit from August 2020 to July 2022 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our

findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

The Scientific Process and Recent Concerns

The scientific accumulation of knowledge is an iterative process of discovery, confirmation, and correction. Trustworthy and reliable results are more likely when studies are conducted with methodological and analytical rigor, and with sufficient transparency to enable follow-up studies.

Key to building a body of knowledge is the process of testing the results of prior work by reproducing the research or by replicating its outcomes. However, the reliability and efficiency of this iterative process has come into question following a number of studies that found it difficult to confirm the results of peer-reviewed work. For example, in 2012 an article was published outlining how scientists from Amgen, Inc. had attempted, over the previous decade, to replicate 53 critical research studies on blood disorders and cancer that had been deemed "landmark" studies. The scientists reported they could replicate the results from six of the 53 studies when they repeated the experiments.[4] In another study published in 2015, researchers could, using their replication standards, replicate 39 of 100 studies published in three psychology journals.[5] More recently, a study published in 2021 described researchers' attempts to replicate 193 experiments in preclinical cancer research published in 53 articles. Researchers were able to attempt replication on 50 experiments from 23 of the papers, in part because many of the original papers failed to report key descriptive statistics. According to the study, none of the 193 experiments were described in sufficient detail to enable researchers to design similar research protocols.[6] As discussed in the 2019 National Academies report, unsuccessful attempts to reproduce and replicate

[4]Begley, C., and L. Ellis. "Raise Standards for Preclinical Cancer Research," *Nature* 483, 531–533 (2012). https://doi.org/10.1038/483531a.

[5]Open Science Collaboration. "Estimating the Reproducibility of Psychological Science," *Science* 349, aac4716 (2015). https://doi.org/10.1126/science.aac4716.

[6]Errington, T., M. Mathur, C. Soderberg, et al., "Investigating the Replicability of Preclinical Cancer Biology," *eLife* (2021). https://doi.org/10.7554/eLife.71601.

research results have been documented across many scientific disciplines, including disciplines funded by NASA, NIH, and NSF.

In 2013, OSTP issued a memorandum that directed each federal agency with more than $100 million in annual research and development expenditures to develop a plan to support increased public access to the results of federally funded research, publications and data in particular.[7] In 2017, Congress directed NSF to commission a report by the National Academies to assess the reproducibility and replicability of scientific research and provide findings and recommendations for improving research rigor and transparency.[8] The National Academies report, released in 2019, included an overview of assessments of reproducibility and replicability in scientific research; a discussion of factors contributing to non-reproducible and non-replicable research; and a summary of efforts to improve research outcomes.[9] The National Academies report discussed the difficulty in assessing the extent of non-replicability and non-reproducibility in research given the different types of studies, differences across scientific disciplines, and the lack of agreed-upon standards for assessing replication, among other things. The report included a series of recommendations for improving research. These included recommendations aimed at increasing transparency, providing additional researcher training, and investing in tools and infrastructure that support reproducibility.

In 2019, OSTP's National Science and Technology Council launched the Joint Committee on the Research Environment to address challenges facing the scientific research enterprise. As part of this effort, OSTP formed a Subcommittee on Rigor and Integrity in Research to address concerns over institutional incentives and systemic practices that undermine rigor and integrity. In addition, the National Science and Technology Council's Subcommittee on Open Science, led by OSTP, NIH, and NSF, helps agencies coordinate with each other to implement public access plans, including through an interagency working group.

[7]Office of Science and Technology Policy, *Increasing Access to the Results of Federally Funded Scientific Research,* Memorandum (Washington, D.C.: Feb. 22, 2013).

[8]American Innovation and Competitiveness Act, Pub. L. No. 114-329, § 116, 130 Stat. 2969, 2994 (2017).

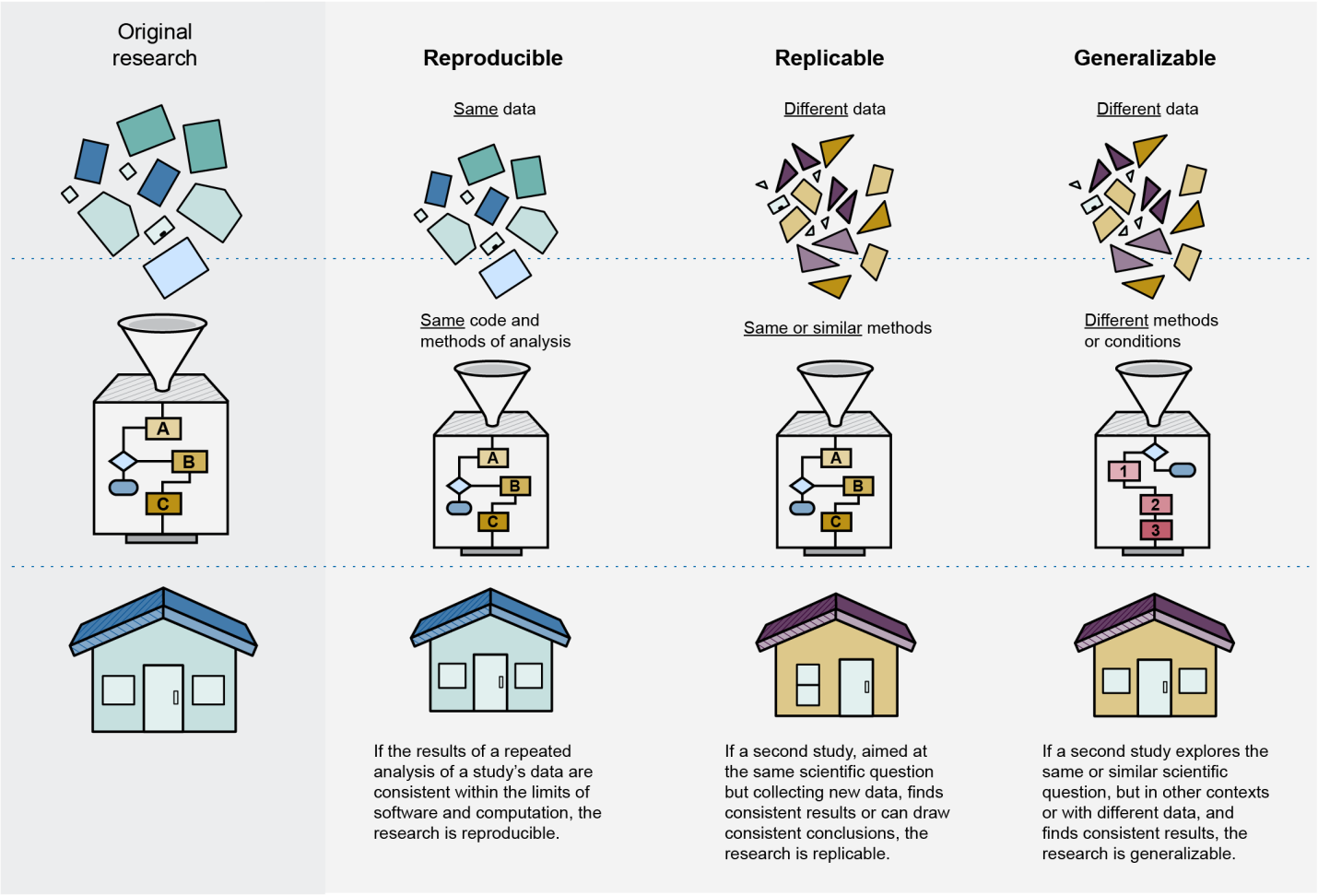
[9]2019 National Academies report.

Reproducibility and Replicability

To reproduce prior research work means to achieve consistent results using the same input data, computational steps, and methods of analysis. To replicate prior research, on the other hand, is to confirm prior results using the same or similar methodology, but different data, to answer the same question. Although the research community has sometimes used the terms reproducibility and replicability interchangeably, the National Academies distinguishes the terms and illustrates the variety of ways that research is tested for reliability.[10] Another characteristic of reliable research is generalizability, which is the extent to which a similar study, with different methods and different data, achieves consistent results. (See figure 1.)

[10]For purposes of this report, we will refer to replication as a general term for efforts to confirm prior research outcomes.

Figure 1: Elements of Reliable Research



Source: GAO analysis of information from the National Academies of Sciences, Engineering, and Medicine. | GAO−22−104411

As the 2019 National Academies report states, replication is neither guaranteed nor always expected, and a variety of factors can affect a replication, including the discovery of an unknown effect, inherent variability in the system, inability to control complex variables, or substandard research practices. Replication is a matter of degree rather than a definitive assessment of success or failure. In addition, as the 2019 National Academies report discusses, a successful replication does not guarantee that the original results of a study were correct, nor does a single failed replication conclusively refute the original claims. Moreover, the ability to test prior work precisely is more likely in the physical and computational sciences—disciplines that tend to involve precise

measurements using standard equipment and controllable samples. By contrast, studies in the social and life sciences more often involve human or animal subjects or microorganisms, which add more variability to the research. Other factors that affect replicability can be related to the intrinsic variation and complexity of the natural world and the limits of current technologies.[11]

Research Rigor and Transparency

As discussed in the 2019 National Academies report, rigor and transparency are key to ensuring that researchers can build on earlier work and continue scientific inquiry.

Rigor. Methodological rigor refers to a well-designed plan and adherence to methodological best practices, according to the National Academies report. Or, as one study defined it, rigor is the soundness and precision of a study in terms of the design, execution, data collection, and analysis.[12] Steps researchers must take to ensure methodological rigor differ across scientific disciplines. Nevertheless, poor research practices can hinder the ability to exploit and build on prior research. As the 2019 National Academies report discusses, striking a balance between the need for methodological rigor and the need for risk taking to advance science and push the boundaries of discovery is a significant challenge.

Transparency. Transparency refers to the clear reporting and dissemination of information related to a study's design, execution, analyses, and results, along with any limitations, sources of uncertainties, or sources of possible bias. According to the National Academies, transparency is fundamental to ensuring that others can understand research results and their implications.[13] In addition, transparency of methods and results is a prerequisite for efforts to replicate or build on previous results. Efforts to increase transparency have generally fallen under the "open science" movement, a term used to identify a set of

[11]For a detailed discussion of the ways in which concerns over how reproducibility and replicability of research can manifest across scientific disciplines and types of research, see the 2019 National Academies report.

[12]Marquart, F., "Methodological Rigor in Quantitative Research," *The International Encyclopedia of Communication Research Methods* (2017). https://doi.org/10.1002/9781118901731.iecrm0221.

[13]2019 National Academies report.

practices aimed at increasing public access to all aspects of the research process.[14]

Rigor and transparency in research increase the likelihood that research results are reliable. They help ensure that results are valid and can be understood, and that other researchers have confidence in interpreting the results and building on them. When researchers provide a clear, specific, and complete accounting of the materials and methods they used, the results they found, and the uncertainty associated with the methods and results, other researchers will know how to interpret the findings.

Efforts to ensure transparency and rigor can be undermined by poor research practices, which can result from time and resource constraints, lack of training, and other factors.

[14]The underlying theme of open science practices is that—with exceptions for privacy, intellectual property, or security considerations—all aspects of scientific research should be accessible and free to all.

Examples of Research Practices that Undermine Rigor and Transparency

Underpowered studies: Cost and time constraints can sometimes lead researchers to conduct studies with fewer subjects than would be ideal. Studies with too few subjects have low statistical "power." In low-powered studies, it is more difficult to rule out chance as the cause of an observed effect rather than the actual variable of interest. Experiments that fall into this category are more likely to deliver spurious results. For example, a researcher might conclude that an ineffective drug actually works because the study was based on too few subjects.

Measurement and other data errors: Researchers may make errors when collecting data, such as incorrectly measuring an outcome, failing to properly calibrate equipment, measuring results with a lack of precision, inaccurately recording measurements, or making errors in calculations. These types of errors can often be difficult to detect by the researcher or others trying to replicate the study.

Publishing pressures: Researchers face immense pressure to publish their research results, which creates incentives for researchers to look for an experimental effect (a "positive" result) when none actually exists. Because journals tend to favor publishing studies with positive effects, it can be difficult to publish a paper on an experiment that failed to find an effect (a "negative" or "null" result).

Inscrutable research data: Researchers may make research data, computer code, and other digital artifacts available but without the proper contextual information it can be difficult or impossible for others to understand or use.

Post-hoc hypotheses: According to academics who study research, it can be tempting for some researchers to develop a hypothesis after they have collected and analyzed data. This seemingly innocuous practice results in a greater likelihood that the study will report spurious results. This practice can come in the form of HARKing (hypothesizing after results are known) or "p-hacking" (manipulating data analyses to enable favored results to be presented as statistically significant). Thus, many actions to improve research rigor aim to have researchers clearly state research plans—particularly the hypothesis—before conducting the experiment.

Unblinded and nonrandomized studies: When studies are unblinded researchers know which participants are part of which group (e.g., who is getting a placebo). This approach can impart unintentional bias in the results. Similarly, when researchers do not randomize their study groups, they may impart unintentional bias when picking which participants should be part of each group.

Source: GAO review of relevant literature. | GAO-22-104411

Peer Review

Peer review is a process used in the scientific community to ensure that the quality of a published manuscript meets the standards of the scientific and technical community. Peer reviewers are specialists in the field of study who were not involved in carrying out the study or producing the manuscript. Peer reviewers generally assess the manuscript for characteristics such as

- the clarity of hypotheses,
- the validity of the research design,
- the quality of data collection procedures,
- the robustness of the methods employed,
- the appropriateness of the methods for the hypotheses being tested,
- the extent to which the conclusions follow from the analysis, and
- the strengths and limitations of the overall product.

Peer reviewers may suggest ways to clarify assumptions, findings, and conclusions. For instance, peer reviewers can help filter out biases and identify oversights, omissions, and inconsistencies and may encourage authors to more fully acknowledge limitations and uncertainties.

Peer review may take a variety of forms, depending upon the nature and importance of the product. For example, the peer reviewers may represent one scientific discipline or a variety of disciplines and the names of each reviewer may be disclosed publicly or remain anonymous to encourage candor. Editors of scientific journals use reviewer comments to help determine whether a draft scientific article is of sufficient quality, importance, and interest to a field of study to justify publication.

Federal Investment in Basic Research

The federal government is the largest funder of basic research in the United States. Federal expenditures for basic research, including defense agencies, were more than $42 billion in fiscal year 2019, or about 40 percent of the total U.S. funding for basic research, public or private. Among the civilian agencies, NIH, NSF, and NASA are the three largest federal funders of basic research. Each has a different federal mission and research focus related to advancing science and technology.

Each of these three agencies funds basic and applied research across their respective missions. NIH is the leading agency for funding basic research in life sciences, psychology, and engineering. NASA's funds basic research primarily in the physical and environmental sciences. Because of its broad mission, NSF has the widest portfolio of basic science funding, including engineering; computer sciences and mathematics; and the physical, environmental, life, and social sciences. (See table 1.)

Table 1: NIH, NSF, and NASA Obligations for Basic Research, by Scientific Discipline, Fiscal Year 2019

(Dollars in millions)

	National Institutes of Health (NIH)	National Aeronautics and Space Administration (NASA)	National Science Foundation (NSF)
Computer sciences and mathematics	$88.9	$ –	$939.2
Engineering	$1,102.4	$ –	$798.3
Environmental sciences	$176.2	$1,623.1	$908.5
Life sciences	$15,776.0	$43.6	$689.9
Physical sciences	$45.6	$4,008.9	$967.5
Psychology	$1,247.6	$ –	$36.2

(Dollars in millions)

	National Institutes of Health (NIH)	National Aeronautics and Space Administration (NASA)	National Science Foundation (NSF)
Social sciences	$58.2	$ –	$190.8
Other	$525.2	$13.3	$681.6
Total	**$19,020.0**	**$5,689.0**	**$5,212.1**

Source: NSF Survey of Federal Funds for Research and Development. | GAO-22-104411

Stakeholders in the Scientific Community

Researchers. Federal grants support the work of hundreds of thousands of researchers. Researchers carry out the studies, usually in teams, led by a principal investigator. In addition, researchers are generally responsible for applying for grants—which includes explaining the objectives of the research, outlining proposed methodologies and how research data will be managed, and providing information about researchers' qualifications to carry out the research—in accordance with agency grant requirements. Researchers are also generally responsible for getting their studies published and making sure manuscripts and research data are publicly accessible.

Nonfederal funders. Corporations, nonprofits, and private foundations also play an important role in funding basic research. Corporations, in particular, are the second leading funder of basic research in the United States, accounting for about $33 billion, or about 31 percent of funding dollars.[15] In addition, according to a 2013 report by the Science Philanthropy Alliance, about $2 billion a year goes to basic science from private foundations.[16]

Research organizations. Thousands of organizations engage in research funded by federal agencies. These include corporations; universities and colleges; state and local governments; and a variety of research institutions.

Professional societies. There are more than 250 professional societies and academies affiliated with the American Association for the Advancement of Science, according to the organization. Many of them

[15]Congressional Research Service. *U.S. Research and Development Funding and Performance: Fact Sheet*, CRS-R44307 (Washington, D.C.: Oct. 4, 2021).

[16]Science Philanthropy Alliance, *Funding Basic Research—The Foundation of Our Future* (Washington, D.C.: July 17, 2013).

publish subscription-based, peer-reviewed journals that focus on research relevant to their membership. They also advocate for their members through public and government affairs and provide professional development services, such as training workshops and collaborative conferences.

Publishers. Over decades, journal publishers have established themselves as an important resource in the dissemination of research findings and coordination of peer review. Publishers can be commercial entities or nonprofits, such as universities and professional societies, and they can be national or international. They can offer strictly online electronic journals, or in some cases paper copies, and can be open access or subscription-based with professional society membership fees. According to an estimate by the International Association of Scientific, Technical and Medical Publishers, about 33,000 active scholarly peer-reviewed English-language journals existed in 2018, publishing more than 3 million articles per year.[17]

Academic and independent research libraries. Libraries, both public and private, are essential stakeholders in the curation and archiving of digital artifacts, such as electronic documents, data, and code related to research. There are more than 3,500 academic research libraries in the U.S.

Journalists and public affairs specialists. Journalists and public affairs specialists play an important role as conveyors and disseminators of research to the public. These stakeholders, some of whom may have backgrounds in STEM, help disclose and present scientific information for general audiences, including policymakers.

[17]STM: International Association of Scientific, Technical, and Medical Publishers. *The STM Report: An Overview of Scientific and Scholarly Publishing*. 5th ed. (The Hague, Netherlands: 2018).

The Scientific Community Has Implemented Several Strategies for Promoting Rigor and Transparency, but Various Factors Discourage Their Widespread Use

The scientific community has developed a variety of strategies to promote more rigor and transparency in the design, execution, analysis, and reporting of research results. However, according to our interviews with stakeholders across diverse disciplines and our review of relevant literature, there are various factors, including misconceptions and poorly aligned incentives, that discourage the use of these strategies and their associated tools.

Strategies Involve Several Tools and Practices to Increase Research Rigor and Transparency

The scientific community has developed and employed a variety of tools and practices to help increase research rigor as well as transparency in the design, execution, analysis, and reporting of their research results.

Preregistration of studies and registered reports to disclose proposed study design. Preregistration of studies and registered reports are tools meant to promote research rigor and transparency by having researchers make certain aspects of the research process—such as hypotheses, planned methodology and data collection strategies, and planned statistical analyses—accessible to others before any research is conducted. Preregistration allows researchers to publicly post their descriptions of their research design in a registry before the research is conducted and results are reported. Describing hypotheses, methodologies, and analysis approaches up front, and publicly posting them, allows others to understand and comment on the study design, which can sometimes result in changes to those plans, and can reduce researchers' ability to misuse statistics, among other benefits. Any subsequent deviations from the specified plan are publicly logged.

Registered reports takes preregistration to another level, according to stakeholders. Similar to preregistration, research plans are registered publicly—but in conjunction with a journal publisher. With registered reports, research proposals and plans are submitted to a publisher prior to data collection, with the assurance that the research results will be published regardless of the study's outcome. Publication is assured as long as the researcher follows the registered plan, or changes to those plans were justifiable and reasonable. More than a dozen governmental and nongovernmental entities worldwide, including NIH and the World

Health Organization, host registries for preregistration of studies, and more than 300 journals accept registered reports.[18]

Preprints to share research results before journal publication. Preprints are public drafts of scientific research manuscripts that have not yet undergone the peer review process that generally precedes publication in a peer-reviewed journal. Preprints improve transparency because they are posted on public servers and provide public access to research results through earlier dissemination of research findings. Preprints also enhance rigor because they allow the author of the manuscript to incorporate feedback from other researchers and take other steps to improve the research, such as collecting additional data, and repeating or redoing analyses. As of April 2022, more than 80 preprint repositories exist worldwide, operated by academic centers, research institutions, and non-federal funders.

Data repositories to share research methods and data. Data repositories are centralized locations where research data, code, and other digital artifacts such as audio and visual media are stored, archived, and made freely available to the research community and the public. Data repositories, which may differ by scientific discipline, are generally managed by data curation personnel—often staffed by research libraries—who ensure that research data are preserved long-term. A wide range of organizations maintain repositories, including the federal government, research institutions, non-profit entities, and for-profit companies. As of March 2022, there were more than 2,000 governmental and nongovernmental data repositories worldwide. In 2013, OSTP issued a memo to heads of executive departments and agencies that called for increasing public access to the research and scientific data resulting from federally funded scientific research.[19] We have previously reported that federal and non-federal entities may maintain repositories where datasets

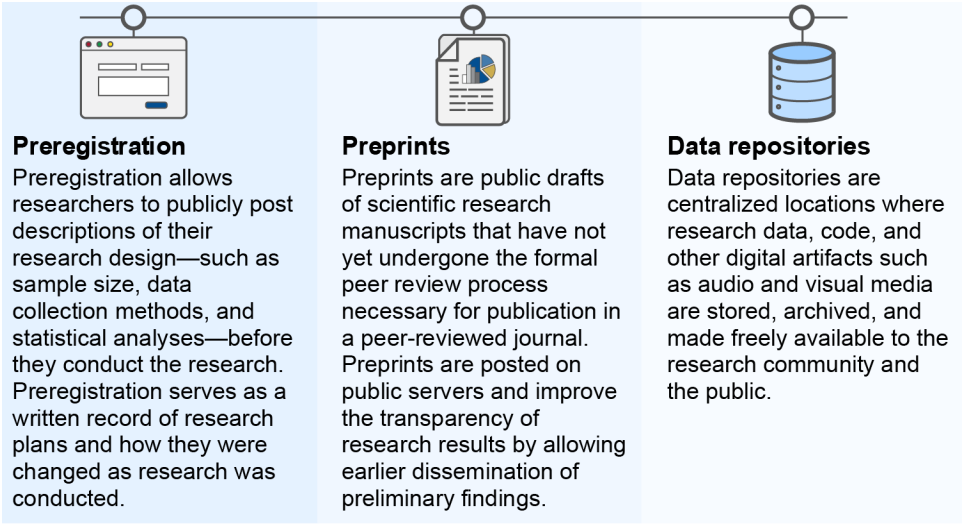
[18]For additional discussion of preregistration and registered reports, see Nosek, B., C. Ebersole, A. DeHaven, and D. Mellor, "The Preregistration Revolution," *Proceedings of the National Academy of Sciences,* 115, no. 11 (2018) https://doi.org/10.1073/pnas.1708274114, and Chambers, C., L. Tzavella, "The Past, Present, and Future of Registered Reports," *Nature Human Behaviour,* 6 (2021) https://doi.org/10.1038/s41562-021-01193-7.

[19]Office of Science and Technology Policy, *Increasing Access to the Results of Federally Funded Scientific Research,* Memorandum (Washington, D.C.: Feb. 22, 2013).

are stored along with metadata to ensure the public can find and use data resulting from federally funded research.[20]

A summary of these tools and practices is presented in figure 2.

Figure 2: Tools and Practices to Increase Rigor and Transparency



Source: GAO analysis of information from the National Academies of Sciences, Engineering, and Medicine. | GAO-22-104411

Improving researcher training on these tools and other topics is also important for increasing research rigor and transparency, according to stakeholders. Federal agencies, private academic and research institutions, and professional societies offer instruction on tools to help improve rigor and transparency in research, including awareness and understanding of tools designed to improve research. Training includes topics such as study design, statistical analysis, steps to minimize researcher bias, the importance of transparency and open science, as well as how to use the tools designed to improve rigor and transparency.

Various Factors Discourage Greater Use of Tools to Increase Rigor and Transparency

According to stakeholders we spoke with and our review of relevant literature, a variety of factors discourage the scientific community's efforts to increase the use of preregistration and registered reports, preprints, and data repositories to increase the rigor and transparency of research. Among these factors are concerns that use of the tools will add to

[20]See GAO, *Federal Research: Additional Actions Needed to Improve Public Access to Research Results*, GAO-20-81 (Washington, D.C.: Nov. 21, 2019).

researchers' administrative workload, researchers' misconceptions about the tools, and misaligned incentives.

Concern over administrative workload. Some researchers may hesitate to preregister their work, publish findings as preprints, or submit their research data to repositories in a comprehensive way because of concerns that this will require more time and resources. For example, stakeholders told us that researchers already feel pressed for time to conduct their research and are likely burdened with other administrative tasks. Preregistering a study adds even more work and time before they are able to begin their research. Stakeholders we spoke with also noted that using the registered reports model can also add significant up-front work—sometimes taking weeks to several months—to have a research plan processed and reviewed. Other logistical challenges that can add more work and time include navigating multiple websites to locate an appropriate data repository or preprint server; gathering all digital artifacts from a study; correctly formatting data; and completing other related tasks to ensure study information is findable, accessible, and usable.[21] Stakeholders said that this additional work may make researchers reluctant to use the tools unless it is clear the benefits outweigh the additional costs in time and resources.

Researcher misconceptions. Some researchers are not fully aware of or do not fully understand some tools and their uses for improving rigor and transparency. For example, according to stakeholders and relevant literature we reviewed, researchers may not associate the use of tools such as preregistration and registered reports with rigor and transparency. They said that this lack of association may be partly because there is little, if any, awareness and understanding of the benefits of those practices, especially at the graduate level. In other cases, researchers may understand why transparency is important, but they may not understand that to achieve research transparency, data and information must be shared in a manner that ensures they are findable, accessible, and usable. For example, stakeholders explained that researchers should consider critical decisions before depositing their data and information, such as whether to include data not presented in final

[21]One set of data management principles that have been developed is the FAIR guiding principles for scientific data management and stewardship. FAIR refers to the Findability, Accessibility, Interoperability, and Reusability of digital information. For additional information, see https://www.go-fair.org/fair-principles/. For the purposes of this report, we use the phrase "findable, accessible, and usable" to refer to practices consistent with the FAIR principles.

publication and how much relevant information is necessary to disclose for the scientific community to understand the steps taken to reach the final reported result.

Furthermore, stakeholders also told us that researchers can be hesitant to use some tools because of misconceptions and may think that the benefits are not worth the costs. For example, some researchers may hesitate to register a study plan because of the upfront work and time spent prior to conducting the research, as mentioned above. However, stakeholders said that because the process usually involves working with scientific journals, researchers can save time during the pre-publication peer review process. In addition, they explained that a common misconception about registered reports is that researchers will lose the ability to change planned research protocols or analyses, when in fact the ability to do so is built into the process. Similarly, according to stakeholders we interviewed and our review of relevant literature, researchers may fear that preregistering their study plan or posting their research manuscript on a preprint server may result in their findings being claimed by another researcher as their own or jeopardize their intellectual property rights. Yet stakeholders we interviewed also stated that research information and data posted electronically online would have time stamps or digital object identifiers that help to permanently track researchers' work and thus further ensure their claim to the research.[22]

Challenges in navigating repositories and servers. According to stakeholders and our review of relevant literature, researchers can find it difficult to navigate the many preprint servers and data repositories that exist around the world. Specifically, researchers may find it daunting to locate which server or repository is most appropriate. Stakeholders explained that with the large number and variety of data repositories and preprint servers to choose from globally, it can be difficult for researchers to ensure they will successfully make their study data findable, accessible, and usable.[23] For example, one stakeholder told us that

[22]Researchers may obtain a Creative Commons license to maintain copyright protection while permitting distribution in an open access platform. The Creative Commons licensing framework provides individual researchers and institutions a standardized method to give the public permission to reuse, within limits, an author's work under copyright law while preserving the author's copyright.

[23]We previously reported on the extent to which federal agencies have implemented plans to help the public locate stored data and digital artifacts from federally funded research. Among other things, we concluded that the diverse landscape of discipline-specific data repositories can make it challenging to access or analyze data sets stored across multiple repositories. GAO-20-81.

researchers sometimes use the wrong data repositories for the type of research data they have generated, even when the research is tied to federal grants that require the data to be placed in a specific repository. Furthermore, stakeholders noted many research libraries are struggling with finding the resources and staff who have the expertise to assist researchers with drafting data management plans and properly formatting and depositing research data into repositories. In addition, stakeholders we interviewed stated that although some preprint servers have included platforms for public commentary, most feedback on preprints is exchanged through private email or social media and, therefore, are usually unavailable to the broader scientific community and public.

Lack of uniform standards and accepted best practices. Effective use of some tools is hindered by a lack of uniform standards, according to stakeholders we interviewed and our review of relevant literature. The scientific community has developed various standards and best practices for preregistration, preprint services, and data repositories. These standards and best practices specify, among other things, how researchers use the services, how data and information are preserved, and how digital materials are identified and linked. However, because different standards and best practices have been developed by the research community, it is not always clear which standard or best practice applies to different servers or repositories, according to stakeholders. Stakeholders added that because there are so many different standards and best practices for preprint servers and data repositories, these standards can become inscrutable when a researcher is trying to determine which server or repository to use. Further, stakeholders noted that researchers sometimes do not comply with requirements to use certain repositories as specified in federal research grants, or may bypass best practices that can help ensure their research data and information are understandable and stored in a freely accessible repository. We previously reported that, according to agency officials, the absence of common standards for data repositories poses a challenge to ensuring ongoing access to data and results from federally funded research.[24]

Misaligned incentives for researchers. Stakeholders and relevant literature noted that competing interests and incentives are not always aligned with good research practices. Specifically, according to stakeholders, transparency and research rigor are not in practice as much as they should be because the culture within the research community is

[24]GAO-20-81.

focused on getting research published. They said that within the scientific community publication in a notable peer-reviewed journal is highly prized, as it can help to advance a researcher's career and create more funding opportunities. As a result, there is less incentive for researchers to take the time to use tools such as preregistration or preprints. For example, stakeholders noted that researchers are not compensated for preregistering their research but are instead rewarded for publishing papers. Similarly, they added, when putting data into a repository there is little incentive for researchers to take the time to properly format data and ensure all digital artifacts are included and that associated methodologies are clearly explained so that other researchers can understand the research process and conclusions. While these tools can go a long way to strengthening rigor and transparency, stakeholders said that researchers may not adopt them if they view them as meaningless practices not worth the added time and resources.

Resistance to changing research norms and practices. Researchers can be resistant to adopting new tools aimed at improving rigor and transparency. In particular, stakeholders mentioned that although more researchers are sharing data they have historically not been asked or required to do so, and that preregistration does not prevent publication bias because researchers can still decide not to publish their negative results.[25] Furthermore, stakeholders stated that there is still debate, particularly within the life sciences, on whether preprints are considered legitimate contributions to the scholarly record. They also said medical journal publishers have been especially cautious about preprints because they are not peer viewed. One stakeholder explained that although preprints can provide insight into how the research was carried out and its outcomes, the scientific community values publication of the final research results that have been peer reviewed. In other words, until research findings posted to a preprint server are recognized and vetted in the same way as those published in a reputable peer-reviewed journal, researchers may remain hesitant to adopt preprints as a regular part of their research practice, stakeholders explained.

[25]Publication bias occurs when researchers only report and publish positive research findings, particularly in prestigious journals, but not research results that are null or negative. According to the 2019 National Academies report, such actions exclude research that could still be meaningful to the scholarly record and can instill bias in the published literature about the research itself.

Wide range of publishing practices. According to stakeholders and our review of relevant literature, the evolving publishing landscape presents challenges to ensuring research rigor and transparency, in part because publishers' peer-review practices, standards for accepting manuscripts, and public access policies vary widely. For instance, according to stakeholders, there are journals that have minimal, if any, standards for peer review, thus allowing ample opportunity for publication of non-rigorous research, among other things.[26] Other journals implement thorough peer-review standards—sometimes running the code with the available data to reproduce the study's publishable results—and require authors to fill out a submission checklist to ensure compliance with editorial policies. Similarly, according to stakeholders, journal standards—especially those of open access journals—for ensuring that the research is publicly accessible, including the manuscript and associated data and digital artifacts, can vary widely.[27] As a result, according to stakeholders we interviewed and our review of relevant literature, as open-access journals become more commonplace, some in the scientific community have expressed concern that some open-access journals and their associated data in repositories may be difficult to locate online and are not properly managed. Further, some open access journals may not be a viable, long-term solution to preserving scientific information and data as they may lack adequate financial support. For example, researchers identified more than 150 open access journals worldwide that ceased

[26]According to researchers, the consensus among the scientific community is that "predatory journals and publishers are entities that prioritize self-interest at the expense of scholarship and are characterized by false or misleading information, deviation from best editorial and publication practices, a lack of transparency, and/or the use of aggressive and indiscriminate solicitation practices." See Grundniewicz, A., D. Moher, K. Cobey, et al. "Predatory Journals: No Definition, No Defense." *Nature,* 576 (2019) https://doi.org/10.1038/d41586-019-03759-y.

[27]The two primary methods of offering freely available, digital information online are through "green" or "gold" open access. Green access refers to digital information that is self-archived prior to publication via an institution's repository, such as the case with preprints or when researchers post their published papers on their university's website, with permission by the publisher. Gold access refers to digital information published generally after the author pays an article processing charge to have research results made freely available through an open-access journal, for instance. Some traditional subscription-based journals have adopted hybrid models that have some freely available articles and other articles hidden behind paywalls that require a fee for reader access.

operations between 2000 and 2019, with published research no longer available.[28]

Difficulty ensuring researchers have needed training. As mentioned above, the scientific community has taken steps to provide additional researcher training in areas such as rigorous study design, strategies for avoiding researcher bias, the proper use of statistical analysis and large data sets, and data sharing techniques. However, stakeholders explained that it can be difficult to provide such training and determine how to deliver it across different scientific disciplines with different training needs. For example, one stakeholder we interviewed noted that researchers often do not have the proper training to fully understand how statistical concepts should be applied and how the results should be interpreted. Stakeholders added that a long-term strategy and significant resources will be needed to address such training issues.

Experts Identified a Wide Range of Actions for Federal Agencies to Improve Research Rigor and Transparency

Leading experts we interviewed and others from our roundtable discussions expressed the view that federal agencies, as funders of approximately 40 percent of the nation's basic science research, could take specific actions to enhance research reliability. Experts stressed the need to better align recognition for research, and researcher awards and promotions, with practices that promote rigor and transparency. To that end, they proposed a variety of measures for federal agencies to consider, while noting that differences in scientific fields must be considered along with the need to avoid federal requirements that are unnecessarily burdensome.

Experts Suggested Ways to Modify Grant Processes

Experts discussed a number of ways agencies might modify their grant processes to promote both rigor and transparency. In particular, they suggested ways to offer incentives and to modify grant application requirements to signal an agency's focus on good practices. Acknowledging that agencies would need to weigh the costs and benefits of new measures in light of differences in research goals, scientific disciplines, and types of research, experts proposed the following for consideration:

[28]Laakso, M., L. Matthias, and N. Jahn. "Open is Not Forever: A Study of Vanished Open Access Journals," *Journal of the Association for Information Science and Technology* (2021) https://doi.org/10.1002/asi.24460.

- Allow applicants to cite prior preprints, or research data and materials, as part of the grant application, which could signal that these are valuable and accepted, and provide an incentive for their use.
- Incorporate open science language in grant application materials and progress reports to let applicants know the agency recognizes the importance of rigor and transparency in the proposed research.
- Require researchers to explicitly address, in grant applications, how they will ensure rigor and transparency in executing the planned research.
- Require institutions that receive grant funding to comply with certain practices, such as open access publishing, publication of null results, preregistration of study protocols, and use of Findability, Accessibility, Interoperability, and Reusability (FAIR) open data principles.
- Eliminate the use of journal impact factor as a merit factor when evaluating grant applications,[29] and have researchers omit journal names from grant proposals to lessen their consideration.
- Encourage or require, as appropriate, researchers to engage with a trained methodologist, statistician, research librarian, or other specialist early in the research process to ensure they address issues with methods, planned analyses, and data management before the research plan is executed.
- Provide specific funding within the grant to ensure that study results—including manuscripts and research data, methodology, and analyses—are reported in an open, transparent fashion.
- Transition, as appropriate, to a requirement that grantees preregister their work and publish manuscripts as preprints to maximize research transparency.
- Encourage the publication of null and negative results, and recognize null research results and replication studies as valuable contributions to research.

Experts Proposed Several Additional Strategies for Improving Research Practices

Use pilot projects for incremental changes tailored to scientific disciplines. Experts from our roundtable discussions stated that broad cultural acceptance takes time, and suggested that an incremental approach using pilot projects and flexible mandates with liberal exceptions could allow federal agencies to learn about barriers to

[29]Journal impact factor is a measure of the number of times an average paper in a journal is cited during a year. Impact factor is frequently used as an indicator of the importance of a journal to its field.

implementing new requirements and to develop alternatives. (See sidebar.) Experts emphasize, however, that funding agencies should develop policies and implement actions based on the specific circumstances of each research program or scientific discipline. For example, experts from our roundtable discussion cautioned that some tools, such as preregistration, should be tailored to the needs of each specific discipline.

Pilot Program for Registered Reports

With a registered report, a researcher's study proposal, including background, study design, methodology, and analysis plan is peer reviewed before the research is undertaken. If the study proposal is accepted by the journal, the study's results will be published, pending a review to verify that the study plan was followed. Experts suggested that agencies would benefit from establishing a registered reports pilot program because of the complexities and up-front costs for setting up and administering this preregistration model. These costs and complexities stem from the need to set up formal arrangements with journals and the upfront effort by researchers to document their study plan in sufficient detail and make changes based on the review.

An agency could begin a registered reports pilot program by choosing a subset of studies, perhaps limited to a specific discipline or sub-discipline. The agency would then identify leading journals publishing work in that discipline, and work with the journals to set up a trial partnership and establish an agreement on how they would carry out the study plan review. Agencies would also determine how much additional funding, if any, is necessary to the grantee for upfront documentation and planning, and for other administrative requirements. Agencies would also plan for an ongoing assessment of the pilot program so adjustments could be made before the program was expanded, as appropriate.

Source: GAO review of information from stakeholders and relevant literature. | GAO-22-104411

Increase compliance monitoring. Experts also expressed the view that pilot programs, as well as new mandates and other agency interventions, should be closely monitored for effectiveness, which will allow the agency to make necessary adjustments before they are expanded. Experts from our roundtable discussions also emphasized the importance of monitoring compliance with new requirements, and cited data sharing requirements as a key example. Without this monitoring, the agencies cannot be sure the actions are having intended effects on rigor and transparency.

Develop standards and best practices for using preregistration, preprints, and data repositories. Experts told us that agencies could also take steps to develop clear, understandable standards and best practices for researchers in using preregistration, preprints, or data repository services to encourage those practices. Such standards could help with ease of use of these services and increase compliance with requirements, given the proliferation of preprint services and data repositories that can make it difficult to know which are credible and reliable. Experts stated that researchers are more likely to comply with new requirements, or implement other practices by choice, if doing so is easy rather than adding significant burden. For example, agencies could develop a clear set of standards for documentation (data, code, and supporting information) that researchers should create and post publicly when releasing a preprint or publishing a paper. Similarly, if an agency coordinated with journals on implementing a registered reports model, the process for a researcher choosing to use that model could be a simple choice when a grant application is accepted. In addition, experts suggested that agencies themselves could take steps to curate preregistration, preprint, and data repository services for accepted practices, making it easier for researchers to identify these services.

Boost support for researcher training. Experts identified researcher training as integral to any strategy to increase rigor and promote open science principles. Experts discussed the need for additional training in research design and the proper use of statistics, as well as training in data sharing and transparency. Experts agreed that funding agencies

should collaborate and coordinate training initiatives so there is some consistency in the training, but also tailor training to specific disciplines, given that specific training needs vary among the scientific disciplines. In addition, experts suggested that any agency initiative or requirement addressing rigor and transparency should include funding for training to build capacity and researcher buy-in.

As a long-term undertaking, experts agreed that developing and providing this training would require agencies to take steps to ensure that training on open science, reproducibility, and reliability is integrated into the regular academic curriculum. They added that any federally funded researcher training must focus on open science practices, research integrity, peer review, and responsible research evaluation. In terms of curriculum, they also discussed the need for a bottom-up approach, with input from various parts of the research community, to increase the likelihood of broader acceptance. In addition, experts stated that federal agencies could play a role in curating and identifying the vast amount of training resources that already exist, and pointing researchers to those resources.

Reprioritize agency funding. The experts from our roundtable discussions stated that federal funding agencies may need to reprioritize and reallocate their limited resources to foster greater rigor and transparency of research, such as through compliance monitoring, pilot programs, standards development, and additional researcher training as noted above. According to these experts, redirecting a small percentage of funding may result in less money directed to research and thus fewer funded studies, but could ultimately result in studies that are more high-power, rigorous, and transparent. Also, while experts agreed that funding replication studies on a wide scale may not be necessary, there may be circumstances whereby replication studies may be called for, either to validate a new finding or to serve as a "spot check" on an agency's quality control process.[30] Or, they said that agencies might fund larger, more powerful studies in which two or more independent teams would address the same research question (either adversarial or complementary). In addition, according to experts, agencies could fund studies on what affects the quality of research, such as larger sample sizes and methods to prevent bias, or fund research that evaluates the impact of interventions in research culture. Additionally, agencies could create a

[30]Experts suggested that some replication studies could be carried out by graduate students, which would serve the dual purpose of spot-checking research results and teaching young researchers about the value of scientific rigor and transparency.

suite of funding mechanisms and a portfolio of grants with the specific goal of increasing the reliability of research.

Agencies Have Taken Steps to Improve Research Reliability, but Lack Information to Assess Research Rigor and Transparency

We found that NIH, NSF, and NASA have taken some steps to promote and support additional rigor and transparency in the research they fund. These agencies largely rely on the grant application review process for vetting grant proposals and the prepublication peer review process to ensure rigor, and they rely on researcher compliance with data management requirements to ensure transparency. However, we found that the agencies do not assess the rigor and transparency of the research they have previously funded to consider strategies for improvement. As a result, agencies lack information that could inform or suggest changes to the grant making process and research funding priorities.

NIH, NSF, and NASA Have Taken Some Actions to Promote Rigor and Transparency

Over the past decade, NIH, NSF, and NASA have taken, to varying degrees, steps to improve rigor in the research they fund and to assure the availability of their results and data. These actions were taken, in part, in response to the 2013 OSTP memorandum on increasing public access to federal research and increasing awareness of persistent challenges to the reproducibility and replicability of research. The agencies have each established requirements for grantees to disclose their research and associated data publicly, and raised awareness of the need for more rigor and transparency by posting information and resources for addressing reliability on agency websites. This information includes relevant guidance for grant applications, information about agency open science initiatives, and links to training sources.

Notably, NIH made changes in 2015 to require that—beginning in January 2016—grant applicants provide an explicit discussion and evaluation of the rigor of the relevant prior research, how they intend to address any weaknesses therein, and how their proposed experimental design and methods will achieve robust and unbiased results.

These agencies have also taken some action with regard to the use of preprints, preregistration, data management, and researcher training.

Preprints. In June 2020, NIH launched a preprint pilot program focused on NIH-funded SARS-CoV-2 and COVID-19 research. The pilot, which is ongoing, makes preprints for findings about SARS-CoV-2 and COVID-19 available in NIH's publication databases. NIH officials we spoke with said that posting preprints of these research manuscripts is intended, in part,

to test the possibility of making them searchable and discoverable. Accordingly, the pilot will also allow the agency to explore approaches for increasing the discoverability of early NIH research results and, according to NIH officials, further establish the basis for NIH's current guidance that encourages researchers to use such interim products to speed dissemination and enhance the rigor of their work.[31] To encourage researchers to post their manuscripts to preprint servers, NIH has also begun allowing potential grantees to cite interim research products, including preprints, in their applications. NSF and NASA do not have specific policies or guidance on the use of preprints, according to officials from each agency, though they do not prohibit their use. NSF and NASA officials cited the lack of quality control over the content of preprints as a concern, and neither indicated that a preprint policy would be forthcoming.

Preregistration and registered reports. NSF and NASA have no policies on preregistration although NSF officials stated that preregistration can be useful for certain types of research, such as randomized control trials. NIH policy, on the other hand, establishes the expectation that all investigators conducting clinical trials funded, in whole or in part, by the NIH will ensure that these trials are registered at ClinicalTrials.gov, and that results of these trials are submitted to ClinicalTrials.gov.[32] Disseminating this information supports the NIH mission to advance the translation of research results into knowledge, products, and procedures that improve human health.

Requirements for public access and data management. NIH, NSF, and NASA have public access and data management policies to help ensure that research results are made publicly available in a timely manner and sufficiently preserved over time. Each agency's policy on providing public access includes guidance for researchers to develop a data management plan outlining how the researchers plan to make results as well as information such as methodologies, data, and code available to the public. Each agency's policy is consistent with OSTP

[31]According to NIH, the agency is monitoring the pilot program for effects on how and when research results are shared, discovered, disseminated, and reported. The agency is also monitoring for evidence of increased researcher awareness of preprints and for the emergence of best practices for preprint sharing.

[32]Applicable drug and device trials are generally required to be registered at clinicaltrials.gov under 42 U.S.C. § 282(j), irrespective of their funding source. For more information, see https://clinicaltrials.gov/ct2/home.

guidance requiring that manuscripts generally be made publicly accessible within 12 months of the publication date. The agencies' policies also meet the OSTP mandate for requiring that grantees generally develop plans for the long-term preservation and availability of their data and results. Because of the wide range of research, these data management policies do not require that data be shared in a specific type of repository, be formatted in a specific way, or outline which specific research data and digital artifacts need to be stored.[33] NIH provides guidance to researchers on how to select a suitable data repository, while NSF and NASA are in planning stages of developing such guidance.

More recently, each of the three agencies have begun to take additional steps to support open science and transparency. For example, in October 2020 NIH published an updated data management policy for its grant recipients, due to take effect in January 2023. NIH officials told us that requirements will be further clarified, but that they plan to encourage researchers to use data management and data sharing practices consistent with the FAIR data management principles. NASA is similarly developing a new data management and sharing policy through its Open-Source Science Initiative. Agency officials described that initiative as a long-term commitment to building an inclusive open science community through what they described as "a comprehensive program of activities to support moving science towards openness." In addition, NSF recently accepted proposals, due by April 12, 2022, for studies intended to address wide-ranging needs in the scientific communities around FAIR data management principles and best practices for open science. According to agency officials, findings from these studies will inform NSF's public access policies.

Training. NIH has taken several new steps in recent years to improve and increase access to researcher training related to rigor and transparency, including changing requirements for some training grants and developing training modules for researchers. Specifically, in 2020 NIH changed grant requirements for institutional research training and for institutional career development to require that applicants describe how the program and faculty will provide training in rigorous research design and relevant quantitative and data science approaches. Similarly, NIH changed requirements for individual career development grants to require

[33]The 2013 OSTP memorandum states, among other things, that "to the extent feasible…digitally formatted scientific data resulting from unclassified research supported wholly or in part by federal funding should be stored and publicly accessible to search, retrieve, and analyze."

applicants to address any new research skills they plan to acquire in the areas of rigorous research design, experimental methods, quantitative approaches, and data analysis and interpretation. In addition, NIH has placed several training modules focused on rigor and transparency on the agency's website. One set is focused, for example, on randomization, sample sizes, outliers, and other topics.[34] NSF officials did not identify any recent changes to its training policies or practices related to rigor and transparency. NASA officials have stated that the agency perceives no serious training deficiencies among its grant recipients. They noted that grant applicants must demonstrate their research qualifications in training and education, and that these qualifications are evaluated as part of the grant application review process.

NIH, NSF, and NASA Do Not Collect Data to Assess the Rigor of Agency Funded Research or Transparency of Research Results

However, we found that NIH, NSF, and NASA do not collect data to assess the methodological rigor and transparency of the proposed research or the research they fund once the experiments have been conducted and outcomes published. As a result, the agencies lack information to better position them to address the broader issues raised by prior replication studies or information that might better inform agency decisions in funding future research.

An August 2021 joint memorandum from the Office of Management and Budget and OSTP stated that federal agencies "should prioritize making federally funded R&D…more rigorous, reproducible, and transparent" which would help to build a trustworthy and engaged U.S. science and technology enterprise.[35] Furthermore, the memorandum stated that federal agencies should ensure that research results are made widely available to other scientists and the public in a findable, accessible, and usable way to achieve transparency. This memorandum complements the 2013 OSTP memorandum to federal agencies, which stated that the results of federally funded scientific research—including peer-reviewed publications and digital data—should be made available to and useful for

[34]https://www.nigms.nih.gov/training/pages/clearinghouse-for-training-modules-to-enhance-data-reproducibility.aspx.

[35]Office of Management and Budget & Office of Science and Technology Policy, *Memorandum for the Heads of Executive Departments and Agencies: Multi-Agency Research and Development Priorities for the FY 2023 Budget*, Office of Management and Budget Memorandum M-21-32 (Washington, D.C.: Aug. 27, 2021).

the public, industry, and the scientific community.[36] Furthermore, *Standards for Internal Control in the Federal Government* calls for agency management to use quality information to achieve agency objectives and calls for agency management to identify and analyze risks to program objectives.[37] In this case, this includes information on the rigor and transparency of the research these agencies fund and whether that research is carried out with sufficient rigor and that results are accessible and useable by others.

Assessing research rigor. Officials from NIH, NSF, and NASA emphasized that it would be time consuming and resource intensive to attempt to replicate or reproduce the work underlying the thousands of published manuscripts each year. In addition, officials from these agencies maintained to us that the grant application review process and the pre-publication peer review process are adequate to ensure appropriate rigor.[38] They noted, in particular, that because the agencies fund only the most promising grant applications (fewer than about 25 percent, on average), there is reason to be confident that the planned research will be rigorous and reliable. Although a study's methodology and analysis plans are reviewed before a grant is awarded, changes to a study methodology and analysis by researchers are common. Further, the agencies' annual grant performance reviews do not include an assessment of how the methodology was implemented or an assessment of the study's analysis or results. Moreover, while agency officials said they depend on the pre-publication peer review for an independent assessment of research methodology, analyses, and conclusions, the concerns raised in the 2019 National Academies report regarding rigor involved peer-reviewed and published work. Pre-publication peer review varies in thoroughness, and important indicators of rigor are not

[36]The 2013 memorandum directed certain agencies to develop plans to support increased public access to the results of research funded by the federal government in ways that maximize the impact and accountability of the federal research investment.

[37]GAO, *Standards for Internal Control in the Federal Government*, GAO-14-704G (Washington, D.C.: September 2014).

[38]NIH officials told us that the agency indirectly assesses the quality of its funded research by tracking how often published studies are cited in future research and how often funded research is not published. According to agency officials, numbers of citations is a proxy measure of how useful a study was in informing future studies, and tracking the number of unpublished studies is a measure of how many studies might have been less rigorously designed and executed and ultimately of lower quality. However, neither of these assessments directly assess research rigor.

necessarily assessed by reviewers before research results are published.[39]

Assessing Rigor in a Scientific Sub-Discipline

A 2018 study by a team of independent researchers provides an example of what a research assessment on rigor might look like. The study assessed almost 3,400 NIH-funded preclinical peer-reviewed cardiovascular studies for characteristics of replicability.

The 2018 study found methodological shortcomings in what the study deemed "crucial" design elements that would help ensure the studies were methodologically sound and rigorous. Specifically, the study found low rates of randomization in 21.8 percent of studies, blinding in 32.78 percent of studies, and sample size estimation in 2.3 percent of studies. NIH officials told us that these findings are valuable, and that they agency is reviewing the findings to inform potential programmatic changes associated with how it funds and oversees certain types of pre-clinical (animal) research.

Source: F. Daniel Ramirez, Pouya Motazedian, et al., "Methodological Rigor in Preclinical Cardiovascular Studies: Targets to Enhance Reproducibility and Promote Research Translation" *Circulation Research,* 120, no. 12 (2017) https://doi.org/10.1161/CIRCRESAHA.117.310628. | GAO-22-104411

We found that NIH, NSF, and NASA do not carry out any manner of retrospective reviews or sampling of completed research to look for recognizable indicators of rigor, a process that would not require the expense and time required for replicating individual experiments.[40] While indicators of rigor vary with scientific disciplines, many have commonalities across scientific disciplines, such as a study's statistical power or sample sizes; the methods used to reduce measurement error; the use of randomization in sampling design; the extent to which researchers have adhered to the research methodology; or the soundness of statistical analyses. In some scientific disciplines, indicators of rigor may be quite different. For example, in experiments dependent on data from one-time natural events, the rigor of research design may stem from redundant data collection or finely calibrated instrumentation.[41] In addition, in studies of this type, rigor may be indicated by the processes

[39]Research methodology, data collection, analyses, and conclusions are generally reviewed during peer review prior to publication in a journal. However, some aspects of rigor—such as a small sample size—would not be possible to change without gathering additional data, for example, and would not necessarily prevent a reviewer from recommending publication. Other indicators of rigor may be difficult for the reviewer to understand or comment on, and some errors may not be caught during peer review, depending on the thoroughness of the review.

[40]Studies that assess reproducibility and replicability can be time consuming and resource intensive. For example, a recent study that aimed to repeat 193 experiments from 53 cancer biology papers was only able to repeat 50 experiments from 23 papers. Still, the effort took researchers approximately 7 years to complete and the costs per study were more than double initial estimates.

[41]For example, the Laser Interferometer Gravitational-Wave Observatory (LIGO) focuses on direct detection of gravitational waves produced from black holes or exploding stars in distant galaxies. LIGO has two separate, redundant detectors, each of which must be calibrated precisely, and its scientists collaborate with "sister" facilities in Germany and Italy.

by which data were collected, processed, and analyzed.[42] These agencies could collect or assess these indicators of rigor routinely or as part of a periodic audit for a sample of funded research, or fund assessments of rigor through academia or the nonprofit sector. (See sidebar.)

Assessing transparency. We found that NIH, NSF, and NASA do not systematically assess the transparency of research results, methods, and data. Officials from all three agencies told us their agencies assess compliance with public access and data management requirements. However, these assessments are not done systematically for the broader purpose of ascertaining an overall level of transparency.[43] None of the agencies systematically review compliance with public access and transparency requirements by, for instance, compiling data across the agency, assessing trends, or assessing causes of non-compliance. More importantly, the agencies do not collect information to understand whether research results and associated data are easy to locate and access, whether research data and related study information are in understandable or usable formats, or whether published studies have adequate supplemental information about methods or analyses that would aid in any potential replication effort.

As with assessments of rigor, agencies could assess the transparency of research results using a sample of published studies, and assessing whether study data and methods are findable, accessible, and usable. Simpler assessments might look at the extent to which data repositories used by researchers follow accepted best practices or accepted data management principles.[44]

[42]In some cases, where indicators of rigor are less clear, agencies might assess studies for evidence of questionable research practices, such as p-hacking or HARKing, or whether researchers collected additional data after initial analyses or stopped data collection earlier than planned.

[43]We previously reported on federal agencies' actions to comply with OSTP's 2013 memorandum on public access to research results, including NIH, NSF, and NASA. The report assessed the extent to which agencies had developed public access and data management plans and whether or not the agencies had compliance mechanisms in place. The report did not assess whether agencies were monitoring the extent of research transparency. See GAO-20-81.

[44]For example, the FAIR principles, as discussed previously. See www.go-fair.org/fair-principles for more detail.

Without taking steps to assess indicators of rigor and extent of transparency for agency-funded research, selected agencies may not have the necessary information to take the most effective steps to improve rigor and transparency. In this case, an understanding of the rigor and transparency of research may inform how these agencies fund future research, including any additional requirements for researchers. It may also inform agency efforts to develop pilot programs, increase researcher training, develop standards, and understand the extent that the agencies may need to reallocate resources to these actions. In addition, without assessing the extent to which research methods, data, and results are findable, accessible, and usable, selected agencies do not have an understanding of whether the research they fund can be found by other researchers who might want to use or otherwise build on the results.

Conclusions

The scientific community has introduced practices to increase rigor and transparency in research. However, stakeholders noted that some practitioners have not used many of these tools due to a range of factors, such as misconceptions about these tools, misaligned incentives, and a desire to avoid administrative burdens. Yet, as experts from our roundtable discussions maintain, federal science funding agencies have significant influence on which research practices are widely adopted. Additionally, they stated that these federal agencies have the potential to shift norms within the research community. Whether this calls for instituting different requirements for rigor and transparency or reprioritizing some funding, such as for training, remains to be seen. As the National Academies have noted, the full scope of non-reproducibility and non-replicability of research is generally unknown and likely varies by scientific discipline.

Nevertheless, national policy directives issued by OSTP and the Office of Management and Budget have called for more assurances from federal funders for rigor and transparency. As experts and others have noted, there are a variety of actions that agencies can take to help ensure methodological rigor and improve access to taxpayer-funded research data and results. Such efforts could help instill a culture that aligns researchers' incentives with sound research practices.

However, without collecting information about relevant indicators of rigor and transparency of the research they are funding, NIH, NASA, and NSF lack the information and insight needed to assess the scope of any problems in their respective fields of study. Without these assessments, these agencies lack information that could inform programmatic changes

to improve research. Taking these steps would better position federal agencies to address public concerns over research reliability, or entertain additional strategies to strengthen the yield of a national enterprise with enormous federal investment.

Recommendations for Executive Action

We are making the following six recommendations to NIH, NSF, and NASA:

- The Director of NIH should collect information on relevant indicators of rigor to assess the research projects the agency funds, and implement steps, as needed, to promote strong research practices in future work. (Recommendation 1)
- The Director of NIH should take steps to collect information to determine whether current policies and requirements are adequate to achieve transparency by ensuring research results and data are findable, accessible, and usable, and implement programmatic or policy changes, if needed. (Recommendation 2)
- The Director of NSF should collect information on relevant indicators of rigor to assess the research projects the agency funds, and implement steps, as needed, to promote strong research practices in future work. (Recommendation 3)
- The Director of NSF should take steps to collect information to determine whether current policies and requirements are adequate to achieve transparency by ensuring research results and data are findable, accessible, and usable, and implement programmatic or policy changes, if needed. (Recommendation 4)
- The Administrator of NASA should collect information on relevant indicators of rigor to assess the research projects the agency funds, and implement steps, as needed, to promote strong research practices in future work. (Recommendation 5)
- The Administrator of NASA should take steps to collect information to determine whether current policies and requirements are adequate to achieve transparency by ensuring research results and data are findable, accessible, and usable, and implement programmatic or policy changes, if needed. (Recommendation 6)

Agency Comments and Our Evaluation

We provided a draft of this report for review and comment to NIH, NSF, and NASA. NIH provided written comments, reproduced in appendix II, in which it concurred with both of the recommendations for action. NSF's Section Head for Integrative Activities, and the Senior Advisor and Scientific Integrity Official, Office of the Director, provided comments

orally, and the NSF Liaison to GAO provided written comments via email. These comments, summarized below, generally conveyed NSF's concurrence with the two recommendations. NASA provided written comments that are summarized below and reproduced in appendix III. In its letter, NASA did not concur with the recommendation to collect information on relevant indicators of rigor and partially concurred with the recommendation to determine whether its policies and requirements are adequate to achieve research transparency.

In concurring with the recommendation to collect information on relevant indicators of rigor, NSF officials stated that the recommendation aligns with its efforts in continuous improvement to support research and maximize the impact of NSF-funded research. NSF officials noted, however, that the agency's research portfolio covers nearly all areas of science and stated that, while some research fields may have clear, well-defined indicators of rigor, the most relevant indicators of rigor for other areas of science are less clearly defined. NSF officials also noted that assessing rigor across a representative sample of research is likely to be a highly resource intensive undertaking. They proposed to begin addressing the recommendation by fostering research to develop robust indicators of rigor across different scientific disciplines, and subsequently to explore ways to encourage researchers to consider their use for demonstrating the rigor of their research. The information gathered from such research would then inform NSF's policies and procedures for improving and promoting sound research practices.

Given the wide range of scientific disciplines that NSF funds, we understand that NSF would first want to work with the scientific community to develop indicators of rigor across a number of research disciplines. Such an approach could be especially beneficial in disciplines where indicators of rigor may not be clear and well established. As acknowledged in this report, different disciplines and types of research present distinct and complex challenges to methodological rigor. We agree that NSF's proposed actions are consistent with the intent of the recommendation and acknowledge that such efforts may call for additional resources. Nevertheless, NSF could assess the rigor of those research projects for which clear and consistent indicators of rigor are already established. Additionally, this report outlines several options—such as an annual audit of a sample of studies—that could help NSF alleviate its resource concerns.

In not concurring with the recommendation to collect information on relevant indicators of rigor in projects funded by the agency, NASA stated

that it uses its peer review process to ensure research rigor, quality, transparency, and relevance, noting it believes the peer review process is the best way to ensure research reliability. NASA also stated the agency is aware of little to no evidence of weak research practices in the research it has funded. Finally, NASA stated that it lacks the resources necessary to collect information on indicators of rigor.

While we recognize that NASA has confidence in its peer review process, this report documented the concerns of the scientific community over the reliability of research that has been through the peer review process. Further, as discussed in this report, studies have documented this lack of reliability across a wide range of scientific disciplines, including disciplines funded by NASA. NASA states that the agency sees little to no evidence of weak research practices; however, without assessing research, the agency does not have assurance that this is the case. Collecting information on relevant indicators of rigor would give NASA assurance that the research it is funding is carried out with sufficient rigor, help the agency identify any deficiencies, and inform actions to address such deficiencies. Finally, while we acknowledge that collecting information on indicators of rigor will require redirection of some agency resources, doing so is important to assure that agency research funds are resulting in rigorous research, as discussed in this report. We maintain that the recommendation is valid.

In its partial concurrence with the recommendation for research transparency, NASA stated that the agency is already engaged in work that will accomplish the objective of this recommendation. We agree that the work NASA is engaged in is consistent with the recommendation and likely meets its intent. It remains to be seen whether, and to what extent, additional resources may be necessary to collect more detailed information after that work is complete. Such a determination could be made when the agency's current work, as described, is complete.

We are sending copies of this report to the appropriate congressional committees, the acting Director of NIH, the Director of NSF, the Administrator of NASA, and other interested parties. In addition, the report is available at no charge on the GAO website at http://www.gao.gov.

If you or your staff have any questions about this report, please contact me at (202) 512-6888 or wrightc@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made key contributions to this report are listed in appendix III.

Sincerely yours,

Candice N. Wright
Director, Science, Technology Assessment, and Analytics

Appendix I: Objectives, Scope, and Methodology

In this report, we (1) outline strategies that, according to stakeholders in the scientific community, are available to promote research rigor and transparency, and factors they say currently discourage wider adoption in the research community; (2) examine what actions, according to experts, federal agencies could take to foster rigor and transparency in the research they fund; and (3) assess the extent to which selected federal science funding agencies have taken actions to improve rigor and transparency.

Review of Strategies to Promote Research Rigor and Transparency

To obtain information on available strategies to promote research rigor and transparency, and factors that discourage wider adoption by the research community, we reviewed relevant literature that addressed research reliability, reproducibility, replicability, and related topics. We also interviewed a wide range of stakeholders selected from across the research community who could speak to these issues.

We used three approaches to identify relevant literature that discussed these existing strategies and factors that hinder their implementation. We performed this work from September 2020 through February 2022. These three approaches, combined, produced more than 200 articles—including empirical studies, commentaries, and trade articles—for our review. For the first approach, we conducted searches using freely accessible online search engines and obtained references from initial background interviews with stakeholders. The purpose of the first approach was to establish an informed understanding of the general landscape and scholarship surrounding research rigor and transparency within the scientific community. This approach included the review of scholarly articles, working papers, commentaries, and news articles. For the second approach, we obtained and reviewed literature referenced in the 2019 National Academies of Sciences, Engineering, and Medicine (National Academies) report, *Reproducibility and Replicability in Science*.[1]

The third approach involved conducting further, formal literature searches. A GAO librarian conducted two formal literature searches that spanned a range of literature published from January 2005 through December 2020—including books, working papers, empirical studies, and review articles. To identify relevant sources, we conducted searches in various bibliographic databases, such as ProQuest, Scopus, and

[1]National Academies of Sciences, Engineering, and Medicine, *Reproducibility and Replicability in Science* (Washington, D.C.: The National Academies Press, 2019). https://doi.org/10.17226/25303.

WorldCat. The search terms used to locate relevant citations included "science," "research," "reproducibility," "replicability," "reliability," "culture," and "change management."[2] We (1) conducted an initial search and selected the most relevant articles on the topic of research reproducibility and replicability; (2) reviewed abstracts from the selected articles to select the most relevant full articles for in-depth review; and (3) examined metadata from the full articles, such as indexing and citation references. The last step helped to identify whether, for example, the indexing and citation references were repeated enough to reach saturation of the topic to the extent possible.[3]

To identify relevant nonfederal stakeholders who represented a diversity of viewpoints across the research community, we selected a non-generalizable sample of stakeholders from seven broad categories. We identified these categories based on information obtained from the 2019 National Academies report, preliminary interviews with stakeholders, and review of relevant literature. We identified seven key nonfederal stakeholder groups: (1) researchers; (2) private funders; (3) journal publishers; (4) academic, corporate, and independent research institutions; (5) professional societies; (6) academic and independent research libraries; and (7) journalists and public affairs specialists. We interviewed 28 stakeholders, chosen based on their level of professional involvement on these topics, which we judged based on individuals' backgrounds, experience, and relevant publications. We also sought to obtain the views of a wide-range of stakeholders that represented various academic disciplines. The stakeholders we interviewed were not a representative sample of all stakeholders with expertise on research reliability, but they each demonstrated extensive involvement on this topic and, together, offered a range of perspectives.

[2]The 2019 National Academies report covered the broad topic of reproducibility and replicability in scientific research as it relates to research reliability, but the report also drew conclusions about ways to improve rigor and transparency. Our literature review search terms focused on reproducibility and replicability and related topics, though it did not explicitly include rigor and transparency.

[3]There was overlap in the results of the informal and more systemic approaches to the literature review.

Examining Actions That Federal Funding Agencies Might Undertake to Foster Rigorous and Transparent Research

To examine what actions federal agencies might undertake to foster rigor and transparency in the research they fund, we held a series of expert interviews as well as four roundtable discussion groups with leading experts on specific topics related to the broad theme of research reliability. We developed discussion topics based on initial reviews of relevant literature and initial interviews with selected stakeholders. We modified and updated our topic list based on additional stakeholder interviews in which we queried interview subjects about the extent to which the topics covered major themes of interest related to research reliability. The discussion topics were as follows:

- Preregistration, the public posting of study design, methods, protocols, and intended analyses prior to conducting the study.
- Registered reports, a convention in publishing where journal publishers provisionally agree to publish the results of a study based on the described methodology, as long as the study is carried out within those parameters.
- Preprints, completed scholarly manuscripts shared on a public server prior to pre-publication peer review.
- Data repositories, the digital infrastructure for the curation and long-term storage of study data, including protocols, equipment identifiers, experimental data, and other study information.
- Researcher training, in areas such as the appropriate use of statistical analyses, robust study design, and open science practices.
- Realigning publishing incentives, to reward robust science conducted with rigor and transparency more than the quantity of published papers.
- Funding strategies agencies can use to promote rigor and transparency.
- Promoting a culture of reliable research, which includes promoting shared attitudes, values, goals, and practices that prioritize robust science conducted with integrity, rigor, and transparency as the norm.

We conducted discussion groups for four of the eight categories: training, publishing incentives, research culture, and funding strategies. We chose these four topics because they are broad in scope and, therefore, conducive to brainstorming discussions as a means to elicit relevant information from the experts. We solicited information from experts on the more concrete topics (preregistration, registered reports, preprints, and data repositories) by conducting interviews with individual experts. For each of these topics, we asked experts about actions that funding

agencies could take, the advantages and disadvantages of those actions, and challenges that the agencies might face in taking actions.

To identify leading experts for our interviews and group discussions, we used a technique referred to as snowball sampling. With this technique, we first identified a core group of five experts through a review of relevant literature and initial interviews with stakeholders. We asked the initial five experts to recommend up to two experts, without regard to any specific scientific disciplines, in each of the eight discussion topics.[4] To do this, we developed a questionnaire that defined each discussion topic and asked each expert to recommend other individuals "who would have the greatest expertise to discuss actions that federal funding agencies could take to support" researchers in each of the eight topics.[5] We also asked each expert to self-rate their own expertise in each of the eight topics.

From these first five questionnaire responses, we compiled a list of recommended experts and asked each of those experts to recommend up to two other experts for each topic. This process continued for several iterations, with each wave of experts being asked to recommend other experts, until a list of more than 700 unique experts was developed. Overall, experts that were most commonly recommended tended to be in the life or social sciences, including those that are actively engaged in the sociology of science—particularly "research on research," which is the study of how science is done across disciplines.[6] We ranked experts in each of the eight topic areas according to the number of recommendations they received in that topic area. For the top 10 to 12 ranked experts in each topic area, we then assessed qualitative information from each, including their biography, relevant publishing history, and experience. We did this to confirm that they had relevant expertise on the topic for which they were recommended. For each

[4]For the purposes of this snowball sampling methodology, we defined an expert as an individual with extensive knowledge about the topic, extensive experience working to address the topic, and balance and objectivity when discussing the topic.

[5]The practical difficulties of developing and administering a survey may introduce errors, such as differences in how respondents interpret particular questions or in the sources of information available to respondents when answering a question. Therefore, we included steps to minimize such errors. In particular, we conducted pretests of the questionnaire with seven experts to ensure that the questions were clear and unambiguous, terminology was used correctly, the information could feasibly be obtained, and the survey was comprehensive and unbiased. We made changes as appropriate in response to feedback we received during these pretests.

[6]A limitation of snowball sampling is that full representation is not guaranteed, as subjects tend to nominate people they already know.

discussion topic, we extended invitations to the most-recommended experts until we received affirmative responses from at least five experts for each topic. We also sought to obtain a balance of domestic and international experts in each group. A total of 22 experts comprised our final groups for the four discussion sessions. See table 2 for a list of experts who participated in these discussion groups.

Table 2: Participants in GAO's Expert Discussion Groups

Session 1: Funding strategies
Dorothy Bishop, Professor of Developmental Neuropsychology, University of Oxford
Philip Bourne, Professor of Biomedical Engineering, University of Virginia
Stuart Buck, Vice President of Research, Arnold Ventures
Marcus Munafo, Professor of Biological Psychology, University of Bristol
Carly Strasser, Program Manager for Open Science, Chan Zuckerberg Initiative
Greg Tananbaum, Open Research Funders Group
Session 2: Realigning publishing incentives
Chris Chambers, Professor of Cognitive Neuroscience, Cardiff University
Anna Hatch, Program Director, Declaration on Research Assessment
Veronique Kiermer, Chief Scientific Officer, Public Library of Science
Frank Miedema, Vice Rector for Research, University Medical Center, Utrecht
David Moher, Senior Scientist, Ottawa Hospital Research Institute
Brian Nosek, Executive Director, Center for Open Science
Session 3: Promoting a culture of reliable research
Ulrich Dirnagl, Director of the Department of Experimental Neurology, Berlin Institute of Health
Maryrose Franko, Executive Director, Health Research Alliance
Marcia McNutt, President, National Academy of Sciences
Brian Nosek, Executive Director, Center for Open Science
Bodo Stern, Chief of Strategic Initiatives, Howard Hughes Medical Institute
Session 4: Supporting researcher training
Richard Ball, Professor of Economics, Haverford College
Steve Goodman, Professor of Epidemiology, Stanford University
Kari L. Jordan, Executive Director, The Carpentries
Marcus Munafo, Professor of Biological Psychology, University of Bristol
Kirstie Whitaker, Program Director, Alan Turing Institute

Source: GAO. | GAO-22-104411

Note: Expert affiliations are listed as of the time of our discussion group sessions.

To collect information to structure each discussion group, we emailed each expert a questionnaire several days in advance. The questionnaires

asked experts to provide their best ideas—up to three different concrete action steps—for actions that federal agencies could take to meet the objective of the discussion session topic. For each action they recommended, experts were also asked to answer follow-up questions. Each question included sub-questions offering additional items for the experts to consider, as follows:

a) *What action would you recommend that U.S. federal funding agencies take to support the scientific community in [discussion topic]? Please consider the following elements.*

- What is the nature of the action? (mandate/requirement, a change in incentives, etc.)
- What root cause does this action address?
- Is this root cause addressed directly or indirectly?

b) *Why do you recommend this particular action? Please consider the following elements.*

- What are the key benefits of this action?
- Will benefits of this action be seen immediately or after period of time?
- What are the potential disadvantages/negative consequences of this action?

c) *How would U.S. federal funding agencies implement this action, and what challenges would they face? Please consider the following elements.*

- Who is the responsible party?
- Would the action involve inclusion of non-federal stakeholders?
- What individual steps are required for implementation?
- What barriers exist that may hinder implementation?
- What are the costs or burdens of implementation?
- Does the action stand alone or would it be more effective in conjunction with other actions?

We administered these questionnaires to experts via email in June and July 2021. After gathering the experts' questionnaire responses, we summarized the action ideas that each expert provided, along with

experts' narrative responses, and emailed the summary to each applicable group of experts in advance of the discussion sessions. We asked the experts to review the summary and be prepared to comment on each of the options.

The four group discussions took place in July 2021 and involved two sessions, each lasting about 2 hours. The discussions were moderated by a GAO research methodologist who sought to ensure broad participation among the experts and sufficient depth on each topic. The first discussion group session, on establishing funding strategies to promote reliable research, helped further streamline the format of the following sessions.

During the group discussions, we proceeded through several steps. We structured the core of the discussion around the options provided by the experts and framed the discussions around the following contexts:

- describe key elements of each proposed action
- describe the rationale for proposing each action
- discuss barriers to implementation and ways to overcome barriers

Specifically, we presented each option on a slide and asked the expert who proposed the option to describe its major features. After each expert presented their proposed option, we facilitated a discussion about the details, advantages, and disadvantages of that option. In addition, during the group discussions, we gathered additional, ancillary information from experts using the videoconference chat function.

Rather than quantify the prevalence of support among experts for any particular view, our qualitative approach sought to capture a full range of views on the discussion topics from a non-generalizable sample of experts. Therefore, to determine the threshold in selecting viewpoints for this report, we assessed the strength of evidence supporting those views and the underlying rationale for those views. We considered, among other factors, the extent to which a particular topic was discussed, the degree to which other experts agreed or disagreed with one another, and whether the experts provided sufficient support for a particular discussion topic. Our analysis of the results of interviews and roundtable discussions, including whether factors cited corroborated with our review of relevant literature, helped characterize the range of factors that met this threshold.

Assessing the Extent to Which Select Federal Funding Agencies Have Taken Actions to Improve Rigor and Transparency

To assess the extent to which select federal science funding agencies have taken actions to improve rigor and transparency, we interviewed officials from the National Institutes of Health (NIH), the National Science Foundation (NSF), and the National Aeronautics and Space Administration (NASA). These agencies are the three largest civilian funders of basic scientific research in the United States, funding research across a wide range of scientific disciplines. We also interviewed officials from the Office of Science and Technology Policy (OSTP) to discuss its ongoing work in addressing issues related to open science, research reliability, and related topics. In addition, we spoke with officials from the National Institute of Standards and Technology—an agency whose mission is to promote U.S. innovation and industrial competitiveness by advancing measurement and standards—to understand what role, if any, the agency plays in creating standards related to data repositories or other open science practices.

In addition, we reviewed policies, procedures, and guidance relevant to research reliability, rigor, transparency, and related topics. In particular, we reviewed policies and guidance from NIH, NSF, and NASA on public access and data management, as well as the agencies' grant application requirements and processes for reviewing grant applications. In addition, we reviewed agencies' policies and procedures for assessing compliance with agency policies and requirements. We also reviewed agency initiatives and requirements relevant to researcher training.

We determined that the quality information and risk assessments components of *Standards for Internal Control in the Federal Government* were significant to this objective.[7] Specifically, we looked at the underlying principles that, to achieve its objectives, management should identify risk and define its risk tolerance; analyze and respond to those identified risks; and use quality information by, for example, collecting relevant data from reliable internal and external sources.[8]

In addition, we examined applicable statutes and regulations concerning the federal government's efforts to invest in and strengthen U.S. innovation by improving, among other things, rigor and transparency of federally funded research. Among the statutes and regulations we

[7]GAO, *Standards for Internal Control in the Federal Government*, GAO-14-704G (Washington, D.C.: September 2014).

[8]GAO-14-704G. Principles 6.06, 7.01, 13.01.

examined were The American Innovation and Competitiveness Act and federal guidance issued in 2013 by OSTP on agencies' development of public access and data management plans for federally funded research.[9]

We conducted this performance audit from September 2020 to July 2022 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

[9]American Innovation and Competitiveness Act, Pub. L. No. 114-329, § 116, 130 Stat. 2969, 2994 (2017) and Office of Science and Technology Policy, *Increasing Access to the Results of Federally Funded Scientific Research,* Memorandum (Washington, D.C.: Feb. 22, 2013).

Appendix II: Comments from the Department of Health & Human Services

DEPARTMENT OF HEALTH & HUMAN SERVICES

OFFICE OF THE SECRETARY

Assistant Secretary for Legislation
Washington, DC 20201

July 5, 2022

Candice Wright
Director, Science, Technology Assessment, and Analytics
U.S. Government Accountability Office
441 G Street NW
Washington, DC 20548

Dear Ms. Wright:

Attached are comments on the U.S. Government Accountability Office's (GAO) report entitled, **"Research Reliability: Federal Actions Needed to Promote Stronger Research Practices"** (GAO-22-104411).

The Department appreciates the opportunity to review this report prior to publication.

Sincerely,

Melanie Anne Egorin

Melanie Anne Egorin, PhD
Assistant Secretary for Legislation

Attachment

GENERAL COMMENTS FROM THE DEPARTMENT OF HEALTH & HUMAN SERVICES ON THE GOVERNMENT ACCOUNTABILITY OFFICE'S DRAFT REPORT ENTITLED — Research Reliability: Federal Actions Needed to Promote Stronger Research Practices (GAO-22-104411)

The U.S. Department of Health & Human Services (HHS) appreciates the opportunity from the Government Accountability Office (GAO) to review and comment on this draft report.

General Comments

Recommendation 1

The Director of NIH should collect information on relevant indicators of rigor to assess the research projects the agency funds, and implement steps, as needed, to promote strong research practices in future work.

HHS Response

HHS concurs with GAO's recommendation.

NIH plans to issue a Notice in the NIH Guide to Grants and Contracts encouraging applicants to include elements of rigor in their grant applications and resulting publications by the end of Fiscal Year (FY) 22. Examples of these elements of rigor include study design, sample size, inclusion and exclusion criteria, randomization, blinding, outcome measures, and statistical methods. While NIH already expect applicants to include these elements in their applications as part of a rigorous study design, we believe that reminding the community to include these elements will increase the consistency of reporting. Periodically, NIH will conduct analyses of grant applications to assess whether applicants are including these elements of rigor.

Examples of Actions Taken to Date

- As noted in the GAO report, in 2015, NIH clarified and revised application instructions and review criteria to enhance reproducibility of research findings through increased scientific rigor and transparency (see NIH Notices, blog posts, and references). These updates took effect for research grants and mentored career development award applications submitted for the January 25, 2016 due date and beyond. With these updates, NIH research grant and career development award application instructions and review language focus on four key areas:
  - The rigor of the prior research
  - Rigorous experimental design for robust and unbiased results
  - Consideration of relevant biological variables
  - Authentication of key biological and/or chemical resources.
- When asking applicants to describe their strategies for ensuring rigorous experimental design for robust and unbiased results, NIH explains that scientific rigor is the strict application of the scientific

method to ensure robust and unbiased experimental design, methodology, analysis, interpretation and reporting of results. NIH expects full transparency in proposing and reporting experimental details so that reviewers may assess the proposed research and others may reproduce and extend the findings.

- NIH also provides a number of resources for applicants as they plan to address rigor in their applications. For example, the Experimental Design Assistant is a free online tool from the NC3Rs, designed to guide researchers through the design of their experiments, helping to ensure that they use the minimum number of animals consistent with their scientific objectives, methods to reduce subjective bias, and appropriate statistical analysis.
- NIH also provides a number of training resources on rigorous study design, including modules on (1) blinding and randomization; (2) sample size, outliers, and exclusion criteria; and (3) biological and technical replicates.
- In a recent presentation to the Advisory Committee to the NIH Director, the NIH Deputy Director for Extramural Research, provided an update on NIH's efforts to address rigor and reproducibility in preclinical animal research. The presentation highlighted numerous ongoing NIH initiatives, including a resource offered by the National Institute on Aging called AlzPED, the Alzheimer's Disease Preclinical Efficacy Database. AlzPED is a publicly available, searchable, data resource that aims to increase the transparency, reproducibility and translatability of preclinical efficacy studies of candidate therapeutics for Alzheimer's disease. AlzPED provides quick access and visibility to integrated preclinical efficacy data from published and unpublished studies. This includes information on indicators of rigor, such as study design, randomization, outcome measures, statistical plans, blinding, and more.

NIH will provide a status update to GAO in January 2023.

Recommendation 2

The Director of NIH should take steps to collect information to determine whether current policies and requirements are adequate to achieve transparency by ensuring research results and data are findable, accessible, and useable, and implement programmatic or policy changes, if needed.

HHS Response

HHS concurs with GAO's recommendation.

Data sharing has been long expected of NIH funded studies. On October 29, 2020, NIH announced the Final NIH Policy for Data Management and Sharing that will be effective for applications received on or after January 25, 2023 (this replaces NIH's 2003 policy). The policy requires researchers to prospectively plan for how scientific data will be preserved and shared through submission of a Data Management and Sharing Plan. In the Notice that describes the elements that should be included in a Data Management and Sharing Plan, NIH "encourages data management and data sharing practices consistent with the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles and reflective of practices within specific research communities." NIH's guidance for writing a Data Management and Sharing Plan asks the applicant to address data type; related tools, software and/or code; standards; data preservation, access, and associated timelines; and access, distribution, or reuse considerations. The Data Management and Sharing Plan will be incorporated into the terms and conditions of the NIH award. NIH will monitor recipients' compliance with

the plan at regular reporting intervals; mechanisms and tools to support oversight are currently under development. Recipients' compliance with their Data Management and Sharing Plans may factor into future funding decisions. As this new policy is implemented, NIH will be assessing the process and outcomes, which would inform any future policy changes.

NIH will provide a status update to GAO in January 2023.

Appendix III: Comments from the National Aeronautics and Space Administration

National Aeronautics and
Space Administration

Mary W. Jackson NASA Headquarters
Washington, DC 20546-0001

June 27, 2022

Reply to Attn of: Office of the Chief Scientist

Ms. Candice N. Wright
Director
Science, Technology Assessment, and Analytics
United States Government Accountability Office
Washington, DC 20548

Dear Ms. Wright:

The National Aeronautics and Space Administration (NASA) appreciates the opportunity to review and comment on the Government Accountability Office (GAO) draft report entitled, "Research Reliability: Federal Actions Needed to Promote Stronger Research Practices" (22-104411), dated May 27, 2022.

In this draft report, GAO makes two recommendations to NASA intended to enhance NASA's ability to gather and review research results of Federal funded research.

Specifically, GAO recommends the following:

Recommendation 1: The Administrator of NASA should collect information on relevant indicators of rigor to assess the research projects the agency funds, and implement steps, as needed, to promote strong research practices in future work.

Management's Response: Non-concur. NASA is committed to ensuring research reliability in the research projects that it funds. NASA believes that the best way to ensure research reliability is the peer review process, which has long been the gold standard for scientific credibility. Accordingly, NASA relies on the peer review process in the scientific community to assess research rigor, quality, transparency, and relevance of science proposals submitted to NASA, as well as the scientific journal publications arising from NASA-funded research. Furthermore, NASA is aware of little to no evidence of weak research practices in the research the Agency funds. Finally, while NASA values the recommendation's goal of promoting strong research practices in NASA-funded research, NASA lacks the resources that would be necessary to collect information on indicators of rigor.

Estimated Completion Date: N/A.

Recommendation 2: The Administrator of NASA should take steps to collect information to determine whether current policies and requirements are adequate to achieve transparency by ensuring research results and data are findable, accessible, and usable, and implement programmatic or policy changes, if needed.

Management's Response: Partially concur. NASA is already engaged in two efforts that will accomplish the objective of this recommendation. First, NASA's Science Mission Directorate has initiated its Year of Open Science, and the goals of that effort include, but are not limited to, enhancing ready access to NASA-funded research publications by applying the FAIR rubric: Findability, Accessibility, Interoperability, and Reuse. A challenge in this area is the increased costs for open-source publications compared to the traditional model. Second, NASA is in the process of revising its Scientific Integrity policy in response to the output of the Science Integrity Fast Track Action Committee (FTAC) convened in response to the January 27, 2021, "Presidential Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking | The White House." The FTAC report, "Protecting the Integrity of Government Science, A Report by the Scientific Integrity Fast-Track Action Committee of the National Science and Technology Council," is available at 01-22-Protecting_the_Integrity_of_Government_Science.pdf (whitehouse.gov). Recommendations from this report include, but are not limited to, increasing transparency and accessibility of all Government-funded research.

The resources being devoted to accomplishing these two efforts are significant, and additional resources to collect more detailed information are not available.

Estimated Completion Date: These two efforts are underway during calendar year 2022 and it is anticipated that they should be complete in early 2023; both initiatives are likely to lead to further activities in these areas.

We have reviewed the draft report for information that should not be publicly released. As a result of this review, we have not identified any information that should not be publicly released.

Once again, thank you for the opportunity to review and comment on the subject draft report. If you have any questions or require additional information regarding this response, please contact LaVerne Drayton on (202) 358-1909.

Sincerely,

David S. Draper
Deputy Chief Scientist
On behalf of Katherine Calvin, Chief Scientist

Appendix IV: GAO Contact and Staff Acknowledgments

GAO Contact

Candice N. Wright at (202) 512-6888 or WrightC@gao.gov.

Staff Acknowledgments

In addition to the contact named above, Robert J. Marek (Assistant Director), Michael Krafve, Kristina Hammon, Cheryl M. Harris, Patricia Miller, and Arvin Wu made key contributions to this report. Also contributing to this work were Sue Bernstein, Mark Braza, Leia Dickerson, Louise Fickel, Ryan Han, Patrick Harner, and Jack Wang.

Selected Bibliography

Alexander, Patrick H. "Open Access and Author Rights: Questioning Harvard's Open Access Policy." *Insights,* 33, no. 1 (2020). Accessed March 15, 2021. http://doi.org/10.1629/uksg.525.

Ball, Philip. "High-Profile Journals Put to Reproducibility Test." *Nature* (2018). Accessed November 3, 2020. https://doi.org/10.1038/d41586-018-06075-z.

Begley, C. Glenn, and John P. A. Ioannidis. "Reproducibility in Science: Improving the Standard for Basic and Preclinical Research." *Circulation Research,* 116, no.1 (2015). Accessed April 29, 2021. https://doi.org/10.1161/CIRCRESAHA.114.303819.

Bourne, Philip E., Jessica K. Polka, Ronald D. Vale, et al. "Ten Simple Rules to Consider Regarding Preprint Submission." *PLOS Computational Biology,* 13, no. 5 (2017). Accessed December 14, 2020. https://doi.org/10.1371/journal.pcbi.1005473.

Cacioppo, John T., Robert M. Kaplan, Jon A. Krosnick, et al. *"*Report of the Subcommittee on Replicability in Science Advisory Committee to the National Sciences Foundation Directorate for Social, Behavioral, and Economic Sciences." *Social, Behavioral, and Economic Sciences Perspectives on Robust and Reliable Science,* Alexandria, Va.: National Science Board, May 2015. Accessed Oct. 7, 2020. https://www.nsf.gov/sbe/advisory.jsp.

Cará, Piera Demma, Rosaria Ciriminna, and Mario Pagliaro. "Has the Time Come for Preprints in Chemistry?" *ACS Omega*, 2, no. 11 (2017). Accessed March 15, 2021. https://pubs.acs.org/doi/10.1021/acsomega.7b01190.

Carroll, Michael W. "Sharing Research Data and Intellectual Property Law: A Primer." *PLOS Biology*, 13, no 8 (2015). Accessed December 21, 2021. https://doi.org/10.1371/journal.pbio.1002235.

Coudert, François-Xavier "The Rise of Preprints in Chemistry." *Nature Chemistry,* 12, no. 6 (2020). Accessed March 25, 2021. https://doi.org/10.1038/s41557-020-0477-5.

DeHaven, Alexander. “Preregistration: A Plan, Not a Prison.” *Center for Open Science*, (2017). Accessed March 7, 2022. https://www.cos.io/blog/preregistration-plan-not-prison.

Ebersole, Charles R., Jordan R. Axt, and Brian A. Nosek. "Scientists' Reputations Are Based on Getting it Right, Not Being Right." *PLOS Biology,* 14, no. 5 (2016). Accessed October 22, 2020. https://doi.org/10.1371/journal.pbio.1002460.

Errington, Timothy M., Alexandria Denis, Nicole Perfito, et al. "Reproducibility in Cancer Biology: Challenges for Assessing Replicability in Preclinical Cancer Biology." *eLife* (2021). Accessed December 7, 2021. https://doi.org/10.7554/eLife.67995.

Heesen, Remco. "Why the Reward Structure of Science Makes Reproducibility Programs Inevitable." *The Journal of Philosophy,* 115, no. 12 (2018). Accessed December 8, 2021. https://doi.org/10.5840/jphil20181151239.

Ioannidis, John P. A. "Why Most Published Research Findings Are False." *PLOS Medicine,* 2, no. 8 (2005). Accessed December 30, 2020. https://doi.org/10.1371/journal.pmed.0020124.

Jin, Yanling, Nitika Sanger, Ieta Shams, et al. "Does the Medical Literature Remain Inadequately Described Despite Having Reporting Guidelines for 21 Years—A Systemic Review of Reviews: An Update." *Journal of Multidisciplinary Healthcare,* 11 (2018). Accessed February 18, 2022. https://dx.doi.org/10.2147%2FJMDH.S155103.

Kabitzke, Patricia, Kristin M. Cheng, and Bruce Altevogt. "Guidelines and Initiatives for Good Research Practice." *Good Research Practice in Non-Clinical Pharmacology and Biomedicine. Handbook of Experimental Pharmacology*, vol. 257, edited by Anton Bespolov, Martin C. Michel and Thomas Stecklen, 19–34. Cham, Switzerland: Springer, 2019. https://doi.org/10.1007/164_2019_275.

Koroshetz, Walter J., Shannon Behrman, Cynthia J. Brame, et al. "Research Culture: Framework for Advancing Rigorous Research." *eLife,* 9 (2020). Accessed January 27, 2021. https://doi.org/10.7554/eLife.55915.

Kühberger, Anton, Astrid Fritz, and Thomas Scherndl. "Publication Bias in Psychology: A Diagnosis Based on the Correlation between Effect Size and Sample Size." *PLOS ONE,* 9, no. 9 (2014). Accessed December 31, 2020. https://doi.org/10.1371/journal.pone.0105825.

Lin, Dawei, Jonathan Crabtree, Ingrid Dillo, et al. "The TRUST Principles for Digital Repositories." *Scientific Data,* 7, no. 144 (2020). Accessed September 14, 2021. https://doi.org/10.1038/s41597-020-0486-7.

Lowenberg, Daniella. "Where's the Adoption? Shifting the Focus of Data Publishing in 2018." *UC Curation Center* (December 18, 2017). Accessed November 8, 2021. https://medium.com/@UC3CDL/wheres-the-adoption-shifting-the-focus-of-data-publishing-in-2018-8506f80371cd.

Meng, Xiao-Li. "Reproducibility, Replicability, and Reliability." *Harvard Data Science,* 2, no. 4 (2020). Accessed March 31, 2021. https://doi.org/10.1162/99608f92.dbfce7f9.

Miller, Gary W. "Data Sharing in Toxicology: Beyond Show and Tell." *Society of Toxicology,* 143, no. 1 (2015). Accessed October 21, 2020. https://doi.org/10.1093/toxsci/kfu237.

Miller, Gary W. "Making Data Accessible: The Dryad Experience." *Society of Toxicology,* 149, no. 1 (2016). Accessed October 21, 2020. https://doi.org/10.1093/toxsci/kfv238.

National Academies of Sciences, Engineering, and Medicine, *Advancing Open Science Practices: Stakeholder Perspectives on Incentives and Disincentives: Proceedings of a Workshop-in Brief* (Washington, D.C.: The National Academies Press, 2020). Accessed December 30, 2020. https://doi.org/10.17226/25725.

National Academies of Sciences, Engineering, and Medicine, *Open Science by Design: Realizing a Vision for 21st Century Research* (Washington, D.C.: The National Academies Press, 2018). Accessed September 14, 2020. https://doi.org/10.17226/25116.

National Academies of Sciences, Engineering, and Medicine, *Reproducibility and Replicability in Science* (Washington, D.C.: The National Academies Press, 2019). Accessed July 22, 2020. https://doi.org/10.17226/25303.

National Information Standards Organization, *Reproducibility Badging and Definitions: A Recommended Practice of the National Information Standards Organization* (Baltimore, Maryland: National Information Standards Organization, 2021), no. RP-31-2021. Accessed January 28, 2021. https://doi.org/10.3789/niso-rp-31-2021.

Neville, Tina, and Camielle Crampsie. "From Journal Selection to Open Access: Practices among Academic Librarian Scholars." *Portal: Libraries and the Academy*, 19, no. 4 (2019). Accessed March 8, 2021. https://muse.jhu.edu/article/735356/summary.

Nosek, Brian A., and D. Stephen Lindsay. “Preregistration Becoming the Norm in Psychological Science.” *Association for Psychological Science,* (February 28, 2018). Accessed November 3, 2020. https://www.psychologicalscience.org/observer/preregistration-becoming-the-norm-in-psychological-science.

Nosek, Brian A., and Daniël Lakens. "Registered Reports: A Method to Increase the Credibility of Published Results." *Social Psychology*, 45, no. 3 (2014). Accessed October 22, 2020. https://doi.org/10.1027/1864-9335/a000192.

Nosek, Brian A., Charles R. Ebersole, Alexander C. DeHaven, et al. "The Preregistration Revolution." *Proceedings of the National Academy of Sciences*, 115, no. 11 (2018). Accessed October 2020, 2020. https://doi.org/10.1073/pnas.1708274114.

Nosek, Brian A., George Alter, George C. Banks, et al. "Promoting an Open Research Culture." *Science,* 348, no. 6242 (2015). Accessed October 21, 2020. https://www.science.org/doi/10.1126/science.aab2374.

Nosek, Brian A., and Timothy M. Errington. "What is Replication?" *PLOS Biology,* 18, no. 3 (2020). Accessed October 22, 2020. https://doi.org/10.1371/journal.pbio.3000691.

O'Grady, Cathleen. "Quality Shines when Scientists Use Publishing Tactic Known as Registered Reports, Study Finds." *Science Insider* (June 21, 2021). Accessed February 9, 2022. https://www.science.org/content/article/quality-shines-when-scientists-use-publishing-tactic-known-registered-reports-study.

Open Science Collaboration. "Maximizing the Reproducibility of Your Research." In *Psychological Science Under Scrutiny: Recent Challenges and Proposed Solutions*, edited by Scott O. Lillienfeld and Irwin D. Waldman, 1–21. West Sussex, UK: John Wiley and Sons, 2017. https://doi.org/10.31234/osf.io/k9mn3.

Pankeev, I. A. "New Issues in the Field of Legal Regulation of an Author's Right to Scientific Works." *Scientific and Technical Information*

Processing, 42, no. 3 (2015). Accessed March 15, 2021. https://doi.org/10.3103/S0147688215030107.

Parashar, Manish. "Leveraging the National Academies' Reproducibility and Replication in Science Report to Advance Reproducibility in Publishing." *Harvard Data Science Review,* 2, no. 4 (2020). Accessed March 23, 2021. https://doi.org/10.1162/99608f92.b69d3134.

Robson, Samuel G., Myriam A. Baum, Jennifer L. Beaudry, et al. "Promoting Open Science: A Holistic Approach to Changing Behaviour." *Collabra: Psychology,* 7, no. 1 (2021). Accessed December 28, 2021. https://doi.org/10.1525/collabra.30137.

van Schalkwyk, May C. I., Thomas R. Hird, Nason Maani, et al. "The Perils of Preprints." *BMJ,* 370 (2020). Accessed December 21, 2020. https://doi.org/10.1136/bmj.m3111.

Wicherts, Jelte M. "Peer Review Quality and Transparency of the Peer-Review Process in Open Access and Subscription Journals." *PLOS ONE,* 11, no. 1 (2016). Accessed February 8, 2022. https://doi.org/10.1371/journal.pone.0147913.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, et al. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data,* 3, no. 1 (2016). Accessed May 19, 2021. https://doi.org/10.1038/sdata.2016.18.

Wolfram, Dietmar, Peiling Wang, and Adam Hembree. "Open Peer Review: Promoting Transparency in Open Science." *Scientometrics,* 125, no. 2 (2020). Accessed March 31, 2021. https://doi.org/10.1007/s11192-020-03488-4.

Wright, Jessica. "Finally, Biologists Get Serious about Preprints." *Spectrum* (June 17, 2016). Accessed February 3, 2022. https://www.spectrumnews.org/news/finally-biologists-get-serious-about-preprints/.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through our website. Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. You can also subscribe to GAO's email updates to receive notification of newly posted products.

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, https://www.gao.gov/ordering.htm.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on Facebook, Flickr, Twitter, and YouTube.
Subscribe to our RSS Feeds or Email Updates. Listen to our Podcasts.
Visit GAO on the web at https://www.gao.gov.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact FraudNet:

Website: https://www.gao.gov/about/what-gao-does/fraudnet

Automated answering system: (800) 424-5454 or (202) 512-7700

Congressional Relations

A. Nicole Clowers, Managing Director, ClowersA@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548

Strategic Planning and External Liaison

Stephen J. Sanford, Managing Director, spel@gao.gov, (202) 512-4707
U.S. Government Accountability Office, 441 G Street NW, Room 7814,
Washington, DC 20548