10 Things to Know About
Reproducibility and Replicability

One of the pathways by which the scientific community confirms the validity of a new scientific discovery is by repeating the research that produced it. When a scientific effort fails to independently confirm the computations or results of a previous study, some argue that such an observed inconsistency can be an important precursor to new discovery while others fear it may be a symptom of a lack of rigor in science. When a newly reported scientific study has far-reaching implications for science or a major, potential impact on the public, the question of its reliability takes on heightened importance.

1 **The terms reproducibility and replicability take on a range of meanings in contemporary usage.** The report distinguishes and defines the terms as follows: **Reproducibility** means obtaining consistent results using the same input data, computational steps, methods, and conditions of analysis; it is synonymous with computational reproducibility. **Replicability** means obtaining consistent results across studies aimed at answering the same scientific question, each of which has obtained its own data.

2 **Reproducibility and replicability matter.** Reproducibility and replicability are often cited as hallmarks of good science. Being able to reproduce the computational results of another researcher starting with the same data and replicating a previous study to test its results facilitate the self-correcting nature of science.

3 **Computational reproducibility is more prominent now than ever because of the growth in reliance on computing across all of science.** When a researcher reports a study and makes the underlying data and code available, those results should be computationally reproducible by another researcher.

4 **A successful replication does not guarantee that the original scientific results of a study were correct, nor does a single failed replication conclusively refute the original claims.** Unlike the typical expectation of reproducibility between two computations, expectations about replicability are more nuanced.

5 **Occasionally, non-replicability may be caused by helpful sources that advance scientific knowledge, such as discovering previously unknown effects or sources of variability. At other times, a study cannot be replicated due to unhelpful sources, ranging from simple mistakes to methodological errors to bias and fraud.**

6 **Not all studies can be replicated.** While scientists are able to test for replicability of most studies, it is impossible to do so for studies of ephemeral phenomena.

7 **One type of scientific research tool, statistical inference, has an outsized role in replicability discussions due to the frequent misuse of statistics and the use of a *p*-value threshold for determining "statistical significance."** Biases in published research can occur due to the excess reliance on and misunderstanding of statistical significance.

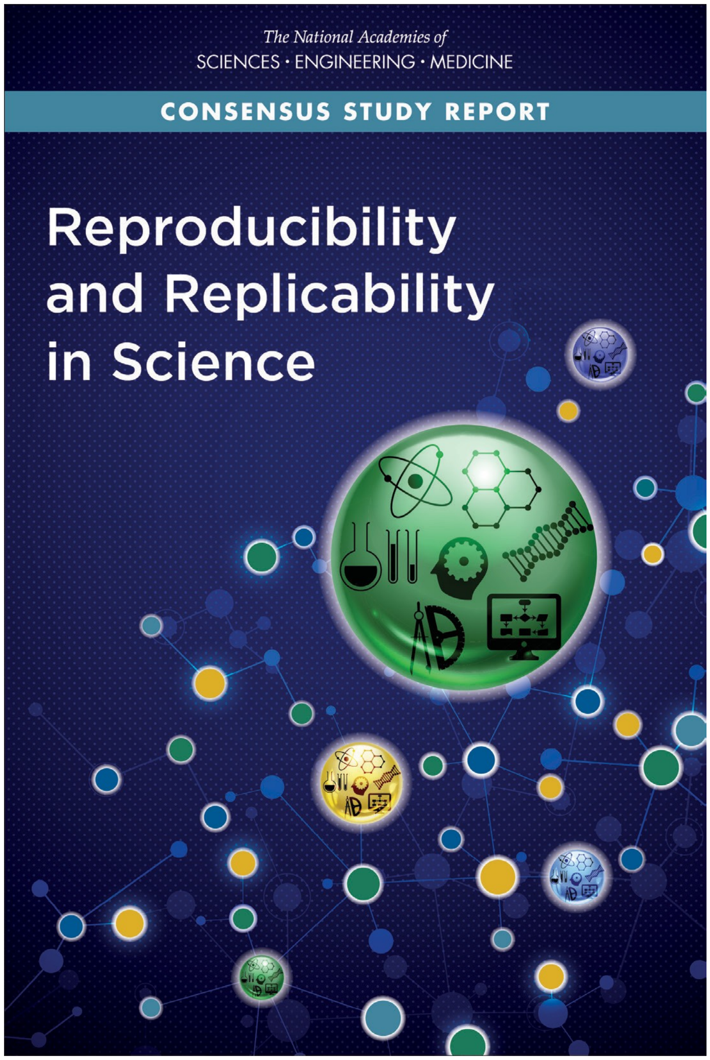
8 **Examining replicability becomes especially important when new findings have strong implications for individual health and well-being, policy choices, or the future course of scientific research.**

9 **Beyond reproducibility and replicability, systematic reviews and syntheses of scientific evidence are among the important ways to gain confidence in scientific results.**

10 **Academic institutions, journals, conference organizers, funding organizations, and policy makers can all play a role in improving the reproducibility and replicability of research. Responsibility begins with researchers, who should take care to estimate and explain the uncertainty inherent in their results and inferences, make proper use of statistical methods, and describe their methods and data in a clear, accurate, and complete way.**

Reproducibility and Replicability in Science
is available at www.nap.edu/25303.

The National Academies of SCIENCES ENGINEERING MEDICINE



http://nationalacademies.org/reproducibilityinscience

Committee Members

HARVEY V. FINEBERG, (*Chair*), Gordon and Betty Moore Foundation

DAVID B. ALLISON, School of Public Health-Bloomington, Indiana University

LORENA A. BARBA, School of Engineering and Applied Science, George Washington University

DIANNE CHONG, Boeing Research and Technology (Retired)

JULIANA FREIRE, Tandon School of Engineering, New York University

GERALD GABRIELSE, Department of Physics, Northwestern University

CONSTANTINE GATSONIS, Center for Statistical Sciences, Brown University

EDWARD HALL, Department of Philosophy, Harvard University

THOMAS H. JORDAN, Department of Earth Sciences, University of Southern California

DIETRAM A. SCHEUFELE, Madison and Morgridge Institute for Research, University of Wisconsin-Madison

VICTORIA STODDEN, Institute for Data Sciences and Engineering, University of Illinois at Urbana-Champaign

TIMOTHY D. WILSON, Department of Psychology, University of Virginia

WENDY WOOD, Department of Psychology, University of Southern California and INSEAD-Sorbonne University

The study is sponsored by the National Science Foundation and The Alfred P. Sloan Foundation.