

Testimony of

Brian A. Nosek, Ph.D.

Executive Director
Center for Open Science

Professor
Department of Psychology
University of Virginia

Before the

Committee on Science, Space, and Technology
U.S. House of Representatives

November 13, 2019

“Strengthening Transparency or Silencing Science? The Future of Science in EPA Rulemaking”

Chairwoman Johnson, Ranking Member Lucas, and Members of the Committee, on behalf of myself and the Center for Open Science, thank you for the opportunity to discuss the role of promoting transparency and reproducibility for maximizing the return on research investments, and responsible management of research transparency with competing interests of privacy protections for sensitive data.

The bottom line summary of my remarks is:

1. **Making open the default** for research plans, data, materials, code, and outcomes will reduce friction in discovery and maximize return on research investments
2. **Extending existing policy frameworks about transparency and openness** across federal agencies will help improve research efficiency. These frameworks can help decision-makers navigate situations in which principles of security and privacy are in conflict with principles of transparency and openness.
3. **Rulemaking should be informed by the best available evidence.** Sometimes the best available evidence is based on data that cannot be transparent, has high uncertainty, or has unknown reproducibility. Developing tools that clarify uncertainty will improve policymaking and shape research priorities.

I joined the faculty at the University of Virginia in the Department of Psychology in 2002. My substantive areas of expertise are research methodology, implicit bias, and the gap between values and practices. In 2013, Jeff Spies and I launched the Center for Open Science (COS) out of my lab as a non-profit technology and culture change organization. COS has a mission to increase transparency, integrity, and reproducibility of research. To advance that mission, COS

maintains the free and open source Open Science Framework (<http://osf.io/>), a cloud-based collaborative management service used by more than 180,000 researchers to improve the rigor and transparency of their research plans, data, materials, code, and outcomes. COS is working to change the incentives landscape in academic science to prioritize accuracy and rigor. COS works with publishers, funders, institutions, and scientific societies to nudge incentives with a policy framework that promotes transparency and reproducibility called the TOP Guidelines (<http://cos.io/top/>), and initiatives that promote visibility of open practices (<http://cos.io/badges/>) and shift publication criteria toward rewarding asking important questions and using rigorous methodologies to investigate them (<http://cos.io/rr/>). Finally, COS conducts metascience--research on the research process--to identify inefficiencies in the process of discovery and to evaluate whether interventions to reduce those inefficiencies are effective.

Lack of transparency creates friction in the pace of discovery and reduces the return on investment of research dollars. For example, in a large-scale replication project of cancer biology research, we initiated replications of 197 experiments and found that the original papers had enough information to design a complete replication protocol for none of them. Moreover, we were able to access the raw data for just 3 of the 197 experiments in public repositories without engaging the original authors. Return on research investments could increase dramatically by promoting greater transparency of a variety of research outputs:

- Transparency and Openness of Materials -- the protocols, materials, and code that generated my research findings -- will make it easier for others to replicate my findings, and build on my research.
- Transparency and Openness of Data will make it easier for others to test the robustness of my findings and to reuse my data for new questions or combine it with related data for more precise assessments of the totality of evidence.
- Transparency and Openness of Research Plans -- registration of the study design, hypotheses, and analysis plans before the results are known -- will make it easier to discover findings that are never published, particularly negative results that are often ignored, and make clear the difference between confirmatory investigations in which hypotheses are being tested and exploratory investigations in which hypotheses are being generated. Mistaking exploratory analyses as confirmatory tests increases bias and is a threat to the credibility of research claims.
- Transparency and Openness of Research Outcomes will make it easier to find all relevant evidence about a research question, and make it easier for researchers, policymakers, and the tax paying public to examine and use the scientific evidence that we all paid to produce.

There is a mature infrastructure of tools and services, like the Open Science Framework and many other repositories, that make it possible for researchers to do these behaviors. There is also growing awareness within the research community about the importance and value for these transparency promoting behaviors. For example, the TOP Guidelines policy framework has been adopted by more than 1,000 scientific journals for authors, and some funders are likewise adapting their policies for grantees. Following the Holdren memo during the Obama

administration, and with continuing interest in promoting rigor and transparency from OSTP in the present administration, many federal agencies have taken steps toward improving policies supporting transparency and reproducibility of research. There is more work to do, but your continuing support for those efforts could have salutary effects on the research culture. Ultimately, COS believes that the biggest opportunity for reducing friction in research progress is **setting the default to open** -- open plans, open materials, open data, and open outcomes.

Flipping the default from closed to open will foster regulatory framework for the exceptions--when other interests outweigh the goal of transparency. Two common occasions in which competing principles can outweigh the principles of openness and transparency are protecting intellectual property and protecting participant confidentiality for sensitive human subjects research. Sensible policies for managing these competing interests will facilitate the culture shift that is already underway in the private sector and with proactive steps by federal agencies such as NIH and NSF.

Also, federal investment in the services and repositories that support research transparency will ensure persistence and accessibility of that content for researchers, policymakers, and the public. Publicly funded research is a public good, and the infrastructure storing and preserving it should be a public good as well.

Finally, there are a variety of technological and methodological innovations that could address goals of transparency and security simultaneously. For example, data enclaves can provide secure storage of sensitive data and workflows for ethical management of reanalysis and reuse without sacrificing that security. Also, there are emerging methodologies that improve privacy by perturbing the characteristics of the underlying data just enough to make it effectively impossible to identify individual data points but still preserve the overall structure of the data for accurate analysis and inference. Supporting such technologies will make it easier to address the otherwise competing principles of transparency and security.

There are important considerations for how best to use scientific evidence in policy making. The EPA rule that prompted this hearing had the positive qualities of identifying the importance of transparency and reproducibility of research, but had the negative quality of suggesting that evidence failing to meet these principles should not be used in policymaking. This approach would degrade the quality of policymaking.

In policymaking, it is important to use the **best available evidence** for rulemaking. There will always be occasions in which the best available evidence is not fully transparent or has unknown reproducibility. Using the best available evidence does not mean using it blindly or overconfidently. There are many factors that affect the quality of research, the certainty of its conclusions, and its generalizability to the policy context. Explicitly representing the uncertainty of evidence will help policymakers make better decisions. When the evidence is more uncertain, policymakers could ensure that implementation of the policy includes mechanisms to evaluate its success. And, by knowing the uncertainty of evidence, policymakers could direct

resources to supporting research to address those certainty gaps and improve the overall evidence base. For example, DARPA's SCORE program is investigating whether machine algorithms could automatically assess the credibility of research claims. If successful, this could provide an initial filter to inform the translation of research evidence into practice, and prioritization of research funding to topics of national and research interest. Your continuing support for programs like DARPA's is a worthwhile for the long-term objective of having evidence-based social and economic policymaking. [Disclosure: COS is funded by DARPA as part of the SCORE program.]

Public investment in science leads to solutions, cures, and unexpected advancements that benefit the national interest. Making open the default for research process, data, materials, and outcomes would transform science, dramatically increase the return on investment from publicly funded research, and accelerate progress. Thank you for your continuing support of science and for the opportunity to speak with you today.