Testimony of

Jeffrey R. Spies, Ph.D.

Chief Technology Officer
Center for Open Science

Visiting Assistant Professor
Department of Engineering and Society
School of Engineering and Applied Science
University of Virginia

Before the

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

March 21, 2017

"National Science Foundation Part II:
Future Opportunities and Challenges for Science"

Chairwoman Comstock, Ranking Member Lipinski, and Members of the Subcommittee, on behalf of myself and the Center for Open Science, thank you for the opportunity to discuss the role of the National Science Foundation in promoting openness and reproducibility in science.

The impact of science and the results of National Science Foundation (NSF) funding are readily observed in our everyday lives: in the food that nourishes us, the transportation that moves us, the buildings that shelter us, the technology that connects us, and the medicine that heals and saves us. The return on investment of NSF-funded science is immense no matter how the impact of that investment is measured, be it dollars, jobs, or lives. This makes sense. NSF has a diverse portfolio, and science is a very safe bet, for it is a process that is ever improving, self-corrects, and results in increased understanding no matter the substance outcome.

I am the co-founder of the Center for Open Science, a nonprofit technology company based in Charlottesville, Virginia. Our mission is to increase openness, integrity, and reproducibility of scholarly research. As the Chief Technology Officer, I direct the technical strategy of our free, open-source products, like the Open Science Framework--a platform for managing workflow as well

as collaborating on and sharing research.[1] I also co-lead SHARE, a partnership with the Association of Research Libraries to build a free, open data set about scholarly research across the research life cycle.[2] I very much have an interest in openness, but it is not as some grand ideal--openness is a practical means of increasing research efficiency, quality, accessibility, and diversity. Openness amplifies the features intrinsic to science--including reproducibility--that make the scientific process such an efficient way of learning about the world.

NSF has had a tremendous record of success by trusting sound scientific process. It is certainly not my intent to claim that in some way science is broken and no longer trustworthy. In fact, to discredit or ignore any body of evidence that comes from such a process would disrespect the same process that has resulted in society-altering advances. Although I will describe challenges that science currently faces, my recommendations to NSF on incentivizing openness and reproducibility through its grant programs is only in service of making an already efficient process work better.

The Challenge of Reproducibility

Scientific results gain credibility by demonstrating that evidence can be independently reproduced (a word I will use interchangeably with replication although their meaning can differ). This means that someone else can take the same data and observe the same outcomes; that someone else can repeat the essential part of the methodology, collect new data, and obtain similar evidence for the claim; and, that someone else can test the same idea with a different methodology and find similar evidence. For example, the same earthquake can never be repeated, but a scientific claim (hypothesis) can be identified for what should occur when another earthquake that shares the essential features occurs again in the future.

For the last few years, science has been characterized as being in a "reproducibility crisis", partly as a consequence of evidence gathered by the Center for Open Science. Collaborating using the Open Science Framework, 270 co-authors attempted to replicate 100 studies from three prominent psychology journals.[3] We found a rate of replication between 39 and 47% depending on the measure. Since then, more than 10 reanalyses of our data have been reported with varying conclusions. These rates are less than one would hope, but "crisis" was not a term we used. Moreover, it is something of a misnomer because this movement is actually just an illustration of the scientific method in practice. This round of self-skepticism is offering new solutions to improve our processes to thus increase the overall efficiency and quality of science.

[1] See http://osf.io/.
[2] See http://share-research.org/.
[3] Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science,* 349(6251). doi: 10.1126/science.aac4716. Everything needed to reproduce the study is available at https://osf.io/ezcuj/.

A requirement for reproducibility is that the process used to obtain a result is described in sufficient detail. And for correction to occur, failed replications must be added to the body of evidence for a given result. Neither is happening frequently enough, and that is making science less efficient than it could be. To be clear, this is not the same as the fraud or misconduct that has been brought to light by failed replications. While those cases certainly make for buzzworthy headlines and exciting stories, they are relatively rare occurrences. And NSF, for example, has in place mechanisms for investigating and appropriately dealing with offenders. But, again, fraud and misconduct are rare. Science does not have a dishonesty problem; it has a communication problem. And it impacts every domain of science.

Much of the problem stems from the simple fact that science is complex. Brief textual descriptions in scientific papers often cannot provide sufficient detail to capture the nuance necessary to facilitate replication. This is the case for openness as the default standard in scientific communication. If as much of the scientific process is open as reasonably possible, then replication can occur more easily, more frequently, and with greater efficacy. If false leads are discovered, they can be discarded, and correction can occur at a faster pace.

Open-by-default is not the current norm. The present scholarly culture is closed by default--we have to justify why something should be open rather than why it should be closed. The current culture does not incentivize a level of description that makes reproducibility efficacious. More so, it does not facilitate or promote the conducting or communicating of failed replications. Much of this is driven by the currency system in science: publication.[4] Scientists are rewarded for publishing as frequently as possible in the most prestigious outlets as possible--this is how they get jobs, promotions and tenure, and funding. Getting published, however, has very little to do with research being open or reproducible; it has to do with novel, positive results and clean, confirmatory narratives. Journals have traditionally avoided publishing studies that did not work as expected or for which the results are messy and ambiguous. But that is exactly how much of science works--to ignore it is to fill the file drawers with unpublished but potentially important findings that no one has access to, reducing the efficiency by which we can make scientific progress.

In a competitive environment, researchers are forced to make choices that increase the likelihood of publication. These are often unwittingly biased decisions that humans have a very difficult time avoiding. The result is increased publications, but decreased accuracy. Transparency maximizes the ability of science to self-correct via critique and replication. And it has the exact same benefit for those rare cases of fraud.

[4] Nosek, B. A., Spies, J. R., & Motyl, M. (2015). Scientific Utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science. 7*(6) 615-631. doi:10.1177/1745691612459058

The Values of Openness

Reproducibility is a core value of science, and its success depends on the ability to understand how a result was obtained. There is often considerable nuance baked into components of the research process, including materials, methods, software, and analyses. Scholarly claims become credible via transparent communication of the supporting evidence and the process of acquiring that evidence. This way, independent observers can evaluate the quality of evidence for supporting the claim. If bias crept into the process or interpretation of results, it would be detectable when the process is open.

While openness can certainly benefit science with respect to amplifying its corrective features, it has another benefit: accessibility and inclusivity. If the goal of scholarly research is the public accumulation of knowledge--if knowledge is a public good--then a default of openness is the first step in removing exclusionary criteria for participation (e.g., monetary cost). If paired with an environment that facilitates and fosters participation through active invitation and education, individuals who would otherwise not be able to contribute to science would be able to participate. Because the individuals most often excluded are minorities in the sciences, science would see efficiency gains not only from the increased number of contributors but from the benefit that diversity brings to collaboration and innovation.

On both of these two dimensions--credibility and accessibility--there are varying degrees of openness that can be incrementally applied to increase the efficiency and quality of science. It is not all or nothing. There are occasions when openness is not possible or when a degree of open is good enough or better than nothing. For example, if data is protected health information and cannot be shared publically in the interest of human participants, there are a number of methods to still increase the credibility of the work. These include opening other components of the research process while excluding the data or making data available to authorized individuals for the purpose of auditing. In the former, you would still see benefits of accessibility, while in the latter, the focus would be on improving the credibility of the work.

Most of the changes that can help science operate more efficiently and maximize knowledge accumulation are related to two simple concepts that everyone learned in grade school: show your work and share.

Show your work. If scientists transparently show how they arrive at their claims, then the marketplace of ideas, critique, and self-correction can operate efficiently. If others cannot see the outcomes, the data supporting the outcomes, and the process by which those outcomes were produced, then it is harder to identify their strengths and limitations.

Share. If scientists share their materials and data openly, others can independently reproduce the findings and reuse the materials to challenge or extend the work. Without sharing, it is much more difficult to accumulate evidence and move toward certainty.

Recommendations

I would make the following recommendations to any funder in the sciences with an interest in increasing research efficiency and quality via openness and reproducibility. They fall into five categories:[5]

- **Metascience**, collecting evidence to encourage change.
- **Infrastructure**, developing technology to enable change.
- **Training**, disseminating knowledge to enact change.
- **Incentives**, promoting reasons to embrace change.
- **Community**, fostering inclusion and connection to propagate change.

NSF has already taken steps to encourage the values of openness and reproducibility, but, as one of the largest funders of the sciences, NSF has a unique platform to continue promoting and thus accelerating the adoption of open and reproducible practices. A cultural shift must take place, and NSF's continued endorsement can quicken the pace of change.

- **Fund efforts to investigate reproducibility through metascience**. Studying reproducibility is not as glamorous as producing new science, but NSF could create strong incentives for this critical work by creating dedicated funding mechanisms to pursue investigations of reproducibility and create a robust science of reproducibility to determine best practices and advance our knowledge of the efficacy of reproducible practices.
- **Fund public goods infrastructure to improve openness and reproducibility across the research lifecycle**. In order to get robust participation in openness and reproducibility efforts, especially while incentives are not aligned with these practices, technology is needed to make that participation as effortless as possible for scientists. At COS we are building open-source platforms for data sharing and access to research description (i.e., metadata). NSF could dramatically expand the infrastructure available to scientists by funding the development and testing of new platforms and collaborations.
- **Fund curation activities and infrastructure to link research workflow, people, and institutions in order to aid in discovery, reuse, analytics, and metascience.** As components of the research process are made openly available (e.g., data, code, software)--likely on disparate platforms--it is important that they are related in a way that discovery of one leads to discovery of the others. This includes the use of persistent

[5] These are the same categories that we use at COS to organize the tasks necessary to meet our mission. See more in our strategic plan at https://osf.io/x2w9h/.

identifiers for content, people, and institutions. This will accelerate reproducibility and reuse and aid metascience efforts.

- **Fund the development of techniques and infrastructure that would allow for the analysis of sensitive data that cannot be made open.** There are a number of research areas and endeavors that would allow for reproducibility and reuse while data remain private. For example, secure (multi-party) computation uses cryptography to conduct analyses on data that is never exposed.
- **Fund development and dissemination of training to improve reproducibility**. Shifts in culture start with training. NSF could add reproducibility training to its research fellowship and traineeship programs, ensuring that new scientists are being mentored in these areas. NSF could also include reproducibility in training requirements like Responsible Conduct of Research training.
- **Fund projects to develop and test new models of scientific investigation and communication, including Registered Reports**.[6] NSF could fund pilots where awards and publications are based upon a review of the importance of the research question and the quality of the methods. Presently, a bias against reporting negative or null results exists because the perceived likelihood of having those results published is very low. Registered Reports include peer review before results are known to eliminate this reporting bias--the rewards are earned regardless of the outcome. Further, peer feedback occurs early enough to meaningfully impact the research rather than after the work has been completed and the manuscript has been written. Another model that could be developed is that of adversarial collaboration, where a study is conducted by two experimenters with competing hypotheses. The experimenters collaborate on the design and methods until both are satisfied that their hypotheses can be fairly tested. This model could be especially useful to make progress on contentious issues.
- **Promote and support the release of preprints for the rapid dissemination of research.** A preprint is a manuscript that precedes a peer-reviewed publication. The latter can take months to years to reach the scientific community. Peer-review is an important aspect of science, but discourse, evaluation, and feedback can occur within the community prior to publication if it is made available to them. Preprints are standard practice in some fields (e.g., physics) and mostly unknown in others (e.g., life sciences). Promotion could include encouragement for the immediate release of manuscripts as preprints as well as citation of preprints in grant applications and reporting.
- **Promote and fund the sharing and reuse of all components of the research workflow including publications and data; study the outcomes of an open approach.** Momentum is increasing around the opening of publications (open access) and data (open data), but there are other components of the scientific process useful for reproducibility and

[6] See https://cos.io/rr/.

reuse including software and analyses. Rather than waiting for momentum to build around each of these individual components, an open workflow approach could be adopted immediately--encouraging the sharing of as much of the research process as possible.

- **Convene communities to discuss and adopt recommendations and guidelines to shift incentives in order to align scientific values with scientific practices**.[7] Cultural change could be accelerated by a multi-faceted approach across stakeholders in the ecosystem, including journals, societies, and tenure committees. Alignment with other stakeholders or recommendations to other stakeholders could create the momentum needed to increase the pace of change.

Closing

The scientific process that continuously improves our current understanding of the world is itself continuously improving. The knowledge acquired by this process is only made better by critique and new evidence because the process can only lead in one direction: towards understanding. At any given time, that understanding may not be what we want to hear, or it may be more ambiguous than we would like--indicating that the problem is perhaps more complex than we thought. Either way, we may choose to go in the opposite direction, ignoring the signs that read "Wrong Way" and "Turn Back". Regardless of how stubborn we are, the process will continue working, and, when we are ready to trust it again, it will steadily lead us back towards understanding.

When we invest in NSF, we are investing in this process. And when we invest in openness and reproducibility, we are making the path towards understanding more recognizable, the terrain easier to navigate, the trek less lonely, and the warning signs more insistent.

This path leads us incrementally towards the next innovation that increases the quality of life here and abroad. I would like to see us get there as quickly as possible, and I believe that an increased focus on openness and reproducibility will do just that.

We at the Center for Open Science would be glad to continue this discussion and support the efforts of this committee and NSF in pursuit of increased openness and reproducibility.

Thank you for the opportunity to speak with you today. I will be happy to answer any questions you may have.

[7] As an example, see http://cos.io/top for the Transparency and Openness Promotion (TOP) Guidelines developed by community stakeholders and endorsed and adopted by journals, funders, and societies.