Testimony of

Brian A. Nosek, Ph.D.

Executive Director
Center for Open Science

Professor
Department of Psychology
University of Virginia

Before the

Senate Committee on Homeland Security and Governmental Affairs
Federal Spending Oversight Subcommittee
U.S. Senate

October 18, 2017

"Broken Beakers: Federal Support for Research"

Chairman Paul, Ranking Member Peters, and Members of the Subcommittee, on behalf of myself and the Center for Open Science, thank you for the opportunity to discuss the funding of scientific research and the role of transparency and reproducibility to maximize the return on those investments.

Science may be humanity's most important long-term investment. The effort to accumulate knowledge has profound consequences for growing the economy, security, and well-being of American society. Some of the impact of scientific investments are the direct aims of the project. But, much of the impact is indirect. Research often leads to unexpected insights and applications that are only appreciated after the discovery and produce many orders of magnitude return on investments. Examples include the creation of the Internet and the core search insights that led to the founding of Google. Science is a long-term investment strategy for the continuing vitality of society, and we can do an even better job with this investment. Today, I will discuss how promoting a culture of transparency and reproducibility will improve the speed and efficiency of scientific discovery.

In 2002, I became a professor at the University of Virginia in the Department of Psychology. My research group does fundamental research on implicit social cognition -- thoughts and feelings that occur outside of conscious awareness or conscious control. Our interest is interested in why behavior does not always align with values and intentions. For example, a citizen may want to be objective in evaluating evidence for public policies, but may nevertheless be

influenced by the party identity of the proposer independent of the policy itself. My team's research has been supported by federal grants from NIH and NSF.

Since 2013, I have been on extended leave from my faculty position because a graduate student and I started a technology and culture change company called the Center for Open Science based on two projects from my laboratory. The Center for Open Science has a mission to increase openness, integrity, and reproducibility of research. To meet that mission, the Center conducts research on the causes of irreproducibility in science; builds free, open-source software to help researchers manage, archive, and share the materials and data underlying their research; and works with stakeholders across the research community (funders, publishers, societies, institutions) to shift the incentives for researchers to reward transparency and reproducibility.

Transparency and reproducibility are core values of science because they are how science advances knowledge. When I make a claim, you could believe it based my authority as an expert, or how confident I seem making it. These are not sufficient for scientific claims. For credibility of scientific claims, I need to show how I arrived at the claim. By showing you my methodology, the data I collected, and how I analyzed and interpreted that data, you can make an independent assessment. You might recognize a flaw, think of an alternative approach, or have an idea about how to extend what I did to learn more about the phenomenon. Moreover, by sharing how I arrived at the claim, I give you the opportunity to reproduce the evidence. If you can independently obtain similar results, then our confidence in the claim increases. The challenge, and the reason for the Center for Open Science's existence, is that the culture of incentives for scientists sometimes undercuts those core values of transparency and reproducibility.

The currency of academic science is the publication. Researchers earn jobs, grants, tenure, and acclaim by publishing frequently and in prestigious journals. In the present culture, novel results are more publishable than replications or additional evidence for existing claims. Positive results--finding a relationship between things or that an intervention did impact an outcome--are more publishable than negative results--finding no relationship or effectiveness. And, tidy results are more publishable than results with exceptions or loose ends. Novel, positive, tidy results are the best kind of results, but achieving them is difficult. Scientists are studying things that they do not yet understand. Progress is slow and uncertain, often with many dead-ends. But, because I am rewarded for publishability, I might make decisions--perhaps without even realizing it--that maximize publishability at the cost of accuracy. For example, we run many experiments, but we might only select to publish the ones that make the best story and rationalize ignoring the others as having flawed methodology. Also, there are many ways to analyze a dataset to decide what has been learned. I could use that discretion and convince myself that the analysis strategy yielding the most publishable analysis strategy is also the correct analysis strategy. If these behaviors occur, then the findings in my published papers will be less credible than they appear.

The solution to these human behavior challenges are transparency and reproducibility. Science is done by scientists. Even with the best of intentions, scientists' behavior is shaped by their culture and incentive systems. Unintended biases permeate decision-making and are insidious because they are so difficult to identify in ourselves. Transparency is a mechanism for identifying bias. If you knew about the other studies and analyses I conducted but did not report, then you would be able to adjust your impression of the credibility of my findings. Also, with transparency, you could examine my methodology and data and try to independently reproduce the findings yourself. If you don't know how I arrived at my claims, then there is no way for you to properly assess their credibility. If others cannot reproduce my claims independently, then either I haven't communicated sufficiently how to obtain the results, or the results are not sufficiently reproducible to be credible. Science relies on transparency and reproducibility for the open, unregulated marketplace of ideas to function efficiently.

By and large, scientists endorse the values of transparency and reproducibility. But, the culture rewards novel, positive, clean results; there are few incentives for being open or transparent or reproducible. Federal research funding agencies are aware of this problem and have only taken initial steps to address it. As a consequence, openness is rare, and reproducibility appears to be lower than is desirable. This can change. If transparency and reproducibility are incorporated into the policies and incentives shaping the culture and researchers' behavior, we may reduce waste and increase the pace and efficiency of discovery and ultimately earn even greater return on investment of taxpayer dollars.

Moreover, with transparency, we can improve the credibility and understanding of how science works to the taxpaying public. For example, it is easy for the public to understand the value of research investment on applied questions such as "Will this drug reduce cancer?", "Will this battery store more energy for more time?", and "Will this cognitive therapy reduce the symptoms of PTSD?" The direct outcome of interest for such investigations is justification enough for supporting the research.

Transparency can help the public also see the benefits of investments in fundamental science questions. Fundamental, or basic, science is the breeding ground for identifying opportunities to answer those applied questions. How can we know what drugs to try for cancer? We need to know more about the biology of cells. How can we know what materials might lead to longer battery life? We need to know more about how materials store electrical energy as chemical energy. How can we know what cognitive techniques are best for PTSD? We need to know more about human reasoning and emotion and the relationship between thoughts and actions.

Fundamental science is a patient, long-term, high-reward investment strategy. If we understand the core of how something works, then it may have immediate impact on known implications. But, it can also have massive impact on entirely unexpected applications. The emergence of Google out of federal grants for advancing digital libraries is a dramatic example. A more modest example is my own experience.

The products and services provided by the Center for Open Science to improve research efficiency is an unexpected outcome of an R01 grant to my lab from the National Institute of Mental Health. I was studying thoughts and feelings outside of conscious awareness and conscious control. We were trying to figure out why people's social values such as "I want to treat people fairly regardless of their race or gender." did not always translate into their behavior. While studying that, we learned how cultures and social circumstances can influence behavior, sometimes even leading people to behave contrary to their own values without recognizing it. Over years, we came to realized that these general principles of decision-making could be applied to how scientific culture shape scientists' behaviors. This was not anticipated in the R01 application. It is only with the benefit of hindsight that we can see how investigating those fundamental questions led to ideas that spawned the Center for Open Science, its products and services, and its progress on the mission to increase openness, integrity, and reproducibility of research.

Because we don't yet know what we will learn, there is little chance of anticipating which fundamental science investigations will lead to the massive returns on investment. If we knew that, then it would be easy to decide what to study and what to fund. Without a crystal ball, science relies on the next best thing--peer review. I write grants proposing to study what I believe are the next most important questions in my field. Others with similar expertise evaluate my proposals for their merit and anticipated impact. Their prediction of impact is necessarily limited to the ability to imagine what will be learned and how that knowledge could be used. Nevertheless, independent domain experts are best positioned to make those judgments because they possess depth insight on both what we know and what we don't know, and thus have some idea of the implications of filling the knowledge gaps.

Some parts of the peer review process are transparent and other parts are privacy protected. As a representative for the Center for Open Science, I believe that there is room to improve transparency of the peer review process and still protect the interests motivating privacy protections. The benefits of greater transparency in peer review would be to demonstrate the credibility and integrity of the process to the taxpaying public to whom science funding decisions can feel daunting, distant, and obscure.

I will close with a specific suggestion that the committee could pursue to help further improve the return on investment in scientific research: **Set the default to open for papers, data, and materials**. In 2013, federal agencies were asked by the White House to make a plan for improving the management and accessibility of data and materials for research that they fund. Most agencies have completed this work. Congress could take the next logical step and require each Federal research funding agency to develop policies that require the research data and materials generated by federal dollars be made publicly accessible by default upon publication of the findings or completion of the grant period. Changing the default from closed to open for research created with public dollars would alter cultural expectations and behavior. Instead of needing to generate reasons to share data, researchers would need to provide justification for delay or not sharing at all due to proprietary or privacy concerns.

Public investment in science leads to solutions, cures, and entirely unexpected advancements that benefit American society. Making open the default for scientific research data would transform science, dramatically increase the return on investment from publicly funded research, and accelerate progress. This is not a very expensive or difficult proposition, but it does require a mandate. This one action would dramatically increase the public benefit from our investments in science.

Thank you members of the committee for your continuing support of science and for the opportunity to speak with you today.