SOCIAL SCIENCE

Promoting Transparency in Social Science Research

Social scientists should adopt higher transparency standards to improve the quality and credibility of research.

E. Miguel,[1]* C. Camerer,[2] K. Casey,[3] J. Cohen,[3] K. M. Esterling,[4] A. Gerber,[5] R. Glennerster,[6] D. P. Green,[7] M. Humphreys,[7] G. Imbens,[3] D. Laitin,[3] T. Madon,[1] L. Nelson,[1] B. A. Nosek,[8,9] M. Petersen,[1] R. Sedlmayr,[10] J. P. Simmons,[11] U. Simonsohn,[11] M. Van der Laan[1]

There is growing appreciation for the advantages of experimentation in the social sciences. Policy-relevant claims that in the past were backed by theoretical arguments and inconclusive correlations are now being investigated using more credible methods. Changes have been particularly pronounced in development economics, where hundreds of randomized trials have been carried out over the last decade. When experimentation is difficult or impossible, researchers are using quasi-experimental designs. Governments and advocacy groups display a growing appetite for evidence-based policy-making. In 2005, Mexico established an independent government agency to rigorously evaluate social programs, and in 2012, the U.S. Office of Management and Budget advised federal agencies to present evidence from randomized program evaluations in budget requests (1, 2).

Accompanying these changes, however, is a growing sense that the incentives, norms, and institutions under which social science operates undermine gains from improved research design. Commentators point to a dysfunctional reward structure in which statistically significant, novel, and theoretically tidy results are published more easily than null, replication, or perplexing results (3, 4). Social science journals do not mandate adherence to reporting standards or study registration, and few require data-sharing. In this context, researchers have incentives to analyze and present data to make them more "publishable," even at the expense of accuracy. Researchers may select a subset of positive results from a larger study that overall shows mixed or null results (5) or present

exploratory results as if they were tests of prespecified analysis plans (6).

These practices, coupled with limited accountability for researcher error, have the cumulative effect of producing a distorted body of evidence with too few null effects and many false-positives, exaggerating the effectiveness of programs and policies (7–10). Even if errors are eventually brought to light, the stakes remain high because policy decisions based on flawed research affect millions of people.

In this article, we survey recent progress toward research transparency in the social sciences and make the case for standards and practices that help realign scholarly incentives with scholarly values. We argue that emergent practices in medical trials provide a useful, but incomplete, model for the social sciences. New initiatives in social science seek to create norms that, in some cases, go beyond what is required of medical trials.
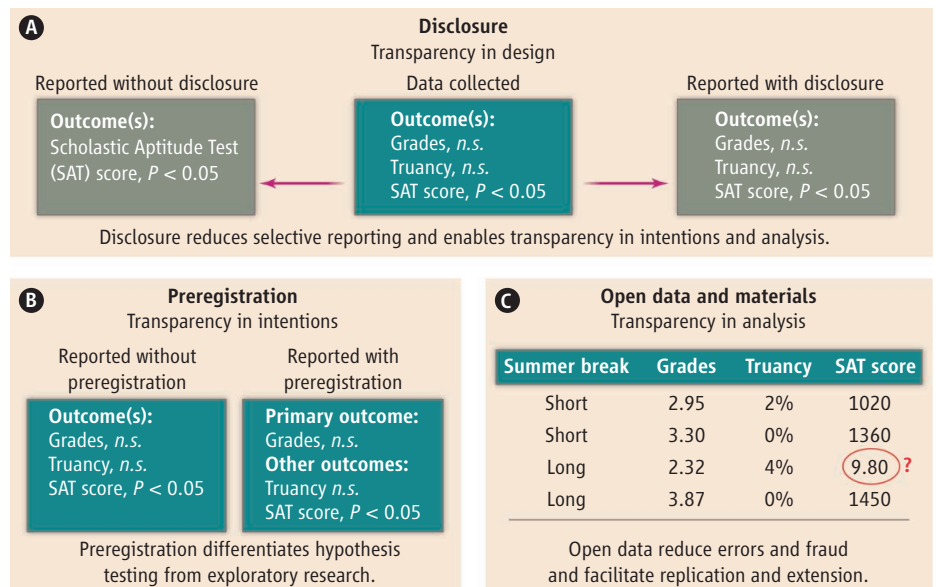
Promoting Transparent Social Science
Promising, bottom-up innovations in the social sciences are under way. Most converge

on three core practices: disclosure, registration and preanalysis plans, and open data and materials (see the chart).

Disclosure. Systematic reporting standards help ensure that researchers document and disclose key details about data collection and analysis. Many medical journals recommend or require that researchers adhere to the CONSORT reporting standards for clinical trials. Social science journals have begun to endorse similar guidelines. The *Journal of Experimental Political Science* recommends adherence to reporting standards, and *Management Science* and *Psychological Science* recently adopted disclosure standards (6). These require researchers to report all measures, manipulations, and data exclusions, as well as how they arrived at final sample sizes (see supplementary materials).

Registration and preanalysis plans. Clinical researchers in the United States have been required by law since 2007 to prospectively register medical trials in a public database and to post summary results. This helps create a public record of trials that might otherwise go unpublished. It can also serve the

[1]University of California, Berkeley, CA, USA. [2]California Institute of Technology, Pasadena, CA, USA. [3]Stanford University, Stanford, CA, USA. [4]University of California, Riverside, CA, USA. [5]Yale University, New Haven, CT, USA. [6]Massachusetts Institute of Technology, Cambridge, MA, USA. [7]Columbia University, New York, NY, USA. [8]University of Virginia, Charlottesville, VA, USA. [9]Center for Open Science, Charlottesville, VA, USA. [10]Wellspring Advisors, New York, NY, USA. [11]University of Pennsylvania, Philadelphia, PA, USA. *Corresponding author. emiguel@berkeley.edu

A Disclosure
Transparency in design

Reported without disclosure | Data collected | Reported with disclosure

Outcome(s):
Scholastic Aptitude Test (SAT) score, $P < 0.05$

Outcome(s):
Grades, *n.s.*
Truancy, *n.s.*
SAT score, $P < 0.05$

Outcome(s):
Grades, *n.s.*
Truancy, *n.s.*
SAT score, $P < 0.05$

Disclosure reduces selective reporting and enables transparency in intentions and analysis.

B Preregistration
Transparency in intentions

Reported without preregistration | Reported with preregistration

Outcome(s):
Grades, *n.s.*
Truancy, *n.s.*
SAT score, $P < 0.05$

Primary outcome:
Grades, *n.s.*
Other outcomes:
Truancy *n.s.*
SAT score, $P < 0.05$

Preregistration differentiates hypothesis testing from exploratory research.

C Open data and materials
Transparency in analysis

Summer break	Grades	Truancy	SAT score
Short	2.95	2%	1020
Short	3.30	0%	1360
Long	2.32	4%	9.80 ?
Long	3.87	0%	1450

Open data reduce errors and fraud and facilitate replication and extension.

Three mechanisms for increasing transparency in scientific reporting. Demonstrated with a research question: "Do shorter summer breaks improve educational outcomes?" n.s. denotes $P > 0.05$.

purpose of prespecification in order to more credibly distinguish hypothesis testing from hypothesis generation. Social scientists have started registering comprehensive preanalysis plans—detailed documents specifying statistical models, dependent variables, covariates, interaction terms, and multiple testing corrections. Statisticians have developed randomized designs to address the problem of underpowered subgroup analysis using prespecified decision rules (11, 12).

Open data and materials. Open data and open materials provide the means for independent researchers to reproduce reported results; test alternative specifications on the data; identify misreported or fraudulent results; reuse or adapt materials (e.g., survey instruments) for replication or extension of prior research; and better understand the interventions, measures, and context—all of which are important for assessing external validity. The American Political Science Association in 2012 adopted guidelines that made it an ethical obligation for researchers to "facilitate the evaluation of their evidence-based knowledge claims through data access, production transparency, and analytic transparency." Psychologists have initiated crowd-sourced replications of published studies to assess the robustness of existing results (13). Researchers have refined statistical techniques for detecting publication bias and, more broadly, for assessing the evidentiary value of a body of research findings (14, 15).

Other recent initiatives create infrastructure to foster and routinize these emerging practices. Organizations are building tools to make it easier to archive and share research materials, plans, and data. The Open Science Framework is an online collaboration tool developed by the Center for Open Science (COS) that enables research teams to make their data, code, and registered hypotheses and designs public as a routine part of within-team communication. The American Economic Association launched an online registry for randomized controlled trials. The Experiments in Governance and Politics network has an online tool for preregistering research designs. *Perspectives on Psychological Science* and four other psychology journals conditionally accept preregistered designs for publication before data collection. Incentives are being created for engaging in transparency practices through COS supported "badges" to certify papers that meet open-materials and preregistration standards.

Improving on the Medical Trial Model

The move toward registration of medical trials, at a minimum, has helped reveal limitations of existing medical trial evidence (16).

Registration could similarly aid the social science community. But study registration alone is insufficient, and aspects of the medical trial system can be improved (17). The current clinical trial registration system requires only a basic analysis plan; the presence of a detailed plan varies substantially across studies. Best practices are less developed for observational clinical research.

The centralized medical trials registration system has benefits for coordination and standards setting but may have drawbacks. For instance, the U.S. Food and Drug Administration's dominant role in setting clinical research standards arguably slows adoption of innovative statistical methods. A single registration system will have similar limitations if applied as a one-size-fits-all approach to the diverse array of social science applications. Active participation from all corners of social science are needed to formulate principles under which registration systems appropriate to specific methods—lab or field experiment, survey, or administrative data—may be developed.

Several authors of this article established the Berkeley Initiative for Transparency in the Social Sciences (BITSS) to foster these interdisciplinary conversations. Working with COS, BITSS organizes meetings to discuss research transparency practices and training in the tools that can facilitate their adoption. Interdisciplinary deliberation can help identify and address limitations with existing paradigms. For example, the complexity of social science research designs—often driven by the desire to estimate target parameters linked to intricate behavioral models—necessitates the use of preanalysis plans that are much more detailed than those typically registered by medical researchers. As medical trials increasingly address behavioral issues, ideas developed in the social sciences may become a model for medical research.

The scope of application of transparency practices is also important. We believe that they can be usefully applied to both nonexperimental and experimental studies. Consensus is emerging that disclosure and open materials are appropriate norms for all empirical research, experimental or otherwise. It is also natural to preregister prospective nonexperimental research, including studies of anticipated policy changes. An early preanalysis plan in economics was for such a study (18).

Further work is needed. For instance, it is unclear how to apply preregistration to the analyses of existing data, which account for the vast majority of social science. Development of practices appropriate for existing

data—whether historical or contemporary, quantitative, or qualitative—is a priority.

Exploration and Serendipity

The most common objection to the move toward greater research transparency pertains to preregistration. Concerned that preregistration implies a rejection of exploratory research, some worry that it will stifle creativity and serendipitous discovery. We disagree.

Scientific inquiry requires imaginative exploration. Many important findings originated as unexpected discoveries. But findings from such inductive analysis are necessarily more tentative because of the greater flexibility of methods and tests and, hence, the greater opportunity for the outcome to obtain by chance. The purpose of prespecification is not to disparage exploratory analysis but to free it from the tradition of being portrayed as formal hypothesis testing.

New practices need to be implemented in a way that does not stifle creativity or create excess burden. Yet we believe that such concerns are outweighed by the benefits that a shift in transparency norms will have for overall scientific progress, the credibility of the social science research enterprise, and the quality of evidence that we as a community provide to policy-makers. We urge scholars, journal editors, and funders to start holding social science to higher standards, demanding greater transparency, and supporting the creation of institutions to facilitate it.

References and Notes

1. M. F. Castro *et al.*, "Mexico's M&E system: Scaling up from the sectoral to the national level" (Working paper, World Bank, Washington, DC, 2009).
2. Office of Management and Budget, "Use of evidence and evaluation in the 2014 budget" (Memo M-12-14, Office of the President, Washington, DC, 2012).
3. B. A. Nosek *et al.*, *Perspect. Psychol. Sci.* **7**, 615 (2012).
4. A. S. Gerber, N. Malhotra, *Q. J. Polit. Sci.* **3**, 313 (2008).
5. K. Casey *et al.*, *Q. J. Econ.* **127**, 1755 (2012).
6. J. P. Simmons *et al.*, *Psychol. Sci.* **22**, 1359 (2011).
7. J. P. A. Ioannidis, *PLOS Med.* **2**, e124 (2005).
8. W. Stroebe *et al.*, *Perspect. Psychol. Sci.* **7**, 670 (2012).
9. M. Humphreys *et al.*, *Polit. Anal.* **21**, 1 (2013).
10. D. M. Lane, W. P. Dunlap, *Br. J. Math. Stat. Psychol.* **31**, 107 (1978).
11. M. Rosenblum, M. J. Van der Laan, *Biometrika* **98**, 845 (2011).
12. D. Rubin, *Ann. Appl. Statist.* **2**, 808 (2008).
13. Open Science Collaboration, *Perspect. Psychol. Sci.* **7**, 657 (2012).
14. A. S. Gerber *et al.*, *Polit. Anal.* **9**, 385 (2001).
15. U. Simonsohn *et al.*, *J. Exp. Psychol.* 10.1037/a0033242 (2013).
16. S. Mathieu *et al.*, *JAMA* **302**, 977 (2009).
17. D. A. Zarin, T. Tse, *Science* **319**, 1340 (2008).
18. D. Neumark, *Industr. Relat.* **40**, 121 (2001).