

## ONE

# Introduction

### THE NEED FOR TRANSPARENT SOCIAL SCIENCE RESEARCH

Contemporary society is complex and rapidly changing. Leaders of government, corporate, and nonprofit institutions all face a constant stream of choices. Thankfully, these leaders are increasingly investing in data acquisition and analysis to help them make good decisions. Researchers are often charged with providing this information and insight, in areas ranging from environmental science to economic policy, immigration, and health care reform. Success often depends on the quality of the underlying research. Inaccurate research can lead to ineffective or inappropriate policies, and worse outcomes for people's lives.

How reliable is the current body of evidence that feeds into decision making? Many believe it is not reliable enough. A crisis of confidence has emerged in social science research, with influential voices both within academia (Manski 2013) and beyond (Feilden 2017) asserting that policy-relevant research is often less reliable than claimed, if not outright wrong. The popular view that you can manipulate statistics to get any answer you want captures this loss of faith in the research enterprise, and the sense that too many scientific findings are mere advocacy. In this era of “fake news” and the rise of extremist political and religious movements around the world, the role of scientific research in establishing the truth as common ground for public debate is more important than ever.

Let's take, for example, the case of health care reform in the United States—the subject of endless partisan political debate. This tension can

be partly explained by the simple fact that people feel strongly about health care, a sector that affects everyone at one time or another in their lives. But there are also strong ideological disagreements between the major U.S. political parties, including the role government should play in providing social services, and the closely related debate over tax rates, since higher taxes generate the revenue needed for health programs.

What role can research play in such a volatile debate? The answer is “It depends.” Some people—and politicians—will hold fast to their political views regardless of evidence; research cannot always sway everyone. But data and evidence are often influential and even decisive in political battles, including the 2017 attempt by congressional Republicans to dismantle the Affordable Care Act (ACA), or Obamacare. In that instance, a handful of senators were swayed to vote “Nay” when evidence from the Congressional Budget Office estimating the likely impact of ACA repeal on insurance coverage and health outcomes was released. Media coverage of the research likely boosted the program’s popularity among American voters.

The answers to highly specific or technical research questions can be incredibly important. In the U.S. case, findings about how access to health insurance affects individual life outcomes—including direct health measures, as well as broader economic impacts such as personal bankruptcy—have been key inputs into these debates. How many people will buy insurance under different levels of subsidies (i.e., what does the demand curve for health insurance look like)? How do different institutional rules in the health insurance marketplace affect competition, prices, and usage? And so on.

When the stakes are this high, the accuracy and credibility of the evidence used become extremely important. Choices made on the basis of evidence will ultimately affect millions of lives. Importantly, it is the responsibility of social science researchers to assure others that their conclusions are driven by sound methods and data, and not by some underlying political bias or agenda. In other words, researchers need to convince policymakers and the public that the statistical results they provide have evidentiary value—that you can’t just pick out (or make up) any statistic you want.

This book provides a road map and tools for increasing the rigor and credibility of social science research. We are a team of three authors—one sociologist and two economists—whose goal is to demonstrate the role that greater *research transparency and reproducibility* can play in uncovering and documenting the truth. We will lay out a number of specific

changes that the research community can make to advance and defend the value of scientific research in policy debates around the world. But before we get into the nitty-gritty or “how,” it is worth surveying the rather disappointing state of affairs in social science research, and its implications.

### **HOUSTON, WE HAVE A PROBLEM: RESEARCH FRAUD AND ITS AFTERMATH**

If you thought we’d have research methods all figured out after a couple centuries of empirical social science research, you would be wrong. A rash of high-profile fraud cases in multiple academic disciplines and mounting evidence that a number of important research findings cannot be replicated both point to a growing sense of unease in the social sciences. We believe the research community can do better.

Fraud cases get most of the headlines, and we discuss a few of the most egregious cases here. By mentioning these examples, we are *not* claiming that most researchers are engaging in fraud! We strongly believe that outright fraud remains the exception rather than the rule (although the illicit nature of research fraud makes it hard to quantify this claim or even assert it with much confidence). Rather, fraud cases are the proverbial canaries in the coal mine: a dramatic symptom of a much more pervasive underlying problem that manifests itself in many other ways short of fraud. We will discuss these subtler and more common problems—all of which have the ability to distort social science research—at length in this book.

The field of social psychology provides a cautionary tale about how a lack of transparency can lead to misleading results—and also how the research community can organize to fight back against the worst abuses. In recent years, we have seen multiple well-publicized cases in which prominent tenured social psychologists, in both North America and Europe, were caught fabricating their data. These scholars were forced to resign from their positions when colleagues uncovered their misdeeds. In the circles of scientific hell, this one—simply making stuff up and passing it off as science—must be the hottest (Neuroskeptic 2012).

Perhaps best known is the case of Diederik Stapel, former professor of psychology at Tilburg University in the Netherlands. Stapel was an academic superstar. He served as dean of social and behavioral sciences, was awarded multiple career prizes by age 40, and published 150 articles, including in the most prestigious journals and on socially important topics, including the psychology of racial bias (Carey 2011; Bhattacharjee

2013). Academic careers rise and fall largely on the basis of publishing (or not publishing) articles in top research journals, which is often predicated on successful fund-raising, and according to these metrics Stapel was at the very top of his field.

Unfortunately, Stapel's findings and publications were drawn mostly from fabricated data. In his autobiography, written after the fraud was discovered, Stapel describes his descent into dishonesty, and how the temptation to alter his data in order to generate exciting research results—the kind he felt would be more attractive to top journals and generate more media attention—was too much for him to resist:

Nobody ever checked my work. They trusted me. . . . I did everything myself, and next to me was a big jar of cookies. No mother, no lock, not even a lid. . . . Every day, I would be working and there would be this big jar of cookies, filled with sweets, within reach, right next to me—with nobody even near. All I had to do was take it. (quoted in Borsboom and Wagenmakers, 2013)

As Stapel tells it, he began by subtly altering a few numbers here and there in real datasets to make the results more interesting. However, over time he began to fabricate entire datasets. While Stapel was certainly at fault, we view his ability to commit fraud undetected as an indictment of the entire social science research process. Still, there were many warning signs. Stapel never shared his data with others, not even his own graduate students, preferring to carry out analyses on his own. Over time, suspicions began to snowball about the mysterious sources of his data and Stapel's "magical" ability to generate one blockbuster article after another, each with fascinating constellations of findings.

Ultimately, a university investigation led to Stapel's admission of fraud and his downfall: he retracted at least 55 articles (including from leading research journals like *Science*), was forced to resign from his position at Tilburg, and was stripped of his Ph.D. Criminal proceedings were launched against him (they were eventually settled). The article retractions further discredited the work of his students and colleagues—collateral damage affecting dozens of other scholars, many of whom were supposedly ignorant of Stapel's lies.

Stapel's autobiography is a gripping tale of his addiction to research fraud. At times it is quite beautifully and emotionally written (by all accounts, though we have not read it in the original Dutch). It emerged after the book was published, however, that several of the most moving

passages were composed of sentences that Stapel had copied (into Dutch) from the fiction writers Raymond Carver and James Joyce. Yet he presented them without quotes and only acknowledged his sources separately in an appendix! Even in his mea culpa, the dishonesty crept in (Borsboom and Wagenmakers 2013).

How many other Stapels are out there? While it is impossible to say, of course, there are enough cases of fraud to provoke concern. No academic field is immune.

Roughly a quarter of economics journal editors say they have encountered cases of plagiarism (Enders and Hoover 2004). Political science was rocked by a fraud scandal in 2015, when David Broockman, then a graduate student at the University of California, Berkeley, discovered that a *Science* paper on the impact of in-person canvassing on gay rights attitudes, written by Michael LaCour and Don Green, contained fabricated data (Broockman, Kalla, and Aranow 2015). While Green was cleared of wrongdoing—he had not collected the data and was apparently unaware of the deception—the incident effectively ended LaCour’s promising academic career: at the time, he was a graduate student at the University of California, Los Angeles, and had been offered a faculty position at Princeton, which was later withdrawn.

These cases are not ancient history: they took place just a few years back. While some progress is already being made toward making research more transparent and reproducible (as we will discuss in detail throughout this book), it remains likely that other instances of data fabrication will (unfortunately) occur. Many of the problems with the research process that allowed them to occur—such as weak data-sharing norms, secrecy, limited incentives to carry out replications or prespecify statistical analyses, and the pervasive publish-or-perish culture of academia—are still in place, and affect the quality of research even among the vast majority of scholars who have never engaged in outright fraud. Even if rare, cases of scholarly fraud also garner extensive media coverage and are likely to have outsize influence on the perceptions of social scientists held by the general public, policymakers, and potential research donors.

How can we put a lid on Stapel’s open cookie jar to prevent research malpractice from happening in the future? With science already under attack in many quarters, how can we improve the reliability of social science more broadly, and restore public confidence in important findings? This book aims to make progress on these issues, through several interconnected goals.

**BOOK OVERVIEW**

First, we aim to bring the reader up to speed on the core intellectual issues around research transparency and reproducibility, beginning with this introduction and continuing in Chapter 2 with a detailed discussion of the scientific ethos and its implications for research practices.

Next, we present existing evidence—some classic, some new—on pervasive *problems* in social science research practice. One such problem is publication bias (Chapter 3), whereby studies with more compelling results are more likely to be published, rather than publication being based solely on the quality of the data, research design, and analysis. Another distinct, but closely related, problem is specification searching during statistical analysis (Chapter 4). Specification searching is characterized by the selective reporting of analyses within a particular study, generating misleading conclusions. By now, there is ample evidence that both of these problems are real and widespread, leading to biased bodies of research evidence.

The documented existence of these problems sets the stage for a series of methodological *solutions* designed to address them. Some of these solutions are well known, including approaches that enable scholars to use all possible data across studies (through study registries and meta-analysis) to reach more robust conclusions (Chapter 5). The use of prespecified hypothesis plans to discipline analysis and boost accountability harkens back to our most fundamental understanding of the scientific method (Chapter 6). We present a “how-to” guide for utilizing pre-analysis plans in practice. Meanwhile, sensitivity analyses and other antidotes to specification searching often rely on recent advances in statistics and econometrics (Chapter 7). We illustrate these tools using current examples from across the social sciences—economics, political science, psychology, and sociology.

Unfortunately, these well-intended solutions are only as effective as they are widely adopted. For outcomes to change, *practices*, norms, and institutions must also change. One change discussed in this book is the adoption of reporting standards and disclosure practices that structure the presentation of data and the design of studies (Chapter 8). Another is replication, a practice critical for enhancing accountability and discovering problems in existing work (Chapter 9). Beyond discussing the technicalities of each practice, we note how the incentives that researchers encounter often discourage replication and suggest ways to move fields toward more productive research norms.

Another critical practice for enhancing accountability is the open sharing of data and other research materials (Chapter 10). Still, there are many unresolved questions around safely sharing personal data without violating individual confidentiality. This is an area of current interest across disciplines. Thankfully, social scientists are finally beginning to adopt beneficial reproducible coding and workflow practices from computer science and data science. We discuss the adaptation of these practices to the social sciences in Chapter 11. Throughout the book, we provide technical material for readers interested in the statistical and computational details of these approaches, and for those seeking to apply them to their own research.

Finally, we discuss the evolving landscape in the areas of research transparency and reproducibility, the institutional changes that could buttress recent progress, and the importance of changing research norms in order to achieve sustainable progress (Chapter 12).

The audience for this book is intentionally broad (although we are happy to preregister our hypothesis that it is unlikely to end up a national best seller sold in airport magazine stands). Doctoral and master's-level students are perhaps its most natural users. We hope that young scholars will find the ideas presented here both inspiring and useful as they build up their technical skill set and develop their own research workflow. Given the numerous applications and examples we provide, the material should fit nicely into graduate curricula on research methods, study design, statistics, and econometrics, as well as in more specific field courses.

We believe this work will serve as a valuable bookshelf reference for more seasoned scholars who have completed their training, including faculty, postdoctoral scholars, and staff scientists in academic settings, government agencies, and the private sector, as well as for research funders, publishers, and the end consumers of social science research. Gaining a better understanding of the threats to and solutions for improving the credibility of social science is critical for anyone producing or consuming research evidence. While some of the problems we discuss are fairly well known (if not yet widely taught), many of the solutions and practices that aim to enhance research transparency and reproducibility are new to the social sciences and could be useful for scholars at all career stages.

Highly motivated undergraduates with strong training in statistics and some familiarity with social science research can also gain from reading this book. We relegate some of the more technical material to

appendices and text boxes throughout, specifically to make the core text more widely accessible to undergraduates and others who are not (yet) professional researchers.

Additionally, we envision this book as a resource for graduate and undergraduate research assistants (as well as more open-minded coauthors) who are just becoming acquainted with scientific ideals and practices. Just as there is more to capturing the spirit of a great musician than learning to play the right notes, there is more to being a good scholar than simply learning how to code in R or Stata, or memorizing your field's canonical papers. The best scholars carry out research in an intellectually balanced way, with the right ethos and an open mindset. With this book, we aim to crystallize these ideals and put them into practice.

Finally, some of the material in this book has been incorporated into a massive open online course (MOOC) entitled “Transparent and Open Social Science Research” with UK-based FutureLearn, using audio from a graduate course recently taught at UC Berkeley by two of the authors. This online course contains homework exercises, videos, and discussion forums that complement this textbook. We encourage readers interested in digging deeper to check it out.