

1 Introduction

John Gerring, James Mahoney and Colin Elman

In recent years, methods of data collection in the social sciences have expanded in range and sophistication. New data sources (many of them hosted on the worldwide web) and data harvesting techniques (e.g., web crawlers) have been discovered, leading to big-data projects of a sort previously unimaginable (Steinert-Threlkeld 2018). One can now read and electronically code foreign newspapers, government reports, interviews and even archival material without leaving one's office. Techniques for measuring obscure or sensitive attitudes and activities, and developing scales for composite indices, have been refined (Bandalos 2018). Relationships among individuals can be probed with social network tools (Borgatti, Everett and Johnson 2018). The location of subjects and events can be tracked with GIS points and polygons (Steinberg and Steinberg 2005). Surveys can be implemented with less time and cost (especially when subjects are recruited through online platforms like M-Turk, Facebook or Crowdfunder), most field sites are more accessible than ever and qualitative data of all sorts can be recorded in their original form (Kapiszewski, MacLean and Read 2015).

Methodological tools for data analysis have likewise undergone major changes. A revolution in thinking about causal inference has occurred (Morgan and Winship 2015). Quasi-experimental and experimental techniques are now brought to bear on topics previously regarded as purely observational (Druckman and Green 2011; Dunning 2012; Kagel and Roth 2016). Machine learning allows one to make inferences from huge quantities of data (King 2014). New frameworks promise to broaden our thinking about causal inference through causal graphs (Pearl and Mackenzie 2018), Bayesian probability (Fairfield and Charman 2020; Humphreys and Jacobs 2021) and set theory (Schneider and Wagemann 2012). Finally, specialized software (Python, R, SAS, Stan, Stata and so forth) facilitate all of these tasks of data collection and analysis.

Clearly, a great deal of progress has been made over the past several decades. Even so, nagging worries persist about the course of social science.

Progress in these disciplines is hard to assess and core scientific goals such as discovery, transparency, reproducibility and cumulation seem frustratingly out of reach. Despite their impressive technical acumen and the many tools at their disposal, today's social scientists may be only slightly better equipped to vanquish error and construct an edifice of truth than their forbears – who conducted analyses with slide rules and wrote up results with typewriters.

As an example, let us consider the problem of reproducibility. A key issue in the production of knowledge is the reproduction of knowledge. If a finding cannot be reproduced by someone else, it cannot be directly confirmed or disconfirmed. Worryingly, many published findings cannot be repeated – or, if repeated, cannot be reproduced. Publication biases, giving precedence to methods or results that are novel, accentuate this problem. Indeed, published results may have a *lower* probability of being true than unpublished results. Insofar as replication serves as a hallmark of science, social science is falling woefully short.¹

Even if studies are replicated perfectly, one may doubt that knowledge in these disciplines would cumulate neatly into a progressive body of research – one that discards false theories and preserves true theories, reaching consensus on an issue and constructing a unified theoretical framework upon which new knowledge can grow. Instead, one often finds that old findings are forgotten (Gans 1992), and the same theories – many of them decades or centuries old – are recycled, without ever being decisively proven or disproved, to be joined by new theories, which enjoy their time in the sun and then a slow demise.² Social science follows fashion, and it is not entirely clear that this year's fashion brings us closer to the truth than last year's fashion.

Our purpose in this volume is to consider the challenges facing the social sciences, as well as possible solutions to those challenges. In doing so, we adopt a *systemic* view of the subject matter. Entire disciplines, with all of their moving parts, constitute our units of analysis. We begin this chapter by laying out the approach. Next, we consider the scope of the volume, which

¹ See Atmanspacher and Maassen (2016), Camerer et al. (2016), Chang and Li (2015), Dewald, Thursby and Anderson (1986), Ioannidis (2005), Open Science Collaboration (2012, 2015). Although there is considerable ambiguity about what it means to replicate, or fail to replicate (Parts III–IV, this volume), we take it for granted that the problem is non-trivial.

² On problems of cumulation and reaching consensus in the social sciences, see Abbott (2001), Chernoff (2014), Cole (1994), Collins (1994), Geller and Vasquez (2005), Johnson (2003), Rule (1997), Simowitz (1998), Sjöblom (1977, 1997), Smith (2005, 2008). On the rise and fall of intellectual movements see Frickel and Gross (2005).

encompasses all of the social sciences. In the third section, we outline the contents of the volume. The final section offers a brief conclusion.

A Systemic Approach

Traditionally, social science methodology has focused on individual studies, i.e., how to conduct a study and how (ex post) to judge the adequacy of that study. This is what one finds, for the most part, in the pages of methods journals and textbooks and in methodology classes. We have learned an enormous amount from this finely honed approach. Indeed, many of the advances signaled at the outset of this chapter may be credited to it.

Even so, the current disorderly state of social science suggests that a piecemeal approach to scientific progress may not be entirely satisfactory. Truth and falsehood are often difficult to discern, even after the most vigilant peer review. Efforts to reduce error (the publication of studies whose findings are untrue) by raising the bar to publication inevitably edge out the most innovative work, which (almost by definition) is less likely to be regarded as true, in favor of work in established research traditions, where the payoff to scientific progress is lower. Replications of work already published are rare, and their results often ambiguous (findings may be supported to some degree or in some respects, but not in others, or the results hinge upon assumptions that cannot be tested). Finally, the bits and pieces of truth that we feel fairly confident about do not fall neatly into place within a larger theoretical scaffolding. Cumulation is not easy.

A study-centered approach to social science will not solve these larger problems. In response, we propose a broader approach, one that focuses on the system within which studies are produced and vetted.

This “system” is hard to bound, as it includes multiple organizations – departments, universities, journals, presses and professional associations – each of which enjoys some degree of independence but all of which interact in crucial ways, affecting each other’s behavior. For present purposes, an academic system will be understood to include all of these interacting parts. It is at least as large as a discipline (e.g., economics) and in some respects transcends individual disciplines (which are in any case overlapping, as discussed below).

We do *not* consider the broader society within which social science is situated. This lies beyond the scope of the volume, though we readily acknowledge that social science is affected by pretty much everything that goes on in society. It is not a closed system.

A systemic framework is intended to complement – not replace – methodology’s traditional focus on individual studies. One sits within the other, and neither makes sense without the other. Specifically, methodological advice with respect to conducting and appraising individual studies must be consistent with meta-level advice about how to conduct one’s professional life.

Unfortunately, conflicts between these two levels are rampant in today’s academy. For example, signals emanating from the community of methodologists suggest that scholars should seek, above all, to avoid Type-1 errors – falsely rejecting a true null hypothesis. At the same time, signals from the academy with respect to hiring, salary and promotion suggest that innovation and productivity are the principal keys to success (Whicker et al. 1993).

A systemic approach to social science brings these potential conflicts into view, forcing us to consider how they might be harmonized and how we might avoid the whack-a-mole dynamic that often ensues when a reform focuses on one scientific goal without considering the effect of that reform on other scientific goals – or when a reform focuses on one corner of the scientific universe without considering the effect of that reform on other corners.

An example of the latter is the current call for replication, intended (among other things) to combat publication bias in favor of studies that reject the null. While the number of replications has grown in recent years it is still miniscule. Nor is it clear that it will solve the problem of publication bias. Indeed, one recent study suggests that replication studies are also subject to a distinctive publication bias of their own – in favor of studies that disconfirm a published study (Berinsky et al. 2018). This sort of ping-pong game (“True”-“Not True”) is not likely to lead to a broad consensus on the topic under investigation.

A systemic approach suggests that the shortcomings of social science are not the fault of insufficiently sophisticated methods. While a study-based methodology certainly has room for improvement, further refinements are probably insufficient to solve the core problems facing the social sciences today.

Nor do we believe that social science falls short because of irresponsible or poorly trained researchers. Granted, some individuals may have a poor grasp of the methods at their disposal and thus make suboptimal choices among them or reach interpretations unwarranted by the evidence at hand. A few individuals falsify data or misreport results. Nonetheless, we believe that most researchers pursue their craft in a reasonable and honest manner.

To the extent that social scientists fail to fulfil the ideals of a progressive science, we believe the blame lies primarily with institutions rather than with

individuals.³ It is not a product of a few bad apples. It is the product of a system with fundamental design flaws.

Institutional analysis is commonplace in the analysis of politics, economics and society (DiMaggio and Powell 1991; Hall and Taylor 1996; North 1986), and it seems reasonable to apply the same analytic lens to the community of scholars. If one wishes to understand the behavior of individuals (e.g., citizens, consumers or terrorists) or organizations (e.g., governments, political parties, interest groups, firms or schools), one must understand the official rules and informal norms that structure behavior within these contexts.

An academic discipline is one such context. A priori, there are good reasons to suppose that institutions structure the behavior of social scientists. Although *academe* is often described as an “individualistic” enterprise, the competitive struggles that characterize our work fit within a common template. Indeed, the structure of rewards – based primarily upon publication in top journals and presses – is nearly identical across fields and at various stages of one’s career. Moreover, the search for truth (define it how you will) is necessarily a communal enterprise. The questions scholars ask are framed by a field’s existing stock of knowledge, they are addressed using research techniques they hold in common and their answers are promulgated to and judged by peers (Merton 1973b). Academics is a highly professionalized and institutionalized field of endeavor.

What, then, are the rules and norms governing behavior in the social sciences? What kinds of research, and what sort of researcher, wins and loses under the current system? In what ways does this incentive structure serve – or subvert – the goal of scientific progress? A wealth of research on institutions has shown that formal rules and informal norms often persist even when they do not facilitate what most participants would regard as an optimal outcome. Institutions are sticky, even when inefficient (David 1985). This leads to our final question. Can institutions that govern the production of knowledge be altered so as to better serve the goals of science?

These are the questions animating the present volume. Thankfully, the systemic aspects of social science have begun to receive greater attention. Responsibilities that individual scholars owe their research communities are being spelled out in a clearer fashion, and conformity to those norms is more closely monitored. The infrastructure needed for scholars to deliver on these obligations is under development. Scholars are beginning to take more

³ One might also blame the subject matter, an issue discussed in the next section.

seriously the ways in which sociological aspects of knowledge production can hinder progress in the social sciences, and they have begun to devise institutional responses to those problems. We seek to take stock of this work and to push it forward – with special focus on the communal context in which individual studies are generated, digested and disseminated.

Social Science as a Field of Endeavor

Problems in realizing scientific progress are not unique to social science (Laudan 1977). However, it seems fair to say that methodological obstacles are more profound in the context of the social sciences than in the context of the natural sciences. As an example, one might consider the so-called “replication crisis.” Researchers in medicine, physics and chemistry often face obstacles in replicating each other’s results. Nonetheless, attempts at replication occur on a regular basis in the natural sciences, while they are rare in the social sciences. Moreover, replications in the natural sciences usually manage, over time, to sort out the good from the bad, or merely ugly – which cannot be said with confidence of the social sciences.

In these respects, we take the traditional view that natural science disciplines are more consistently on track with expectations about scientific progress than social science disciplines. This justifies our focus on the latter, a perennial problem child within the family of sciences.

It is understandable to find confusion and ambiguity in research focused on the decisional behavior of human beings, where outcomes are subject to myriad causes and to contextual variation (including variation over time), where questions of theoretical interest cannot always be studied experimentally, where categories may depend on collective agreement for their meaning and where the results of any study are available to the subjects of interest and may shape their future behavior. Social science is hard. In this light, the non-progressive features of social science are endemic to the enterprise (Collins 1994; Hacking 1999; Winch 1958).

Nonetheless, to say that pathologies are intrinsic to social science does not mean that they are always present in equal degrees. The practice of social science has fundamentally transformed over the past half-century, suggesting a high degree of malleability and the possibility of improvement. That issues of knowledge production have received attention in recent years is testament to the willingness of social scientists to examine their own routines and credentials, and to consider possible reforms. It should be clear that the

authors of this volume take the goal of science seriously and do not view it as incompatible with the “naturalistic” goals associated with biology, chemistry, engineering and physics. Indeed, we draw on the experience of those fields wherever relevant.

Social science is in some ways different and in some ways the same as natural science. We do not feel the need to stake out a precise position in this perennial debate. The important point is that this book is written by social scientists and for social scientists, and does not purport to reflect upon the production of knowledge in the natural sciences except by way of an occasional comparison or contrast.

Social science, for present purposes, includes the core fields of economics, political science and sociology along with their many offshoots – business, management, communications, demography, education, environmental policy, international relations, law, social work and so forth. Fields like psychology and public health sit astride the social/natural science divide, while other fields like cultural anthropology and history straddle the social science/humanities divide. These areas lie on the periphery of our concerns.

So delimited, the social sciences share a common subject – understanding social behavior in a scientific fashion. As such, these fields have a great deal in common. They share concepts and approaches, and encounter similar methodological obstacles (Gerring 2012b). Increasingly, academics form partnerships that stretch across disciplines and publish in journals that are not restricted to any single discipline. Their topics intermingle. Accordingly, economists, political scientists and sociologists who study the same subject usually have more to say to each other than to their colleagues studying other subjects. Disciplinary boundaries are increasingly hazy – except insofar as they govern academic institutions (e.g., PhD programs, departments, conferences and journals). From a sociology-of-science perspective it matters greatly whether one’s field is economics, political science or sociology. From a methodological or substantive perspective, not so much.

That is why this book is formulated in an inclusive fashion, encompassing all the social sciences. Contributing authors hail from all three major social science fields. Although some chapters are centered on one or two of these disciplines, this is a matter of familiarity and logistics. It is hard to cover all disciplines with equal facility, especially within the confines of a single chapter. Authors naturally gravitate to what they know best, and this means that each chapter is likely to tilt toward the author’s home turf. In any case, none of the issues discussed in this volume are discipline-specific; they pertain broadly to all the social sciences.

Outline

Having defined the scope of this project, we turn to its content – the production of knowledge. We divide this subject into five areas, corresponding to the five sections of the book: (a) discovery, (b) publishing, (c) transparency and reproducibility, (d) appraisal and (e) diversity.

Discovery

While most work on social science methodology focuses on the task of appraisal, one must not lose sight of the importance of discovery. For social science to progress, researchers must push upon the frontiers of what is known. Innovative work, work that takes risks and conceptualizes topics in new ways, must therefore be valorized and rewarded – and, if possible, taught.

In Chapter 2, Richard Swedberg focuses on exploratory research, which he defines as research aimed at the discovery of something new and important. Swedberg argues that exploratory research is at the heart of all good research, yet exploratory research is risky (because it is often unsuccessful) and undervalued in the social sciences. He distinguishes different kinds of exploratory studies and provides guidelines for their effective use going forward. In Chapter 3, Evan Lieberman contrasts the place of exploratory work within the broader research cycle in political science versus biomedical research. Whereas exploratory studies in the biomedical sciences are valued because of their substantively important role, these studies are assigned a marginal position in political science's more truncated research cycle. Lieberman identifies the source of the problem with the disproportionate attention allocated to research focused on the precise estimation of causal effects in political science.

Publishing

Publishing in a top journal or press has long been regarded as a hallmark of professional success in the academy, and there is no sign that is likely to change in the near future. Consequently, journals and presses perform a central gatekeeping function, and rules governing access to these venues are critical to scientific progress within a discipline.

In Chapter 4, Tim F. Liao draws on his own experience as a journal editor to critically evaluate the peer review system of journal publication. He identifies several major problems with peer review as currently practiced, ranging

from unconstructive, slow and cranky reviews to systemic norms that privilege normal science over new developments and paradigmatic shifts. In response, Liao proposes six specific solutions, assessing the feasibility of each. In Chapter 5, John Gerring and Lee Cojocarú consider the consequences for social science of length limits – i.e., word or page limits – in scholarly journal publications. They summarize journal practices in political science and sociology, showing that length limits are pervasive though arbitrary (in partial contrast to economics). Gerring and Cojocarú argue that length limits bias research toward topics that can meet those limits and stand as an obstacle to knowledge production in the social sciences.

Transparency and Reproducibility

Transparency and reproducibility are core components of scientific progress. Without detailed information about the data and analysis employed in a study, it is impossible for other researchers to gauge its probable truth-value or to reproduce the result.

In Chapter 6, Garret Christensen and Edward Miguel frame the nature of the problem, presenting a simple model for estimating the likelihood that research findings are true under different assumptions. They then consider several specific issues: publication bias, specification searches and the inability to reproduce findings. Using empirical data, they show that these problems are pervasive in the social sciences across all disciplines. In Chapter 7, Christensen and Miguel turn to the question of solutions. They offer several plausible if partial solutions, including study registration, pre-analysis plans, improved statistical practices and better data sharing. They suggest these solutions provide grounds for optimism going forward. In Chapter 8, Kapiszewski and Karcher argue that the benefits that data sharing has the potential to bring will accrue more quickly if scholars make their research data *meaningfully* accessible – interpretable and analyzable by others, and shared via the increasingly sophisticated infrastructures being constructed for publishing and preserving research data. They also consider a series of steps that could and, they argue, should be taken in order to establish making research data accessible as a scholarly norm and encourage more social scientists to share their data. In Chapter 9, Alan M. Jacobs explores whether pre-registration and results-free review is a solution to a serious problem in the social sciences: published results often cannot be believed. In particular, Jacobs explores whether the practices of pre-registration and results-free review can inform observational research, especially qualitative

research in the social sciences. He shows why and how benefits of these practices often can be realized – and should be realized – in observation and qualitative research.

Appraisal

Scientific progress depends upon replicating the results of previous studies. Without such replications, there is no way to verify the truth-value of findings, or to extend theories to new areas (i.e., new populations). However, individual replications, by themselves, are unlikely to secure scientific progress. Findings must also be integrated into a larger body of work (e.g., meta-analysis) and a broader theoretical framework (synthesis).

In Chapter 10, Jeremy Freese and David Peterson use the crisis concerning replication in social psychology as a springboard for discussing the challenges of replication in quantitative research more generally. They note that replication itself can mean different things: the use of the same procedures, the use of the same data and/or the reproduction of findings. Freese and Peterson see issues of replication as inherently complex and raising tradeoffs for researchers that defy simplistic one-size-fits-all solutions. In Chapter 11, Dan Reiter focuses on the specific area of measurement replication, which he shows is crucial to knowledge accumulation in both qualitative and quantitative research. Reiter uses measurement replication as a lens to consider three types of errors common in the social sciences – errors in fact, errors in interpretation and context and errors in consistency of application. He concludes with concrete suggestions for improving measurement replication in the future. In Chapter 12, Tasha Fairfield and Andrew Charman consider analogs of replication for qualitative research grounded in Bayesian reasoning. They show how Bayesianism directs these researchers to explore whether previous scholars may have overstated the weight of evidence in support of the advocated argument by failing to assess how likely that evidence would be if a rival hypothesis were true. They focus on practices of appraisal that can help qualitative scholars improve inferences, building more consensus and better promoting knowledge accumulation.

In Chapter 13, John Gerring offers a specific proposal to address some of the problems raised by the replication crisis: coordinating reappraisals at an institutional level. He defines reappraisals broadly to include a large range of follow-up studies. His specific proposal includes the creation of a reappraisal institute, an accompanying infrastructure to help oversee the creation and collection reappraisals, and a scorecard system to keep track of findings and