

this volume comments little on the relationship between these techniques and theory. Do they allow us to better test existing theories? Or do they enable us to construct richer theories and ask big new questions?

Grimmer's chapter, an *excellent* introduction to topic models, exemplifies this lack of engagement with theory. While the data set used in this chapter contains 170,000 press releases from 2005 to 2010, the test of the model depends on a much smaller sample and much less sophisticated analysis where Grimmer looks at 20 "spikes" in three of the 44 granular topics recovered by his model, and whether or not some of them correspond to real-world events. He concludes that in 2005, the Republican members claimed credit for spending on various projects in 23.2% of press releases, and by 2010 this had dropped to 9%. During this time, Republicans also increased the frequency of their attacks on the Obama administration, particularly on the issue of Obama's health-care plan. BIG Data, new techniques, but nothing new or even that interesting theoretically.

Similarly, the chapter by Joshua Tucker et al., takes on a fascinating topic: the connection between social media use and protest movements in Turkey and Ukraine. Again, we have a huge data set of ~30 million tweets in the case of Turkey and ~11 million for Ukraine. But the goal of the chapter is not to establish a causal link between participation in protests and social media use or to look at how social media use changed these protests. Rather, when one boils down the standards these authors present to validate their model, the dual goal is to determine whether or not people tweet more when there is a protest and whether important or violent events that occur midprotest cause a spike in tweets. Unsurprisingly, people do tweet as expected, but the proof in both cases is whether three (for Turkey) or four (for Ukraine) events correspond to a subset of spikes in Twitter usage during each protest.

Tucker et al. avoid the more interesting question of how Twitter and Facebook impacted the dynamics of the protests. In passing at the end of the chapter, the authors note that they contacted 16 people involved in the protest and found that 14 of them heard about the protests on social media. Ironically, they apply BCD to small questions and small data to the big question. We intend this not as a criticism but as a comment on the opportunity before us. The granularity of BCD may well allow for testing theories of the role of social media in political uprisings. For that to occur, we need more interaction between empiricists and theorists.

We believe that empirically minded researchers should read, teach, and engage this volume. We applaud the editor and authors for their creative use of previously unexplored sources of data and for mapping out the edges of the frontier of a new social science. We also encourage more theoretical-minded researchers to contemplate how new data connect to long-standing theoretical questions

in social science about why people participate in politics, why nation-states engage in conflict, or how we might maintain the commons. Alvarez and the other authors are reaching out to you, the broader community. We hope you respond in kind. Without communication between empiricists and theorists, the revolution of BIG data could lack significance.

Process Tracing: From Metaphor to Analytic Tool.

By Andrew Bennett and Jeffrey T. Checkel. New York: Cambridge University Press, 2014. 344p. \$99.00 cloth, \$36.99 paper.
doi:10.1017/S1537592716003339

— Alexander Lee, *University of Rochester*

Social scientists spend much of their time making statements about cause and effect, and developing complex theories of causal relationships. The most basic way to test such theories is comparison of cases, whether a small number of case studies (in a qualitative setting) or a larger number of observations (in a quantitative setting). Practitioners of comparison techniques have tended to discourage making causal inferences within single cases (Gary King, Robert O. Keohane, and Sidney Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research*, 1994), due to problems of generalizability. However, there are many instances where a single case study may be the only viable research design.

In dealing with problems of causal inference in single cases settings, the "central" (p. 4) technique in political science is process tracing, a term first developed in cognitive psychology and appropriated by Alexander George (Alexander L. George, "Case Studies and Theory Development: The Method of Structured, Focused Comparison" in Paul Gordon Lauren, ed., *Diplomatic History: New Approaches*, 1979). Where multi-case studies attempt to infer causation from the correspondence of cause and effect, process tracing seeks to use the mechanism itself as evidence, to examine "whether the causal process a theory implies is in fact evident in the sequence and values of the intervening variables" (Alexander L. George, and Andrew Bennett, *Case Studies and Theory Development in the Social Sciences*, 2005, p. 6). Quite commonly, this involves examining the statements of decision makers involved in the process, but the technique can potentially be applied to the actions of individuals and groups as well.

As process tracing has become more popular in recent years, it has suffered something of a "buzzword problem," with the term being promiscuously applied to a wide variety of qualitative techniques with little link to the original idea. Moreover, much of the literature on the topic has been polemical in tone, advocating process tracing's efficacy relative to other techniques, particularly quantitative ones, rather than distinguishing good and bad examples and providing advice on techniques.

Bennett and Checkel's *Process Tracing: From Metaphor to Analytic Tool* seeks to address these shortcomings in the literature, and to provide a roadmap for scholars to use in designing their own research. The introduction sets out the editors' definition of process tracing and suggests a set of best practices for its implementation. The next six chapters examine the application of process tracing in specific literatures, including European integration, international institutions, ideological effects, civil war, and the end of the Cold War. To varying degrees, each of these contributions applies the best practices of the introduction, showing how existing work does, or occasionally does not, use these techniques. In separate chapters, Thad Dunning develops a set of ideas on how process tracing can be applied to substantiating the assumptions of quantitative natural experiments, while Vincent Pouliot articulates a non-positivistic conception of process tracing quite different from that in the rest of the volume.

The volume represents a major advance on the current process-tracing literature in emphasizing practice over theory, and provided a clear and lucid introduction to the technique. However, the book falls somewhat short of being a "how-to" manual that an applied scholar could implement easily, since in many cases it appears that existing practice in process tracing falls short of the rigorous techniques advocated by the editors. Areas in this gap especially noticeable include the relationships between process tracing and evidence gathering, evidence presentation, theory building, and other analytical techniques.

Evidence Gathering. Successful process tracing, or indeed most types of research, requires three steps:

1. the gathering of evidence,
2. the use of this evidence to access the truth of hypotheses, and
3. the presentation of this evidence in a way calculated to convince readers of the truth of the hypothesis.

Bennett and Checkel focus heavily of the second of these questions. However, there is some reason to believe that the first and third also are significant obstacles to the success of the technique.

In most existing process-tracing projects, the evidence-gathering process involves either a series of field interviews, the use of written primary sources (whether archival or non-archival), or (less admirably) the perusal of secondary sources. All of these techniques take up a large amount of a researcher's time, and inevitably choices must be made: what shall be read, and who shall be interviewed? And how shall their evidence be assessed? Bennett and Checkel's advice on this question, while unobjectionable, is vague: Scholars must be "relentless" in gathering evidence, but then make a "justifiable decision on when to stop" (p. 27), and "consider the potential evidentiary biases of sources"(p. 24).

This lacuna is especially problematic by two factors: Firstly, American graduate schools in political science often do not teach either archival or interview methods, with coursework emphasizing either quantitative techniques or research design. Students are often sent to the field to "figure it out."

Such figuring may be especially dangerous in the context of a process-tracing design. Since process tracing (in Bennett and Checkel's view) involves testing the implications of a theory, students may understandably focus on gathering information that supports their hypothesis, rather than that which disproves it. A scholar with such a biased source base may be tempted to ignore Bennett and Checkel's cautions on interpreting "absence of evidence as evidence of absence" and report that no evidence contradicts her view, even when such evidence exists. In such circumstances, process tracing might devolve into the type of "just so stories" that many of the contributors disparage.

Archival sources present these problems in a particularly dangerous form. Not only is the material in archives shaped by the biases and agendas of their creators, but the material available in archives is often not a representative sample of the material originally created, shaped by the selective deletions by agencies and governments. These deletions are not always obvious to scholars, but may well be associated with particular categories of documents, and thus with the evidence to support or disprove particular hypotheses.

To restate this objection, while Bennett and Checkel's best practices may be appropriate, there is reason to be skeptical of the ability of political science as a discipline to implement them at the present time. Until political scientists become as sophisticated consumers and producers of archival information as historians, and as sophisticated consumers of interviews as anthropologists, using such information to make causal inferences may present difficulties.

Presentation: After making a determination about the correctness of the hypothesis, a scholar (of any methodology) must then convince the reader of the correctness of his conclusions by the presentation of evidence. This is complicated by the fact that scholars may have just as much reason to obfuscate their presentation as their sources do to obfuscate their motivations. Having invested in developing an original hypothesis and working out its implications, a scholar may be unwilling to find that an alternative hypothesis was right all along. It is notable that none of the process-tracing works discussed in this book come to such a finding.

Given this incentive, how can users of process tracing convince their readers that the conclusions they come to are not spurious, and that the evidence they present is not cherry-picked? Like other qualitative scholars, they may be inhibited by the length or confidentiality of their

evidence from presenting it in full, a problem that has been at the crux of the recent debate on reproducibility in qualitative research.

The contributors to this volume provide little insight on this question. While in many cases the evidence in support of a particular argument is described as “rich,” there is little discussion of how this richness may be conveyed. There is also little discussion on how a scholar may appropriately convey exhaustion—the “I’ve looked at hundreds of pages of testimony and have never come across a discussion of a mechanism” that may be necessary to disprove implications of alternative hypotheses.

Discussion of evidence presentation is most explicit in Thad Dunning’s chapter on the use of process tracing to substantiate the assumptions of natural experiments. The credibility problem in the presentation of process-tracing results is especially severe in this application, and the need to present evidence to support the conclusions of the process-tracing analysis is thus especially keen. The rewards of substantiating the exogeneity of the assignment process, and being able to implement a well-identified natural experiment, are often quite high relative to the rewards of finding that assignment is a product of endogenous social processes. Dunning’s suggestions for increased transparency and public availability of qualitative evidence would surely be helpful in this regard, but it is doubtful whether they will fully solve this problem, given the amount of supporting material that these projects frequently produce, and the level of contextual information necessary to interpret it correctly.

Theory and Process Tracing. At the heart of Bennett and Checkel’s approach to process tracing is the articulation of clear theories. Scholars must then work out the observable implications of these theories, and search for evidence confirming or disproving them. Some of the contributors conceptualize (and diagram) these implications as nodes or “observations” of a causal process, each leading to the others through a logical progression. A theory confirmed by such an analysis must then be compared to alternative theories, which must be subjected to an equally rigorous analysis of the evidence for their observable implications.

Several points might be made about such an approach. Firstly, the leap from accessing the truth of specific observable implications to accessing the truth of the overall hypothesis is often a large one. Bennett and Checkel (in the appendix) recommend a Bayesian framework for reassessing hypotheses in light of evidence, though, as they recognize, in practice it can be quite difficult to know the various prior probabilities necessary to calculate the posterior probability, particularly in the unique cases on which process tracing often focuses. This means that in many cases the assessment of probability must begin based on “subjective guesses” (p. 280), rather than the known probabilities used in teaching examples.

Despite this problem, Bennett and Checkel argue that Bayes law is valuable for this purpose, and even suggest that scholars make explicit numerical estimates of the priors. However, no scholar, either in this volume or elsewhere, has ever implemented this suggestion (p. 298), preferring a more ad hoc way of accessing Bayesian posteriors. Indeed, the choice is an unenviable one: Either make hard estimates based on weak or partial data (all exposed to the criticism of colleagues) or obfuscate the logical process by which hypotheses are accessed. Bennett and Checkel’s advocacy for the first option may well be correct, but in the absence of successful examples it may be for some time unheeded.

One of the claimed strengths of process tracing is its openness to inductive learning, allowing scholars to reassess probabilities in light of evidence (pp. 29–30). However, Bennett and Checkel do not make clear how this inductive process should be combined with the type of rigorous testing of observable implications that they advocate. If the theory generation process is fully inductive, the scholar might fall into the type of behavior that in a quantitative application would be called data mining—developing ex-post logic to justify observed empirical patterns. While such a theory might find all its observable implications confirmed, this association might well be spurious or non-generalizable.

The tension between inductive and deductive theory building is an old one in the philosophy of science, but the close interplay between theory and empirics in process tracing make the problem a knotty one. While the authors’ appear aware of the problem (p. 22), more advice is necessary on how to balance the need for tight tests of theories with the “surprise” and richness often attributed to process-tracing techniques (p. 30).

The Boundaries of Process Tracing. Bennett and Checkel are very clear about the definitions of the technique, and what does not “count” as process tracing. In particular, causal inference through comparisons of cases, either within or across time, is not process tracing. However, many of the contributors to this volume, and many of the works they discuss, use the term much more loosely to describe a wide variety of qualitative research techniques. Alan Jacobs’ discussion of Sheri Berman’s book, for instance, describes the strategic rigidity of the German SPD over time, despite changing electoral incentives, as evidence for an ideational theory of party positioning. While this may well be correct, it represents an over-time comparison rather than process-tracing proper.

A similar confusion is apparent with reference to the relationship between process tracing and case study evidence. While Bennett and Checkel strongly advocate for the use of cases studies to supplement process tracing, the mechanics could use fleshing out. Is process tracing simply a sophisticated technique for understanding cases, which will then be compared with each other using standard techniques? Or will case studies somehow be

integrated into the Bayesian assessment of the original hypothesis? Individual chapters often obscure this question, describing particular case analyses within texts but not the larger framework within which they are embedded.

Qualitative-Quantitative Conflict: Discussion of process tracing has frequently tended to coincide with broader controversies in the discipline on the value of qualitative methods, and of “rational choice” theoretical perspectives. This has led many scholars to conclude that process tracing is not only incompatible with quantitative methods but inseparable from theories emphasizing the importance of ideology and cognitive factors.

One excellent feature of this volume is that the contributors reject both stereotypes, and treat process tracing as a technique suitable for testing a wide range of theoretical perspectives in coordination with a wide range of additional techniques. Indeed, the review chapters illustrate that process tracing can be used to test a wide range of hypotheses. It is notable, however, that ideological and cognitive explanations bulk larger in this volume than in the discipline as a whole, as the collections of personal testimonies necessary to substantiate such explanations are also important evidence—and often the only evidence—of how political processes unfold.

Conclusion: The contributors to this volume have made a notable contribution to the evolution of process tracing into the standard technique for the evaluation of hypotheses in single case studies. The techniques that they advocate and discuss represent an improvement on the more theoretical discussions common in the qualitative methods literature, and will make this book a valuable tool in graduate teaching. However, more remains to be done. Until the applied tools necessary to process tracing—archival analysis, qualitative reliability—are fully worked out, process tracing will tend to remain a vague label or a magnet for epistemological debate, rather than an ordinary research tool.

**Explanation and Progress in Security Studies:
Bridging Theoretical Divides in International Relations.**

By Fred Chernoff. Stanford: Stanford University Press, 2014. 328p.

\$90.00 cloth, \$27.95 paper.

doi:10.1017/S1537592716003340

— John M. Owen IV, *University of Virginia*

Raise topics of philosophy of science and meta-theory around scholars of international security, and most will begin checking their phones. It is fair to say that a majority of the security studies guild finds questions about research programs and epistemology irrelevant and a distraction. That is a pity. As Fred Chernoff makes clear in this important book, philosophy of science has a great deal to say about a question that all students of security—all social scientists—should care about: Are we making progress? Can we explain more than previous generations could?

The question has arisen before, of course, and most IR scholars have some familiarity with what the works of Thomas Kuhn and Imre Lakatos imply about their field. But Chernoff takes a fresh approach to the question of progress by conducting exhaustive surveys of three important literatures—on nuclear proliferation, alliances, and democratic peace—and teasing out a different kind of answer.

Chernoff reformulates the question, asking: Why are so many decades-old debates in security studies unresolved? The debates he has in mind are not over paradigms. Chernoff has no interest in thawing the frozen conflicts between realism and liberalism, or rationalism and constructivism. He has in mind, rather, the intractability of so many empirical puzzles. After nearly fifty years of study, for example, why can we not agree on why states form one set of alliances rather than another, or ally at all?

In three chapters that could serve as introductions to graduate students on diversity in IR, he selects ten or eleven highly cited authors in each of the literatures, summarizes the arguments and methods of each, points out their similarities and differences, and offers some critiques of individual works. Some of the works are quantitative, others qualitative; some are realist, others liberal; some are rationalist, others constructivist.

The chapter on nuclear proliferation, for example, surveys ten amply cited works of scholarship, starting with George Quester's 1973 book *The Politics of Nuclear Proliferation*, running through work by Kenneth Waltz, William Potter, Stephen Meyer, Etel Solingen, Scott Sagan, T. V. Paul, Sonali Singh and Christopher Way, Jacques Hymans, and Dong-Joon Jo and Erik Gartzke. Chernoff finds wide disagreement across these authors over how best to explain proliferation, with some holding to a strategic hypothesis about external threats, and others arguing for bureaucratic variables, domestic politics, national identity, or the characteristics of decision makers. Policy makers could use closure on this academic debate, but none is forthcoming.

One might think that empirical impasses such as this are caused by prior theoretical disagreements, especially between realists and non-realists. Chernoff certainly recognizes the importance of theory, and in his final chapter thoughtfully considers the possibility that scholarly stalemates are produced by lack of agreement over the precise question being asked, which in turn could be tied to theoretical diversity.

His own answer, however, looks to philosophy of science. Where security scholars cannot agree on the best explanation of a phenomenon, it is in large part because they do not agree on what constitutes a good explanation to begin with. Few security scholars explicate criteria for an adequate causal argument. (This seems to irritate Chernoff, but surely economists, geologists, physicists, and others in more progressive sciences are similarly