# Part I

## Introduction

# 1    Process tracing
## From philosophical roots to best practices

Andrew Bennett and Jeffrey T. Checkel

## Introduction

Why did the Cold War end peacefully, without a shot being fired? Why did some European democracies survive the interwar period while others were replaced by fascist dictatorships? In the post-Cold War world, civil conflicts have replaced interstate war as the dominant form of organized political violence, with rebel groups – instead of intercontinental ballistic missiles (ICBMs) – as a key focus of both policy and scholarship. Yet what makes such groups tick? Why do some engage in wanton killing and sexual violence while others do not? The European Union is a unique experiment in governance "beyond the nation state," but how are its supranational governance structures being crafted and with what effect on the ordinary citizens of Europe?

Contemporary political science has converged on the view that these puzzles, and many more on the scholarly and policy agendas, demand answers that combine social and institutional structure and context with individual agency and decision-making. This view, together with recent developments in the philosophy of science, has led to an increasing emphasis on causal explanation via reference to hypothesized causal mechanisms. Yet this development begs the questions of how to define such mechanisms, how to measure them in action, and how to test competing explanations that invoke different mechanisms.

This book argues that techniques falling under the label of process tracing are particularly well suited for measuring and testing hypothesized causal

mechanisms. Indeed, a growing number of political scientists now invoke the term. Despite or perhaps because of this fact, a buzzword problem has arisen, where process tracing is mentioned, but often with little thought or explication of how it works in practice. As one sharp observer has noted, proponents of qualitative methods draw upon various debates – over mechanisms and causation, say – to argue that process tracing is necessary and good. Yet, they have done much less work to articulate the criteria for determining whether a particular piece of research counts as good process tracing (Waldner 2012: 65–68). Put differently, "there is substantial distance between the broad claim that 'process tracing is good' and the precise claim 'this is an instance of good process tracing'" (Waldner 2011: 7).

This volume addresses such concerns, and does so along several dimensions. Meta-theoretically, it establishes a philosophical basis for process tracing – one that captures mainstream uses while simultaneously being open to applications by interpretive scholars. Conceptually, contributors explore the relation of process tracing to mechanism-based understandings of causation. Most importantly, we articulate best practices for individual process-tracing accounts – for example, criteria for how micro to go and how to deal with the problem of equifinality (the possibility that there may be multiple pathways leading to the same outcome).

Ours is an applied methods book – and not a standard methodology text – where the aim is to show how process tracing works in practice. If Van Evera (1997), George and Bennett (2005), Gerring (2007a), and Rohlfing (2012) set the state of the art for case studies, then our volume is a logical follow-on, providing clear guidance for what is perhaps the central within-case method – process tracing.

Despite all the recent attention, process tracing – or the use of evidence from within a case to make inferences about causal explanations of that case – has in fact been around for thousands of years. Related forms of analysis date back to the Greek historian Thucydides and perhaps even to the origins of human language and society. It is nearly impossible to avoid historical explanations and causal inferences from historical cases in any purposive human discourse or activity.

Although social science methodologists have debated and elaborated on formal approaches to inference such as statistical analysis for over a hundred years, they have only recently coined the term "process tracing" or attempted to explicate its procedures in a systematic way. Perhaps this is because drawing causal inferences from historical cases is a more intuitive practice than statistical analysis and one that individuals carry out in their everyday lives.

Yet, the seemingly intuitive nature of process tracing obscures that its unsystematic use is fraught with potential inferential errors; it is thus important to utilize rigorous methodological safeguards to reduce such risks.

The goal of this book is therefore to explain the philosophical foundations, specific techniques, common evidentiary sources, and best practices of process tracing to reduce the risks of making inferential errors in the analysis of historical and contemporary cases. This introductory chapter first defines process tracing and discusses its foundations in the philosophy of social science. We then address its techniques and evidentiary sources, and advance ten best-practice criteria for judging the quality of process tracing in empirical research. The chapter concludes with an analysis of the methodological issues specific to process tracing on general categories of theories, including structural-institutional, cognitive-psychological, and sociological. Subsequent chapters take up this last issue in greater detail and assess the contributions of process tracing in particular research programs or bodies of theory.

## Defining process tracing

The term "process tracing" originated in the field of cognitive psychology in the United States in the late 1960s or early 1970s.[1] As used in psychology, process tracing refers to techniques for examining the intermediate steps in cognitive mental processes to understand better the heuristics through which humans make decisions. In 1979, the Stanford University political scientist Alexander L. George appropriated the term to describe the use of evidence from within case studies to make inferences about historical explanations (George 1979).

Because much of George's own research was in political psychology, and because the term "process tracing" originated in cognitive psychology, it has sometimes been viewed as applying mostly or only to analyses of individual level decision-making. Although process tracing does apply well to the individual level and cognitive theories (see Jacobs, this volume, Chapter 2), George made clear in subsequent writings that it can also be used to make inferences on structural or macro-level explanations (George and Bennett 2005: 142, 214; see also Waldner, this volume, Chapter 5). For example, many economic theories hypothesize relationships and sequences among macroeconomic

---

[1] The very first usage of the term remains unclear; the earliest relevant citation on Google Scholar is Hobarth 1972, a Ph.D. thesis at the University of Chicago.

variables that can be tested through process tracing at the macro level as well as that at the micro or individual level.

Similarly, because of its origins in cognitive psychology and because many of its early practitioners in that field went on to explore the errors that individuals make and the biases they exhibit in their decision-making, process tracing is sometimes viewed as incompatible with rational choice theories. We concur, however, with the many prominent rational choice theorists who argue that their hypotheses should bear some correspondence with the actual processes through which individuals make decisions, and that they should therefore be amenable to process tracing (Bates *et al.* 1998; see also Schimmelfennig, this volume, Chapter 4).

The essential meaning retained by the term "process tracing" from its origins in cognitive psychology is that it refers to the examination of intermediate steps in a process to make inferences about hypotheses on how that process took place and whether and how it generated the outcome of interest. In previous work together with George, one of us defined process tracing as the use of "histories, archival documents, interview transcripts, and other sources to see whether the causal process a theory hypothesizes or implies in a case is in fact evident in the sequence and values of the intervening variables in that case" (George and Bennett 2005: 6). We added that "the process-tracing method attempts to identify the intervening causal process – the causal chain and causal mechanism – between an independent variable (or variables) and the outcome of the dependent variable" (ibid.: 206).

The authors then used a metaphor to expand on this definition. If one had a row of fifty dominoes lying on the table after they had previously been standing, how could one make inferences about whether the first domino caused the last to fall through a domino process, or whether wind, a bump of the table, or some other force caused the dominoes to fall? The answer, George and Bennett argued, was to use evidence on the intervening processes posited by each of the alternative explanations. Did anyone hear a succession of dominoes? Do the positions of the fallen dominoes shed light on how they fell? And so on.

While we feel this definition is still an excellent starting point, it is necessary to point out a weakness in both it and the accompanying metaphor. The term "intervening variable" opens the door for potential confusion because social scientists are accustomed to thinking of variables as either causal (independent) or caused (dependent). However, both the term and the metaphor of dominoes falling suggest that an intervening variable is both fully caused by the independent variable(s) that preceded it, and that it transmits this causal

force, without adding to it, subtracting from it, or altering it, to subsequent intervening variables and ultimately through them to the dependent variable.

When the observable events that intercede between hypothesized causes and observed effects have this character, they constitute "diagnostic evidence," not "variables." Diagnostic evidence indicates the kind of process taking place, but does not transmit any independent effects to the dependent variable. This is analogous to a diagnostic medical test, such as a dye injected into a patient to enhance a CAT scan of blood flow. Ideally, the dye does not transmit any noteworthy side-effects to the patient, but it provides evidence on the processes taking place in the patient. Similarly, in social and political life, the ways in which actors privately frame or explain their actions may provide diagnostic evidence on their motives without independently affecting the outcomes of interest.

Quite often, however, the events that lie temporally and/or spatially between the independent variable and the dependent variable are *not* fully determined by the specified independent variables and these events *do* have independent effects on the nature, timing, or magnitude of the dependent variable. In such instances, researchers have to make theory-building choices. Are they going to model these intervening events as variables? If so, should they treat them as exogenous, complementary, or endogenous to the theory or explanation of interest? Exogenous variables are those excluded from a model because they are either not powerful or frequent enough, or too complex or unpredictable, to be brought into the theory. Complementary variables are those that add to or subtract from the effects of the main variables of interest, but do so independently, or without interaction effects related to the main variables. When such variables are sufficiently simple to be theorized, they can be added to a model without changing the main variables or mechanisms of interest. Additional variables that interact with the independent variables of interest in more complex ways need to be either brought into the model (endogenized) or identified but set aside from the model (exogenized) for the sake of simplicity. Methodologically, whatever way additional variables are brought into or set aside from the theory that aims to explain the case, this modification can be tested through additional process tracing.

We thus drop the term "intervening variable" and define process tracing as the analysis of evidence on processes, sequences, and conjunctures of events within a case for the purposes of either developing or testing hypotheses about causal mechanisms that might causally explain the case. Put another way, the deductive theory-testing side of process tracing examines the observable implications of hypothesized causal mechanisms within a case to test whether a

theory on these mechanisms explains the case (Schimmelfennig, this volume, Chapter 4, emphasizes such a procedure). The inductive, theory development side of process tracing uses evidence from within a case to develop hypotheses that might explain the case; the latter hypotheses may, in turn, generate additional testable implications in the case or in other cases (Pouliot, this volume, Chapter 9, stresses inductive research procedures).[2]

It is important as well to define "case" and "within a case" as we use them. Following George and Bennett, we define a case as "an instance of a class of events" (George and Bennett 2005: 17). This definition recognizes that classes of events – revolutions, democracies, capitalist economies, wars, and so on – are the social constructions of both political actors and the social scientists who study and define political categories. They are not simply given to us by history, but defined by our concepts, and much contestation in interpreting the results of case-study research concerns disagreements over which "cases" should or should not be included in a defined population.

We define within-case evidence as evidence from within the temporal, spatial, or topical domain defined as a case. This can include a great deal of evidence on contextual or background factors that influence how we measure and interpret the variables within a case. Henry Brady and David Collier provide a useful distinction here between data-set observations and causal-process observations (see also Dunning, this volume, Chapter 8). Data-set observations are "an array of scores on specific variables for a designated sample of cases," and these observations provide the basis for statistical analyses. Causal-process observations are "observations on context, process, or mechanism" and are used in within-case analyses such as process tracing (Brady and Collier 2010: 12).

With these definitions in hand, we note that process tracing is closely related to historical explanation, as that term is used by the historian Clayton Roberts. In Roberts's view, an historical explanation is not simply a detailed description of a sequence of events; rather, it draws on theories to explain each important step that contributes to causing the outcome. Roberts distinguishes between macro-correlation and micro-correlation, the latter of which is quite similar to process tracing. Macro-correlation involves an attempt to explain historical cases at a high level of generality through universalistic theories, similar to Hempel's notion of theories as covering laws.

---

[2] Beach and Pedersen 2013a suggest three different types of process tracing: theory testing, theory building, and outcome explaining. The first is primarily deductive, the second more inductive, and the third uses both kinds of logic with the goal of causally explaining an individual case.

Roberts argues that historical events are too complex to fit easily under exception-less covering laws, and efforts to explain history in this way "have met with little success" (Roberts 1996: 15). He urges instead that researchers should use micro-correlation, which involves "the minute tracing of the explanatory narrative to the point where the events to be explained are microscopic and the covering laws correspondingly more certain" (ibid.: 66).

One difference between Roberts's approach to process tracing and our own is that Roberts felt that – at the micro-correlational level – the theories underlying an historical explanation would be "platitudinous." Historians, he thus argues, rarely reference them explicitly because to do so would "hopelessly clog the narrative" (ibid.: 66–67, 87–88). We emphasize instead the importance of making explicit the hypotheses about underlying causal mechanisms that are theorized to have caused an outcome so that these can be rigorously assessed, even if this results in political science narratives that are more clogged – and alas, less likely to become best-sellers – than those of historians (see also Evangelista, this volume, Chapter 6, for analysis of works that focus their process tracing as much on explaining an important historical case as on developing and testing general theories).

Yet, these disciplinary differences need not be viewed in zero-sum terms. That is, it is possible to have an application of process tracing that is simultaneously rigorous, explicit, and transparent, and that also reads well – say, by placing the process tracing tests in an appendix separate from the main narrative (Fairfield 2013 provides an excellent example).

Our concept of process tracing differs even more sharply with time series cross-sectional analysis, which involves the correlational study of data across a variety of units (often, annual data across a range of countries). Although this form of analysis might be confused with process tracing because it involves temporal data from within cases over time, it is still a form of cross-case and correlational inference, rather than the study of hypothesized processes within individual cases, and it is thus fundamentally different from process tracing (see also the discussions and examples in Lyall, Chapter 7; and Dunning, Chapter 8, both this volume).

In sum, process tracing is a key technique for capturing causal mechanisms in action. It is not simply glorified historiography, nor does it proceed by the logic of frequentist statistics. And – as we argue below – there are metrics and best practices that allow one to distinguish good process tracing from bad. However, since standards flow from underlying philosophical positions, it is important first to clarify the meta-theory of process tracing.

## Philosophy of social science and process tracing

On a philosophical and epistemological level, process tracing is closely related to the turn toward social science explanations based on reference to causal mechanisms (Elster 1998; Gerring 2007b; Mayntz 2004), or the underlying entities that generate observed processes and outcomes. Much of the thinking about causality and causal explanation over the last 200 years has been strongly influenced by David Hume's argument that constant conjunction – the frequent conjoint occurrence of variables A and B – is the essence of causal inference. More recent work by pragmatist (Johnson 2006) and scientific realist (Wight 2006) philosophers of science, however, provides a meta-theoretical foundation more amenable to thinking in terms of mechanisms. Indeed, for these scholars, a causal explanation is built around contiguity and sequencing of events – concepts that Hume mentioned, but gave insufficient attention. These open a methodological space for process tracing.

One difficulty in making use of contemporary discussions in the philosophy of science is that there are at least a half-dozen variants of scientific realism (Chernoff 2002) and even more different definitions of causal mechanisms (Mahoney 2001; see also Hedström and Ylikoski 2010). While a full discussion of scientific realism is beyond our present purposes, we concur with the emphasis it places on causal processes and causal mechanisms as central elements of causal explanation.

More important for this volume is the task of sorting out the competing definitions of causal mechanisms. These divide along three fundamental issues: (1) Are causal mechanisms in some sense unobservable? (2) Does explanation via reference to causal mechanisms involve a commitment to methodological individualism, or beyond that, to explaining human behavior by neuroscience and ultimately by sub-atomic physics? (3) Are causal mechanisms sufficient to explain outcomes in specified circumstances or contexts, or might mechanisms be inherently probabilistic or stochastic?

On the first issue, most discussions of mechanisms place them on the ontological level. This means we make hypotheses or theories about how such ontological entities as mechanisms might work, and we test the observable implications of these hypotheses, but we do not observe causal mechanisms directly. Some proponents of mechanisms take a different view, arguing that they are at least somewhat observable. Hedström and Ylikoski, for example, critique Mahoney for the view that mechanisms are unobservable,