CHAPTER 5

Causal Inference and Process-Tracing Methods

Social scientists use theories in the same manner that we use maps—to simplify an immensely complex reality. Yet whereas cartographers engage in descriptive inference when they make maps, social scientists are also interested in going a step beyond describing what happens to study causal relationships, explaining how and why social phenomena occur. Theories of causal mechanisms are, for example, simplifications of reality that predict what causal forces are important in explaining the occurrence of a phenomenon. Yet social scientists are not interested in theories as purely hypothetical thought experiments. Theories need to be empirically validated to make sure that they accurately represent reality.

This chapter deals with the interrelationship between theory and empirical reality. After a theory is developed, we need to empirically validate its accuracy. We do so by testing whether empirical evidence strengthens or weakens our confidence in the validity of the theory as an explanation of a phenomenon.

Given the difficulty of measuring complex social phenomena such as democracy, we cannot fully measure what is happening in reality. Even with the best measuring instruments, we can gain only a selected sample of observations of a given phenomenon. Therefore, we are forced to infer from a small set of empirical observations that a theory was the cause of the phenomenon. In other words, we make an inferential leap from what we can observe empirically to conclude that an underlying causal explanation exists.

What enables us to jump from a set of empirical observations to infer whether a causal mechanism exists in a specific case in process-tracing research? Is the logic that enables inferences to be made the same in process-tracing as the frequentist logic of inference that underlies the methodological prescriptions for qualitative case study research as suggested by King, Keohane, and Verba (KKV)(1994)?

In process-tracing case studies, we attempt to analyze whether a theorized causal mechanism exists in an individual case. Therefore, we are interested in making what can be termed *within-case inferences,* meaning that we use empirical evidence collected from a particular case to infer that all of the parts of a hypothesized causal mechanism were actually present in that case. Process-tracing methods cannot be used to make cross-case inferences that involve concluding based on evidence drawn from a sample of comparable cases that a causal relationship exists across a population of a given theoretical phenomenon. Other forms of inferential tools, such as comparative methods, are necessary to make cross-case inferences.

In this chapter, we discuss why mainstream inferential tools used, for example, in classical statistical analysis and comparative cross-case methods cannot be used to make within-case inferences about causal mechanisms. Here, we continue the argument that our methodology must be brought in line with ontology (Hall 2003). In particular, we illustrate that given that the ontology of causality used in process-tracing differs from that of both classical statistical analysis and comparative methods, the inferential tools used in these methods do not apply in process-tracing. This discussion is followed by the presentation of the Bayesian logic of inference as a solution to the problem of making within-case inferences in process-tracing. The chapter concludes by discussing in more detail the types of inferences that can be made using process-tracing methods and, equally important, what types of inferences cannot be made.

5.1. The Type of Inferences Made in Process-Tracing

Process-tracing methods are used when we want to investigate whether causal mechanisms are present or absent in an in-depth case study (theory-testing), when we want to build a theorized causal mechanism based on an in-depth examination of the evidence in a single case (theory-building), or when we want to explain a particular outcome, such as why Nazi Germany was not effectively deterred from aggression against Poland by threats from France and the United Kingdom (explaining-outcome).

In all three variants of process-tracing, we are investigating a causal mechanism (or mechanisms) in what is effectively a single-case study. An important distinction across the variants is whether the ambition is to generalize beyond the single case (theory-centric) or is more focused on accounting for the outcome in the single case (case-centric). In a theory-centric single-case study, the researcher is interested in studying a delimited phenomenon such as the occurrence of war and making generalizations from the single case to the population level. The single case of war is treated as a selected example of the broader population of the phenomenon. Both theory-centric variants employ cases studies in this manner, although inferences within the single case about the presence/absence of a causal mechanism are made using process-tracing methods whereas inferences beyond the single case are not made using these methods. Instead, cross-case inferences are based on comparative methods or frequentist methods (see chapter 8). Cross-case inferences to a broader population are, for example, made possible using comparative cross-case methods such as fuzzy-set QCA that allow us to detect least-likely cases. Here, the logic would be that if evidence of a causal mechanism is found in a least-likely case, we can expect to find the mechanism elsewhere in the population.

Case-centric ambitions are prevalent in explaining-outcome process-tracing studies, where the scholar is interested in accounting for the key aspects of particularly important and puzzling outcomes, such as why the Soviet Union peacefully relinquished hegemonic control of Eastern Europe in 1989–90 or why EU governments decided to adopt the Euro in the Treaty of Maastricht. Indeed, many qualitative-oriented scholars contend that outcomes such as the end of Soviet domination of Eastern Europe have so many unique features that they are noncomparable with other cases of the broad phenomenon, such as the collapse of the Habsburg Empire in the aftermath of World War I (see also Collier and Mahoney 1996). Yet all three variants of process-tracing share the ambition of making within-case inferences.

Before we turn to the elaboration of why process-tracing methods build on a different inferential logic than other social science methods, it is instructive to develop an example of what a typical causal mechanism looks like—in particular, the types of evidence that must be gathered to enable an inference that all of the parts of a hypothesized causal mechanism are present in a particular case.

Owen (1994, 1997) has developed a five-part causal mechanism that explains how democracy produces peace.[1] Owen identifies the entities and activities that make up each part of the hypothesized causal mechanism that transmits causal forces from mutual democracy (X) to produce peace (outcome Y). Owen (1994) tests whether these phenomena are present in four case studies, but given that the evidence gathered in each case study is not comparable across the cases, these are in effect four parallel single-case process-tracing studies. Table 5.1 depicts the causal mechanism, with entities underlined and activities in italics.

Part 5 of his causal mechanism best illustrates the transmission of causal forces from X to Y that we want to study empirically in process-tracing analysis. Part 5 says that we should expect to see liberal elites agitate on behalf of their policies during war-threatening crises. The entities are liberal elites (n_5), and their activities (→) that transmit causal forces are their agitation against war with other liberal democratic states.

In his case study of the Franco-American crisis of 1796–98, Owen investigates whether available evidence supports the idea that liberal elites were agitating, with the theory being operationalized in a case-specific manner that directs his attention toward the actions of liberal elites in the news media and other public forums. Observations include speeches by a key liberal actor (Republican vice president Thomas Jefferson) along with the views of prominent liberal newspapers (Owen 1997: 86–87). After these observations are assessed for their accuracy and interpreted in their context, Owen uses them as evidence to infer that the part of the causal mechanism was present in the case.

Two crucial differences between process-tracing case studies and other social science methods can be seen here. The first relates to the types of material being collected; the second is in the analytical aims of causal inference.

TABLE 5.1. The Five Parts of Owen's Causal Mechanism whereby Democracy Produces Peace

Part of the Mechanism	Conceptualization of Mechanism and Its Parts (entities and activities)
Context	Crisis between states that can result in war
Independent variable (X)	Pair of states where analyzed state is democratic and where opponent is either democratic (liberal) or autocratic (illiberal) state
Part 1 (n_1 →)	<u>Liberals</u> will *trust* states they consider liberal and *mistrust* those they *consider* illiberal
Part 2 (n_2 →)	When <u>liberals</u> *observe* a foreign state becoming liberal democratic by their own standards, they will *expect* pacific relations with it
Part 3 (n_3 →)	<u>Liberals</u> will *claim* that fellow liberal democracies share their ends and that illiberal states do not
Part 4 (n_4 →)	<u>Liberals</u> will *not change* their assessment of foreign states during crises unless those states change their institutions
Part 5 (n_5 →)	<u>Liberal elites</u> will *agitate* for their policies during war-threatening crises
Outcome (Y)	During crises, <u>statesmen</u> will be constrained to follow liberal elites, thereby not going to war with other liberal states

Source: Owen 1994.

Quantitative statistical studies attempt to make inferences about the size of the causal effects that independent variables have on a dependent variable in the population of the phenomenon (cross-case inferences), whereas process-tracing research aims to make inferences about the presence/absence of hypothesized causal mechanisms in a single case (i.e., within-case inferences are being made).[2]

What Type of Empirical Material Is Used to Make Inferences?

First, the predicted evidence that has to be collected to test whether part 5 of the mechanism is present is very different from the types of evidence that are relevant to test the presence/absence of the other parts of the mechanism. For example, evidence that can test whether liberals trust other states that they consider liberal is very different from evidence for part 5. Put simply, evidence of part one is noncomparable with evidence of part 5 within the individual case. Evidence is also noncomparable across the four cases that Owen investigates. The evidence that Owen collects for other cases, such as Anglo-American relations from 1803 to 1812, differs greatly from that collected for the Franco-American crisis. Furthermore, pieces of evidence are attributed different inferential weight based on their case-specific probability (chapter 7), with the pronouncements of key liberal actor Jefferson seen as a particularly important bellwether for liberal elite agitation.[3]

Collier, Brady, and Seawright (2010b: 184–88) have introduced a useful distinction between the types of empirical material gathered in quantitative statistical analysis and qualitative case study research. In statistical analysis, the ambition is to gather what can be termed *data-set observations* (DSOs), defined as all of the scores in a row in a rectangular data set for a given case for both the dependent and all of the independent variables (see also Bennett 2006). If we transformed Owen's mechanism into a standard X/Y theory, the data-set observations of his study could be depicted as seen in table 5.2.

TABLE 5.2. Data-set Observations of Owen's Liberal Democratic Peace Theory

Case	X = Perceptions of Opponent as Liberal Democracy?	Y = Armed Conflict?
F-US crisis (1796–98)	Yes	No
UK-US crisis (1803–12)	No	Yes
UK-US crisis (1861–63)	Yes	No
UK-US crisis (1895–96)	Yes	No

Source: Based on Owen 1994.

In the data set, we can see that there is a clear correlation between perceptions of an opponent as a liberal democracy (X) and a lack of armed conflict (Y). However, the mechanism whereby causal forces are transmitted from X to produce Y is black-boxed.

In contrast, in process-tracing, another type of empirical material is gathered. Collier, Brady, and Seawright suggest that this type of material be termed *causal process observations* (CPO), defined as "an insight or piece of data that provides information about the context or mechanism and contributes a different kind of leverage in causal inference. It does not necessarily do so as part of a larger, systematized array of observations" (2010b: 184).[4]

Unfortunately, the term *causal process observation* conflates observations with evidence. Observations are raw material data; they become evidence only after being assessed for accuracy and interpreted in context (chapter 7). This distinction is analogous to how the two terms are used in criminal proceedings, where observations by themselves have no evidentiary importance; empirical material must be assessed for accuracy and interpreted in its context to be admitted to court as evidence for/against a specific theory about why a crime occurred.

In process-tracing, this assessment is undertaken using case-specific contextual knowledge; therefore, we can depict the evaluation process symbolically as $o + k \rightarrow e$, where o is an observation, k is case-specific knowledge, and e is the resulting evidence produced by the evaluation process. After evaluation, empirical material can be termed *evidence*, which can then be used to make inferences that update our confidence in the presence of a hypothesized causal mechanism (see chapter 7).

With regard to part 5 of Owen's hypothesized mechanism, evidence includes speeches by leading liberal elites (e.g., Jefferson) and editorials published in leading liberal newspapers. This evidence seeks to measure the activities (agitation) by the theorized entity (liberal elites) and thus can capture the transmission of causal forces from X to Y in ways that are impossible to measure using standard DSOs.

Even more important regarding the possibility of making cross-case inferences, evidence in the Franco-American crisis case is not comparable with evidence of the same part in the other three cases, making these in effect four parallel single-case studies. Mechanisms have different empirical manifestations in different cases. For example, in testing part 5 of the mechanism in the 1796–98 Franco-American crisis case, relevant evidence includes actions by Jefferson-led Republicans in Congress against a war declaration along with the finding that "the Republican press shrieked in protest" (Owen 1994: 107). In contrast, the evidence produced in the Anglo-American crisis during the Civil War includes the statement that "the [Emancipation] Proclamation energized evangelical Christian and other emancipation groups in Britain" and that mass rallies took place in Manchester at the end of 1862 and in London in the spring of 1863 against British recognition of the Confederacy (112). Given the very case-specific nature of the observable implications of the mechanism in the different cases, what counts as evidence in one case is not necessarily evidence in the other. What is improbable evidence in one case can be probable in another, also making observations noncomparable across cases (see chapter 7). Consequently, the cases cannot be compared directly, making within-case inferential tools necessary.

The Nature of Cross-Case versus Within-Case Inferences

Within qualitative methods, a distinction exists between quantitative-inspired researchers such as KKV, who argue for a "unified" logic that seeks to make inferences about the size of causal effects of independent variables on dependent variables in a population of a given phenomenon (King, Keohane, and Verba 1994; see also Gerring 2005), and qualitative scholars, who contend that in-depth case study methods such as process-tracing have a fundamentally different inferential ambition, which is to detect whether a hypothesized causal mechanism is present in individual cases (Bennett 2008b; Collier, Brady, and Seawright 2010a; Mahoney 2008; McKeown 2004; Munck 2004).

Arguments for making inferences about the size of mean causal effects in a population of a phenomenon build on the combination of the regularity and probabilistic understandings of causality. In KKV's understanding, causal effect is defined as "the difference between the systematic component of observations made when the explanatory variable takes one value and the systematic component of comparable observations when the explanatory variables takes on another value" (King, Keohane, and Verba 1994: 81–82). Here the aim is to make cross-case inferences about mean causal effects of systematic independent variables for the whole population of a given phenomenon based on a selected sample of cases (usually five to twenty), making an inferential leap from the magnitude of the mean causal effects of the independent variables in the sample to the entire population of a given phenomenon.

KKV argue strongly for the position that a single-case study is an invariant research design that prevents us from making causal inferences about causal effects: "Nothing whatsoever can be learned about the causes of the dependent variable without taking into account other instances when the dependent variable takes on other values" (King, Keohane, and Verba 1994: 129). To overcome this problem, they argue that scholars should disaggregate single cases by creating multiple DSOs of the independent and dependent variables. This can be achieved by, for example, disaggregating spatially by transforming a country-level study into its geographical subunits (states, counties, and so forth) or by disaggregating a case into a series of cases over time by testing variables at different points in time (t_0, t_1, t_2 . . . t_n) (217–28). However, KKV view the disaggregation strategy as second-best as a consequence of the lack of independence between DSOs made in the disaggregated "cases." For example, values of variables in t_0 will be expected to affect values at t_1, thereby violating the conditional independence of observations rule that makes inferences possible across cases in the frequentist logic of inference (94–97).

Process-tracing research has a different inferential ambition. What we are doing in process-tracing is making inferences about whether causal mechanisms are present in a single case. We naturally want to know that the independent variable(s) and the ensuing causal mechanism have nontrivial causal effects on the outcome, but we are not interested in assessing the magnitude of the causal effects of the combination of an independent variable and the ensuing mechanism on a dependent variable. Instead, we are interested in how a causal mechanism contributed to producing Y or, in explaining-outcome process-tracing, in how a combination of mechanisms produced the particular outcome. We want to know whether evidence suggests that a causal mechanism is present in a particular case.

In contrast to what quantitative-inspired methodologists contend (King, Keohane, and Verba 1994), we contend that strong within-case inferences are possible based on a single in-depth case study using process-tracing methods. This involves adopting a different inferential logic than that suggested by KKV. The difference in the inferential logics in the two understandings of case study research can best be understood by analogy. The understanding of case study research espoused by KKV resembles a medical experiment that studies whether a treatment given to a group of patients has a substantial impact in comparison to a control group that receives placebo treatments. In contrast, inference in process-tracing is more analogous to a court trial, where the researcher assesses our degree of confidence in the existence of a causal mechanism linking X with Y based on many different forms of evidence collected to test the existence of each part of the hypothesized causal mechanism (McKeown 2004). Here, one piece of evidence can be enough to infer that a part of a causal mechanism exists based on the Bayesian logic of inference (Bennett 2008b).

5.2. What Type of Inferential Logic Can Be Used in Process-Tracing?

How can we make within-case causal inferences about whether the parts of a theorized causal mechanism are actually present? We investigate whether mainstream inferential logics such as the frequentist logic of inference espoused by King, Keohane, and Verba or the comparativist logic of elimination used in comparative methods can be used to make strong within-case inferences about the presence/absence of the parts of a causal mechanism. We argue that neither provides a basis for making within-case inferences about whether a causal mechanism is present. As a result, we need to adopt a different inferential logic to bring methodology in line with process-tracing's ontology of causality.

Table 5.3 illustrates the differences in the three types of inferential logics: the frequentist logic as it has been adapted by KKV to qualitative case study research, the comparativist logic of elimination, and Bayesian logic of subjective probability. Both the frequentist and the comparative method build on a regularity understanding of causality, although they diverge regarding whether we should adopt a deterministic or probabilistic understanding.[5] In contrast, process-tracing utilizes a mechanismic and deterministic understanding.

The types of evidence that form the basis for inferences are also different, with the frequentist logic drawing on homogenous DSOs, whereas process-tracing inferences are made using noncomparable process-tracing evidence. The comparative method utilizes a form of DSO, although they are typically a thicker and more in-depth controlled comparison of two or more cases than the type of observations advocated by KKV (see Mahoney and Rueschemeyer 2003).

The frequentist logic of inference assesses the magnitude of causal effects of X on Y, or the degree to which the presence of X raises the probability of Y in a population (Gerring 2005). In contrast, the comparative method aims at assessing necessary and/or sufficient conditions that produce Y (either in a population or a small number of cases) (Mahoney 2008), whereas process-tracing seeks to detect the transmission of causal forces through a causal mechanism to produce an outcome in a single case.

TABLE 5.3. Three Different Logics of Inference in Social Research

	Frequentist Logic in Qualitative Case Study Research (KKV)	Comparativist Logic of Elimination	Bayesian Logic of Subjective Probability (process-tracing)
Ontological understanding of causality	Regularity and probabilistic	Regularity and deterministic	Mechanismic and deterministic
Inferences made using:	Classic probability theory and predicted probability that a found association is random or systematic	Mill's methods of agreement and difference and variants of them	Bayes's theorem about the expected likelihood of finding specific evidence in light of prior knowledge
Types of causality assessed	Mean causal effect of X's upon Y	Necessary and/or sufficient conditions that result in Y	Presence/absence of causal mechanism (i.e., transmission of causal forces from X to produce Y)
Types of observations used to make inferences	Relatively large set of "thin" data-set observations (5–20 according to KKV)	Smaller set of "thicker" DSO's (typically 2–5 in-depth case studies)	Evidence ($o + k \rightarrow e$), where one piece of evidence can be enough to make an inference depending upon its probability
What counts as an observation?	Independent observations of X and Y across a comparable and relatively randomly selected set of cases	Individual cases of the phenomenon (e.g., Denmark and Sweden)	Observed evidence of whether we find the expected observable manifestations of each part of a causal mechanism
Analytical priority	Theory-centric focus, generalization from sample of observations to general population of phenomenon	Both theory- and case-centric focus	Both theory- and case-centric focus
Types of inferences made	Cross-case inferences (to population of phenomenon) or within-case inferences if congruence method used	Cross-case inferences (but smaller scope population) [contextualized])	Within-case inferences

Beach, Derek; Pedersen, Rasmus Brun. Process-Tracing Methods : Foundations and Guidelines.
Ann Arbor, MI, USA: University of Michigan Press, 2013. p 86.
http://site.ebrary.com/lib/fuberlin/Doc?id=10658497&ppg=86

The frequentist logic of inference can be used to make population-wide inferences about the mean causal effects of independent variables on a dependent variable, inferring from the sample of observations to the broader population of the phenomenon using the frequentist logic adapted from classic probability theory. Comparative methodology and process-tracing are typically more case-oriented in that thicker studies of each particular case are made. However, only process-tracing can in our view be used to make within-case inferences about causal mechanisms; comparative methods use the comparativist logic of elimination that by definition can only make cross-case inferences.

We now turn to a discussion of why the frequentist logic and the comparativist logic of elimination cannot be used in process-tracing research.

The Frequentist Logic of Inference in Qualitative Case Study Research

KKV contend that "the differences between the quantitative and qualitative traditions are only stylistic and are methodologically and substantively unimportant. All good research can be understood—indeed, is best understood—to derive from the same underlying logic of inference" (King, Keohane, and Verba 1994: 4). The inferential logic that KKV advocate is in essence the same frequentist logic that is used in classic statistical analysis.[6]

For frequentist-oriented scholars, the gold standard of scientific research is the experimental design. In a medical experiment, for example, researchers examine the size of the effects that a given treatment (X) has on a group of patients in comparison to a control group that receives placebo treatments (not X). The size of the causal effect of the treatment is then found by assessing the value of Y in the group where the treatment is administered in comparison to the value where it is not given (the control group). However, even in the best-designed experiment, we cannot measure the effect of treatment and nontreatment on the same patient at the same time. This is what Holland (1986) terms the fundamental problem of causal inference—it is impossible to observe the effect of X and not X on the same unit of Y.

Building on classic probability theory, Holland offers what can be termed a frequentist logic of inference as a second-best solution to the fundamental problem of causal inference. The gist of the logic is that when certain assumptions hold, we can proceed as if we were able to measure both treatment and nontreatment on the same unit, using classic probability theory to assess whether any found associations between X and Y are the product of chance or are the result of a systematic correlation that can be interpreted as a causal effect of X on Y. Four main assumptions need to be fulfilled when using the frequentist logic of inference to make casual inferences (Holland 1986). First, the units of a sample must be homogenous, enabling comparisons that approximate a treated group/control comparison of the same unit. Second, X must temporally occur prior to Y. Third, the DSOs of each unit must be independent of each other, usually the result of random selection of the sample. Finally, we must be able to assume that the size of the average causal effect of X is constant on every unit in the population.

In the frequentist logic of inference in quantitative social research, a large number of DSOs are analyzed using statistical models such as linear or logistic regression to test whether there are any systematic correlations in the empirical patterns of association between X and Y. In terms of classic probability theory, the association is tested to see whether it is the product of nonsystematic (i.e., random) or systematic variation. Social scientists can then base their inferences that a causal relationship exists between X and Y on the classic statistical probability that the found correlation is not a product of chance (expressed in terms of statistical significance and p-levels). In the frequentist logic, the inferential heavy lifting is done by the laws of classic statistical probability theories about the distributions of populations, such as the central limit theorem.

For example, in a study in international relations of whether three liberal variables (democracy (X_1), economic interdependence (X_2), and membership in international institutions (X_3) resulted in more or less conflict between states (Y), Oneal, Russett, and Berbaum (2003) use a logistic regression model to analyze the association between the liberal independent variables and the dichotomous dependent variable (war/peace), controlled for a series of competing independent variables such as alliances. The size of the sample is 231,618 DSOs. They find a statistically significant relationship among all three Kantian Xs and the dependent variable. More important, they also estimate the marginal effects measured in terms of the predicted probabilities of war for each of the independent variables, finding a relatively large impact of democracy in comparison to the two other independent variables (table 2, p. 382). What enables Oneal, Russett, and Berbaum to make the inferential leap from their sample of observations to conclude that there is a substantively important causal relationship between democracy and peace more generally in the population are classical statistical laws of probability that assess the degree of probability that the association between X_1, X_2, and X_3 and Y estimated by the statistical model is one of chance.

KKV have adapted the frequentist logic of inference so that it can function as the foundation for methodological prescriptions for qualitative case study research (King, Keohane, and Verba 1994). KKV suggest that to make inferences about the mean causal effect of an X on Y, we should have between five and twenty independent observations of values of X and Y (DSOs). When the assumptions of unit homogeneity and what they term the conditional independence of observations hold,[7] KKV argue that inferences about the magnitude of causal effects can be made using case study methods.

Can this frequentist logic of inference be utilized in process-tracing studies? The short answer is no. As discussed earlier, KKV's frequentist logic of inference builds on logic similar to a medical trial, whereas process-tracing is more akin to a court trial.

In a typical process-tracing research design, the only assumption of Holland's that is usually fulfilled is the temporal one, where most tests of mechanisms involve also testing whether each part of the mechanism is temporally prior to the outcome. The assumption about unit homogeneity is not fulfilled in a process-tracing case study, as the manifestations of causal mechanisms differ both with regard to the individual parts of a mechanism and across cases. In addition, the inferential weight of individual pieces of evidence can also differ markedly.

For example, in Owen's (1994) study, the evidence collected for part 1 is not comparable with the evidence for parts 4 or 5. In contrast, using KKV's understanding, a case study would fulfill the assumption of unit homogeneity, given that by disaggregating over time we could observe the scores of X and Y at t_0, t_1, t_2, enabling them to be treated as comparable with each other. This is how congruence methods enable within-case inferences to be made.[8] Further, given the lack of comparability of evidence across the different parts of the mechanism, the assumption of independence loses its relevance when testing whether causal mechanisms are present. Basically, it matters little whether evidence of parts 1–5 are independent if we cannot compare them.

One could argue that the tests of whether each of the parts of a mechanism is present can be thought of as individual case studies that should be further disaggregated to create a sufficient number of comparable observations to enable the employment of the frequentist logic of inference. However, this approach would conflict with the nature of evidence that is collected in process-tracing research. While we do have multiple pieces of evidence of each part of the causal mechanism, they are noncomparable with each other. The difference between DSOs and evidence results from the different purpose that evidence serves in process-tracing research. In terms of the Owen example, we cannot just disaggregate the examination of part 5 of the mechanism ("liberal elites agitate") to multiple observations of whether or not elites agitate. Indeed, a single piece of highly improbable evidence, if found, can be enough to substantiate that the part of the mechanism was present (Bennett 2008b).

In process-tracing case studies, pieces of evidence should not be thought of in terms analogous to individual patients in a medical trial; instead, they can be seen as resembling evidence at a court trial. For example, in a trial dealing with a robbery, a part of the theory of the crime would be the suspect's motive. If it is theorized that the suspect committed the robbery to fund a drug habit, we would predict that we would find different types of evidence of the habit, such as witnesses who testify to having seen the suspect using drugs and physical evidence such as traces of drugs or drug paraphernalia found at the suspect's residence. Not all of these pieces of evidence would need to be found; indeed, our confidence in the validity of the part of the theory would be substantially increased if even one piece of highly incriminating evidence was found (e.g., a syringe used to shoot heroin with only the suspect's fingerprints on it).

In process-tracing, we never possess a sample of observations that enable us to use the frequentist logic of inference to make causal inferences. However, the nonapplicability of the frequentist logic does not result from a lack of stringent research design, as KKV would argue. Instead, it is an explicit choice. While it limits our ability to make cross-case causal inferences to a population of a phenomenon, process-tracing enables us to make stronger within-case inferences than would be possible using the frequentist logic.

The Comparative Method's Logic of Elimination

Comparative methods deal primarily with finding and/or eliminating necessary and/or sufficient conditions that produce a given outcome. For example, using comparative methods, Skocpol (1979) found that a combination of state breakdown and peasant revolt were sufficient conditions for producing social revolution in agrarian-bureaucratic states. Necessary conditions are defined as causes that must always precede Y for Y to occur; Y will not occur if X is absent. Sufficient conditions are causes that, if present, always produce Y, but Y is not always preceded by X.

While it uses the same regularity understanding of causality as the frequentist logic, comparative methodology utilizes another logic to make causal inferences that can be termed the comparativist logic of elimination. The basic idea of the logic of elimination in comparative methodology is based on John Stuart Mill's methods of agreement and difference. Mill's method of agreement is used to eliminate potential necessary causes. Here, all of the instances of Y (e.g., social revolutions) are examined,[9] and all potential conditions that are not present in all of the cases are eliminated as necessary conditions. While being able to make strong negative inferences (elimination), the method enables only very weak positive inferences. Mill suggested that conditions that survived the test were only possibly associated with the outcome, as it could not be ruled out that future cases would be discovered where the outcome occurred but the posited necessary condition was not present (George and Bennett 2005: 155–56). The method of difference is used to test for sufficient causation, where two or more cases that have different outcomes are compared. Conditions that are present in both types of outcomes are then eliminated as potential sufficient conditions. Mill's methods have been further developed to account for multiple causal paths to the same outcome, using more complex set-theoretical logic that draws on the same notion of a logic of elimination (Ragin 1988, 2000, 2008; Rihoux 2006).

The logic of elimination used in comparative methodology cannot form the basis for causal inferences in process-tracing research for two reasons. First, comparative case studies by definition deal with cross-case inferences based on a controlled comparison of cases, whereas process-tracing studies aim at within-case inferences about whether a causal mechanism is present in a particular case. The comparativist logic of elimination does not give us any inferential leverage to determine whether a part of a mechanism is present in a particular case study, as the comparative method's use of the regularity understanding results in the actual mechanism being black-boxed; instead, patterns of correlations are assessed using the logic of elimination.[10]

Further, even if we could find two fully comparable cases, when we delve into testing whether the theorized parts of a mechanism are present in both, the types of evidence that would be utilized for each part are noncomparable. What counts as strong evidence in one case might differ substantially from another case. Comparative methods offer us no tools for assessing the inferential weight of different process-tracing evidence in particular cases. For example, in Owen's study, to what extent does the observation of the key liberal actor, Jefferson, making statements supporting liberal France in the Franco-American crisis increase our confidence in the presence or absence of part 5 of the causal mechanism? Here, process-tracing—in particular, the Bayesian logic of subjective probability—gives us a set of inferential tools that enable us to evaluate in a transparent manner the inferential weight of different pieces of evidence.

Beach, Derek; Pedersen, Rasmus Brun. Process-Tracing Methods : Foundations and Guidelines.
Ann Arbor, MI, USA: University of Michigan Press, 2013. p 91.
http://site.ebrary.com/lib/fuberlin/Doc?id=10658497&ppg=91

5.3. The Bayesian Logic of Inference and Process-Tracing

Bayesian logic provides us with a set of logical tools for evaluating whether finding specific evidence confirms/disconfirms a hypothesis that a part of a causal mechanism exists relative to the prior expected probability of finding this evidence. Drawing on arguments developed by Bennett (2006, 2008b), along with Collier, Brady, and Seawright (2010b), we argue that Bayesian logic should be utilized as the inferential underpinning of process-tracing methods, enabling us to evaluate transparently and systematically the confidence that we can place in evidence confirming/disconfirming hypothesized causal mechanisms.

In Bayesian logic, the analyst gives greater weight to evidence that is expected a priori to be less probable based on previous knowledge of the phenomenon. "What is important is not the number of pieces of evidence within a case that fit one explanation or another, but the likelihood of finding certain evidence if a theory is true versus the likelihood of finding this evidence if the alternative explanation is true" (Bennett 2006: 341).

This reasoning is based on Bayes's theorem, which is a simple and uncontroversial logical formula for estimating the probability that a theory is supported by evidence based on the researcher's degree of belief about the probability of the theory and probability of finding given evidence if the theory is valid before gathering the data (Buckley 2004; Howson and Urbach 2006; Jackman 2004; Lynch 2005; Western and Jackman 1994).

The simplest version of Bayes's theorem is **posterior ∝ likelihood × prior.** This theorem states that our belief in the validity of a hypothesis is, after collecting evidence (posterior), equal to the probability of the evidence conditional on the hypothesis being true relative to other alternative hypotheses (likelihood), times the probability that a theory is true based on our prior knowledge. Here, we use the term *hypothesis* to refer to hypotheses about the existence of each part of a theorized causal mechanism.

Before we go further into a discussion of the different parts of Bayes's theorem, the idea that a theory, here understood as the existence of the parts of a causal mechanism, can be confirmed contradicts the Popperian falsification ideal of science that still forms the basis of many methodological textbooks in the social sciences. KKV approvingly quote Popper's categorical statement that "theories are not verifiable" (King, Keohane, and Verba 1994: 100). In contrast, Bayesian logic posits that we can both confirm and disconfirm our confidence in the validity of a theory—although given the uncertain nature of empirical observation, we can never be 100 percent confident about either confirmation or disconfirmation. Bayesian logic is closer to the actual practice of science, where we tend to have stronger confidence in the validity of theories that have withstood numerous independent empirical tests (Howson and Urbach 2006: 4). Bayes's theorem also predicts that a theory test that merely repeats existing scholarship with the same data will do little to update our confidence in the validity of the theory. This is due to the Bayesian principle of "updating," where the findings of a previous study form the prior for the next study; given that the data has been assessed in previous studies, finding the data is not surprising (high likelihood), and therefore little or no updating of the posterior takes place (Howson and Urbach 2006). In contrast, our belief in the validity of a theory is most strongly confirmed when we engage in new scholarship and find evidence whose presence is highly unlikely unless the hypothesized theory actually exists. For example, within international relations, confidence in the validity of democratic peace theory increased significantly in the 1990s as new scholarship tested the theory in different ways and with different forms of evidence. However, after initial tests, repeated tests of the thesis using the same data and methods do little to further update our confidence in the validity of the theory.

Confirmation—or, more accurately, an increase in our confidence about the validity of a theory—is achieved when the posterior probability of a theory exceeds the prior probability before evidence was collected. If there is a high prior probability (i.e., existing scholarship suggests that we should be relatively confident about the validity of a theory) and the evidence collected is the same as used in previous studies, additional tests do little to update our confidence in the theory.

The three elements of Bayes's theorem are posterior probability, likelihood, and the prior. The full theorem can be expressed as follows (Howson and Urbach 2006: 21):

$$p(h|e) = \frac{p(h)}{p(h) + \frac{p(e|{\sim}h)}{p(e|h)}p({\sim}h)}$$

The term p(h|e) is the posterior probability, or the degree of confidence we have in the validity of a hypothesis (h) about the existence of a part of a causal mechanism after collecting evidence (e). The term p(h) is the prior, which is the researcher's degree of confidence in the validity of a hypothesis prior to gathering evidence, based on existing theorization, empirical studies, and other forms of expert knowledge. The likelihood ratio is the expected probability of finding evidence supporting a hypothesis based on the

Beach, Derek; Pedersen, Rasmus Brun. Process-Tracing Methods : Foundations and Guidelines.
Ann Arbor, MI, USA: University of Michigan Press, 2013. p 93.
http://site.ebrary.com/lib/fuberlin/Doc?id=10658497&ppg=93

researcher's interpretation of the probability of finding it in relation to the hypothesis and background knowledge informed by previous studies, compared with the expected probability of finding the evidence if the hypothesis is not true ($p(e|\sim h)$). This is expressed in the formula as $p(e|\sim h)/p(e|h)$. A prediction by a soothsayer that you will meet a tall dark stranger is highly probable to occur, meaning that observing it does little to update our confidence in the hypothesis that the soothsayer can predict the future. However, if the soothsayer correctly predicted the number of hairs on the stranger's head, this evidence would seriously increase our confidence in the validity of the hypothesis (Howson and Urbach 2006: 97).

It must be noted that no evidence confirms or disconfirms a theory with 100 percent certitude: Confirmation is a matter of degree. When we develop strong tests for the presence/absence of the parts of a hypothesized causal mechanism that can discriminate between predictions of evidence that confirm h and the alternative hypothesis, and when the observations collected are quite accurate, we can use Bayesian logic to update our degree of confidence in whether a causal mechanism existed.

This formula introduces a degree of subjective choice by the researcher in terms of expectations of the probability of the likelihood of finding certain evidence and the interpretation of our confidence in a theory based on existing theorization (the prior). This has been the subject of extensive critique by non-Bayesian scholars, who contend that it introduces an unacceptable degree of subjectivity into the scientific process. Bayesian scholars contend that these expected probabilities are not purely subjective but are based on the existing body of prior research (Chalmers 1999; Howson and Urbach 2006). In addition, even if priors are somewhat subjective, after a series of empirical tests that increase confidence in the validity of a theory, the final posterior probability would converge on the same figure irrespective of whether two different values of the prior were taken initially (Howson and Urbach 2006: 298).

Further, Bayesians counter that many subjective decisions are also made when using the frequentist logic of inference, such as significance levels, the choice of null hypothesis, and the statistical estimation method. These choices are usually not very transparent and are often quite arbitrary (Howson and Urbach 2006; Lynch 2005: 137; Western 1999: 11). In the Bayesian logic of inference, "subjective" choices are made explicitly and transparently, and they are less subjective than they would appear since they are informed by existing scientific knowledge, documented for example in the form "Based on sources X, Y and Z, we have a relatively high level of prior confidence in theory A."

Given the resemblance of process-tracing to detective work, we illustrate

Beach, Derek; Pedersen, Rasmus Brun. Process-Tracing Methods : Foundations and Guidelines.
Ann Arbor, MI, USA: University of Michigan Press, 2013. p 94.
http://site.ebrary.com/lib/fuberlin/Doc?id=10658497&ppg=94

the relevance of Bayesian logic for causal inference in process-tracing using an example drawn from a classic Sherlock Holmes story, *Silver Blaze* (A. C. Doyle 1975). The crime in the story is the mysterious abduction of a racehorse, Silver Blaze. After Holmes develops a theory of a mechanism that hypothetically can explain the crime based on his prior experience, he travels to the scene of the crime to investigate whether evidence supports the existence of each part of his hypothesized mechanism. One of the parts of this theorized mechanism deals with whether the perpetrator is a stranger or an insider. Holmes utilizes the observation that a particularly obnoxious stable dog did not bark on the night of the crime to evaluate the validity of this specific part of the mechanism.

> "You consider that to be important?" he asked.
> "Exceedingly so."
> "Is there any other point to which you would wish to draw my attention?"
> "To the curious incident of the dog in the night-time."
> "The dog did nothing in the night-time."
> "That was the curious incident," remarked Sherlock Holmes. (A. C. Doyle 1975: 24)

Expressed mathematically, let us say for illustrative purposes that Holmes's degree of confidence in hypothesis h (the hypothesis on the part of the mechanism that an insider was the culprit) is quite low (e.g., 20 percent), meaning that the prior (p (h)) is 20 percent and the more probable alternative hypothesis (~h) is 80 percent probable. The low value for the prior reflects the relative improbability of the part of the mechanism that hypothesizes that an insider abducted his own horse in comparison to more probable alternative explanations such as theft by a stranger of such a valuable asset.

After multiple visits to the stable, Holmes and Watson both experienced firsthand and heard testimony from observers that the stable dog always barked when strangers are near. Therefore, based on these observations and the improbability of alternative explanations of "no bark,"[11] we would expect that it is highly likely that the evidence of the dog that did not bark (e), if found, would support h (p (e|h) = 90%; p (e|~h) = 10%). Note that the inferential weight of evidence is evaluated in relation to case-specific knowledge.

Inserting the prior and the likelihood function into Bayes's theorem, we get the following posterior probability that expresses our degree of confidence in the hypothesis after collecting and evaluating the evidence:

$$69.2\% = \frac{0.2}{0.2 + (0.1/0.9) \times 0.8}$$

This example shows how the Bayesian theorem provides us with an inferential tool that allows us to update the degree of confidence we have in a hypothesis after evaluating the inferential weight of collected evidence. In the Holmes example, after finding the evidence of the dog not barking, based on the probabilities we attached in the preceding, we find that Holmes was justified in holding a much higher degree of confidence in the validity of the hypothesis after evidence was gathered (from 20 percent to 69.2 percent).

The key point here is that one piece of evidence significantly increased our confidence in the validity of the hypothesis—something that never would be the case when using the frequentist logic of inference. In other words, pieces of evidence do not necessarily have the same inferential weight in process-tracing (Bennett 2008b). Inferential weight is instead a function of the expected probability of a piece of evidence given the hypothesis in relation to the expected probability of finding the evidence given the alternative hypothesis.

When we are using the Bayesian logic of inference in process-tracing, we usually do not express priors and likelihood functions in mathematical terms (although there is no logical reason why this cannot be done), but the Bayesian logic plays the same essential role in evaluating causal inferences that frequentist logic plays in quantitative large-*n* research.

We find it unfortunate that many qualitative scholars using process-tracing do not make their priors and likelihood ratios explicit in their analysis. The result is that their work all too easily falls prey to the criticism that their inferential reasoning is soft. In our view, priors and likelihood ratios should be made as explicit as possible in one's analysis. They should be explicitly described in terms of what we believe to be probable based on the existing theoretical debate and empirical studies using the terminology prior probability and the likelihood ratio of finding evidence.

For example, if we are studying the disintegration of the Soviet empire in the late 1980s/early 1990s, based on what we already know from realist theories of declining powers and historical studies (e.g., Gilpin 1981; Kennedy 1988), we would expect that Soviet decision makers would contemplate a last throw of the dice by engaging in foreign military adventures during the power transition. This prior probability would inform our inferences when studying the disintegration of the Soviet empire. Therefore, if we are testing sociological hypotheses about the development of norms of cooperation and we find evidence that matches the predicted evidence that would support the sociological hypothesis (which would predict that Soviet decision makers never even contemplated violence), this found evidence, given that it is not likely unless the sociological thesis is true, would greatly increase the degree of confidence we have in the validity of the sociological hypotheses in comparison to our existing knowledge of past power transitions.

Bayesian logic is what enables process-tracing scholars to make within-case causal inferences in the same manner that frequentist logic and the logic of elimination enables cross-case causal inferences to be made in quantitative and comparative social research.

In the next section, we discuss what type of valid inferences can be made when we are using process-tracing methods. Finally, we discuss in more detail the limits to the types of inferences that can be made using process-tracing methods.

5.4. Making Causal Inferences in Process-Tracing Methods—Uses and Limits

While process-tracing enables strong causal inference to be made with regard to the presence of causal mechanisms in single cases (within-case) and in particular whether the individual parts of a whole mechanism are indeed present in the particular case, it is not compatible with generalizations beyond the individual case (cross-case inferences). When we attempt to generalize beyond the single case, we can no longer rely on process-tracing methods and the underlying Bayesian logic of subjective probabilities but instead need to adopt cross-case comparative methods that are built on either the frequentist logic or the comparativist logic of elimination (see chapter 8).

The types of causal inferences being made in process-tracing depend on which variant of process-tracing is used. As introduced in chapter 2, theory-centric variants of process-tracing (building/testing) have the inferential ambition to detect whether there is evidence suggesting that a causal mechanism was present in a case. Here inferences are first made about whether each part of the mechanism is present. Given that each part is theorized as being individually necessary, if there is evidence that significantly increases our confidence in the presence of each part, we can infer that the whole mechanism is present. Theory-centric process-tracing cannot make inferences about the necessity or sufficiency of the mechanism in relation to the population of the phenomenon. In contrast, explaining-outcome process-tracing seeks to

Beach, Derek; Pedersen, Rasmus Brun. Process-Tracing Methods : Foundations and Guidelines.
Ann Arbor, MI, USA: University of Michigan Press, 2013. p 97.
http://site.ebrary.com/lib/fuberlin/Doc?id=10658497&ppg=97

produce a sufficient explanation of a particular outcome, made possible by the adoption of pragmatic methodological ideas (see chapter 2).

Inferences in Theory-Centric Process-Tracing

The scope of inferences that can be made using theory-centric variants of process-tracing is restricted to whether or not a mechanism is present in a case. Neither inferences about necessity nor sufficiency of a mechanism in relation to the population of a phenomenon can be made. To prove necessity or sufficiency of conditions in relation to a population requires cross-case comparative methods, such as investigating all cases where Y is present to see whether the mechanism is also always present when Y occurs (see Braumoeller and Goetz 2000; Seawright 2002).

However, when we engage in theory-building or testing process-tracing research, the necessity and/or sufficiency of the theorized cause (X) has usually already been tested using more appropriate cross-case methods. In the case of democratic peace theory, many large-*n* cross-case studies have suggested that mutual democracy (X) is a sufficient explanation for peace between two states (Y) (e.g., Russett and Oneal 2001). Confirming the existence of a causal mechanism in a particular case (especially a "least-likely" case) substantiates a robust correlation between an explanatory variable (X) and an outcome that has been found using cross-case comparative or statistical methods, we can infer based on the process-tracing study that the correlation is actually a causal relationship where an X is linked with an outcome through a causal mechanism.

The belief that theory-centric process-tracing can be used to test two competing theories against each other is widespread but erroneous in most situations. In the complex social world, most outcomes are the product of multiple mechanisms acting at the same time. The inferences that can be made with theory-centric process-tracing are therefore restricted to claiming that a mechanism was present in the case and that it functioned as expected. No claims can be made about whether the mechanism was the only factor that resulted in outcome Y occurring—in other words, we cannot claim sufficiency based on a single theory test.

When engaging in theory testing, alternative theoretical explanations of the predicted evidence for a part of the causal mechanism form the ~h element of the likelihood ratio. However, in most situations, alternative theoretical mechanisms will not be able to form ~h for each part of a mechanism.

For example, if we are testing a rational decision-making mechanism, part 1 of our mechanism is that decision makers gather all relevant information. Testing this mechanism might involve a likelihood ratio where h is the part of the rational mechanism, while the most relevant alternative explanation (~h) for the predicted evidence could be a normative explanation. But for part 2, defined as identifying all the possible courses of action, the most plausible alternative explanation (~h) for the predicted evidence might be a bureaucratic politics explanation. In other words, a single competing mechanism will usually not provide the most plausible alternative explanation of evidence for each part of the causal mechanism.

The only exception is when it is possible to conceptualize and operationalize two competing mechanisms in a manner where they are composed of the same number of parts, each of them the polar opposite of the other and mutually exclusive. For example, Moravcsik's (1999) research design (see chapter 4) put forward a supranational entrepreneur mechanism formulated in a manner that was the diametrical opposite of an intergovernmental bargaining mechanism. If the analysis found that one existed, then the other logically could not exist. Each of the parts was operationalized in a manner where the competing theory logically formed ~h, enabling a competing theory test to be performed. Yet this is a rare situation in social research; more common is the situation where theories are acting at the same time. Therefore, we can only infer that a mechanism was present, cognizant that other mechanisms can also contribute to producing Y.

A more viable alternative for testing competing theories is to engage in a two-step research design, where competing theories are evaluated using the congruence method either before or after a process-tracing theory test. If congruence is used first, the competing theory (theory X_2) is conceptualized and tested as a causal theory ($X_2 \rightarrow Y$) instead of being transformed into a mechanism, as in process-tracing. Based on the value of the independent variable in the case (X_2), the congruence method involves testing whether the prediction about the outcome that should follow from the theory is congruent with what is found in the case (Blatter and Blume 2008; George and Bennett 2005: 181–204). If the predicted observable implications of the competing alternative are not found when examining the outcome of the case, then the analyst can proceed with the second step, using process-tracing to test whether a theorized mechanism (X_1 and a mechanism) existed and functioned as expected. If the mechanism is found, the conclusion can be drawn that X_2 was unable to account for the outcome of the case, whereas X_1 was causally related to the outcome through a found mechanism.

Using congruence after a process-tracing theory test is a way to check that the findings are not excessively biased toward what Khong terms "overly subjective interpretations of the raw data" (1992: 66). Here, alternative explanations of the outcome are tested using the congruence method, assessing whether they have greater congruence or consistency with the outcome than theory X_1 and a mechanism. Khong used this type of two-step analysis to bolster his inferences in a theory test about the role that an ideational mechanism (historical analogies) played in the case of U.S. decision making in the Vietnam War. After finding that the predicted outcomes of alternative theories did not match those of the case, he can make a stronger inference that the ideational mechanism was causally related to the decisions taken.

Inferences in theory testing are made for each part of a causal mechanism based on the Bayesian logic of inference, where each part of the mechanism is tested to see whether we can update our confidence in its presence or absence. Each part is theorized as being individually necessary. Therefore, if our confidence in the presence of one part is significantly decreased, we can infer that the mechanism as a whole was not present in the manner in which it was theorized. Our inferences about the presence of the whole mechanism are therefore only as strong as the weakest link in our empirical tests (see chapter 6). If we are unable to infer that the mechanism was present, the result is either the conclusion that the mechanism was not present or a round of theory building should be done to develop a more accurate causal mechanism.

Inferences in Case-Centric Process-Tracing

When the aim of a process-tracing study is to explain a particular outcome, the study is in effect attempting to confirm the sufficiency of an explanation. More precisely, we should seek to provide a "minimally sufficient" explanation with no redundant parts (Mackie 1965). An everyday example of sufficiency is that if the requisite background factors are present (power is available and the circuit and lightbulb are unbroken), when I turn on a switch, the light will always turn on when the mechanism linking the power source and the lightbulb becomes activated. In the same fashion, advocates of democratic peace theory postulate that having two democratic countries is sufficient to produce peace even in situations where the countries otherwise have such strong conflicts of interests that there is a significant risk of war.

Explaining-outcome process-tracing studies can start deductively or inductively (see chapter 2). For example, when we are attempting to explain a given instance of an otherwise well-studied phenomenon, analysts can proceed in a deductive manner, testing existing theories in a comparative theory

Beach, Derek; Pedersen, Rasmus Brun. Process-Tracing Methods : Foundations and Guidelines.
Ann Arbor, MI, USA: University of Michigan Press, 2013. p 100.
http://site.ebrary.com/lib/fuberlin/Doc?id=10658497&ppg=100

test to see what aspects of the specific outcome they can explain. When we are studying a little-studied phenomenon, the analyst can proceed in a manner more analogous with forensic science (as popularized by the TV show *CSI*). Here, the investigator starts with a corpse (the particular outcome) and then works backward to build a plausible theory based on the available evidence gathered from a multitude of different sources. The forensic evidence takes the form of physical evidence, such as hair samples, but the investigator also relies on other forms of evidence, such as the testimony of witnesses. This form of inductive study may appear to start from scratch, but in all but the most extreme instances, the investigator will draw inspiration from past experiences and existing theories.

In most circumstances, a single theorized causal mechanism is not sufficient to explain the outcome and therefore must be supplemented with new parts from other compatible theories or new theories to achieve minimal sufficiency. However, all the parts of this more complex mechanism must be individually necessary for the mechanism, and the overall mechanism need only achieve minimal sufficiency. These requirements focus our analytical attention on achieving as parsimonious an explanation as possible.

How do we know a minimally sufficient explanation when we see it? A good example can be seen in Schimmelfennig (2001). Here he takes as his point of departure two competing theorized causal mechanisms from rationalist and sociological theories of international cooperation to explain the existing EU member states' positions toward eastern enlargement. Not surprisingly, Schimmelfennig finds that neither can fully explain the outcome (neither is sufficient), and he therefore engages in theoretical synthesis, working backward from the outcome to fashion a more complex theoretical mechanism that is logically coherent, involving "rhetorical action." He provides strong evidence that the more complex mechanism is present in the case and that it is sufficient to account for the outcome. Sufficiency is confirmed when it can be substantiated that there are no important aspects of the outcome for which the explanation does not account (Day and Kincaid 1994).

We cannot, however, use process-tracing methods to determine how frequently a given mechanism is sufficient in relation to a broader population of relevant cases as a consequence of the inclusion of nonsystematic parts and case-specific conglomerate mechanisms, and because the outcome is unique (it is not a case of something). Basically, explanations are case-specific here (Humphreys 2010: 269–70). Therefore cross-case comparative designs are more appropriate if we want to study sufficiency across a population (such as a positive on cause design) (see Dion 1998).

Sufficiency naturally does not mean that mechanism X is the only true path to Y but merely that if mechanism X occurs, it is sufficient to produce Y. A classic example used by Mackie (1965) is how we can explain what caused a fire in a house (Y = house fire). In a specific case, a short-circuit, the absence of a suitably placed sprinkler, and the presence of flammable materials were sufficient to produce the fire. However, alternative paths to a house fire are possible, including a gas explosion or lightning coupled with the lack of sprinklers and the presence of flammable materials.

Conclusions

Table 5.4 illustrates what forms of inferences can and cannot be made using process-tracing methods. Process-tracing methods can be used to confirm/ disconfirm the necessity of the individual parts of a causal mechanism. While cross-case comparisons could be utilized if we conceptualized a mechanism as a set of intervening variables that can be compared across cases, we would then lose sight of the causal forces that hypothetically cause the mechanism to produce an outcome.

We have argued that explaining-outcome process-tracing can confirm/ disconfirm the minimal sufficiency of a mechanism in a single-case study but not for a broader population. Process-tracing methods cannot test for the necessity of a mechanism, and we therefore need to rely on other, cross-case comparative methods to test for overall necessity of a mechanism as a cause of a phenomenon.

In practice, most existing applications of process-tracing methods are

TABLE 5.4. Research Situations where Process-Tracing Methods Can Be Used to Make Causal Inferences

Testing for the Necessity of the *Parts* of a Causal Mechanism in Single Case	Testing for the Necessity of a Mechanism as a *Whole* at the Population Level	Testing for the Sufficiency of a Mechanism in Single Case	Testing for the Sufficiency of a Condition at the Population Level
Process-tracing (all variants)	Not process-tracing. Use instead comparative cross-case designs (e.g., positive on outcome design—all Y's).	Explaining-outcome process-tracing	Not process-tracing. Use instead comparative cross-case designs (e.g., positive on cause design, all X's).

Beach, Derek; Pedersen, Rasmus Brun. Process-Tracing Methods : Foundations and Guidelines.
Ann Arbor, MI, USA: University of Michigan Press, 2013. p 102.
http://site.ebrary.com/lib/fuberlin/Doc?id=10658497&ppg=102

hybrid studies. Process-tracing studies often are comprised of a set of two to five individual process-tracing case studies. Internally, each process-tracing study utilizes a mechanismic understanding of causality to engage in causal inferences about whether the individual parts of the mechanism are present and whether the mechanism as such is sufficient either to produce a specific outcome or to test whether the causal mechanism is present in the given case (thereby demonstrating a causal relationship between X and Y instead of a mere correlation). Here, within-case inferences are made using process-tracing. However, cross-case inferences are made using comparative designs, selecting for example a least-likely case or set of cases (see chapter 8).

Beach, Derek; Pedersen, Rasmus Brun. Process-Tracing Methods : Foundations and Guidelines.
Ann Arbor, MI, USA: University of Michigan Press, 2013. p 103.
http://site.ebrary.com/lib/fuberlin/Doc?id=10658497&ppg=103