Article

Sociological Methods & Research
41(4) 570–597

DOI: 10.1177/0049124112437709
http://smr.sagepub.com

The Logic of Process Tracing Tests in the Social Sciences

James Mahoney[1]

Abstract

This article discusses process tracing as a methodology for testing hypotheses in the social sciences. With process tracing tests, the analyst combines preexisting generalizations with specific observations from within a single case to make causal inferences about that case. Process tracing tests can be used to help establish that (1) an initial event or process took place, (2) a subsequent outcome also occurred, and (3) the former was a cause of the latter. The article focuses on the logic of different process tracing tests, including hoop tests, smoking gun tests, and straw in the wind tests. New criteria for judging the strength of these tests are developed using ideas concerning the relative importance of necessary and sufficient conditions. Similarities and differences between process tracing and the deductive-nomological model of explanation are explored.

Keywords

causal inference, case studies, hypothesis testing, deductive-nomological model

Analysts often seek to assess whether some occurrence was a cause of a particular outcome in a specific case; that is, they want to answer the

[1]Northwestern University, Evanston, IL, USA

Corresponding author:
James Mahoney, Departments of Political Science and Sociology, Northwestern University, Evanston, IL 60208-1006, USA
Email: james-mahoney@northwestern.edu

question: "Was X a cause of Y in case Z?" For example, historically oriented social scientists try to explain the occurrence of particular events such as the French Revolution, World War I, and the collapse of the Soviet Union. In doing so, they evaluate hypotheses about the specific causes of these events. Likewise, scholars who study the effects of variables in large populations of cases often explore whether a given value on a variable was a cause of a given outcome in one or more of their cases. They want to locate a general causal effect in particular cases. Yet how can analysts know if some event or value on a variable was a cause of an outcome in a particular case?

In this article, I explore how analysts make these inferences by looking at diagnostic pieces of evidence—usually understood as part of a temporal sequence of events—that have probative value in supporting or overturning conclusions about descriptive and explanatory hypotheses. Following Collier, Brady, and Seawright (2010), these pieces of evidence may be called *causal-process observations* (CPOs). Causal-process observations are used in conjunction with broader generalizations relevant to the case under analysis. In some instances, these generalizations are simply elementary understandings of associations that are nearly universally regarded to be true. In other instances, the generalizations take the form of inferences derived from scientific analysis. The overall methodology of using causal-process observations in conjunction with generalizations can be called *process tracing* (Bennett 2006, 2008; George and Bennett 2005; George and McKeown 1985).

Process tracing can be used as a method for evaluating hypotheses about the causes of a specific outcome in a particular case.[1] It is arguably the most important tool of causal inference in qualitative and case study research (Collier et al. 2010; George and Bennett 2005). The tests associated with process tracing can help a researcher establish that: (1) a specific event or process took place, (2) a different event or process occurred *after* the initial event or process, and (3) the former was a cause of the latter.

Two main kinds of empirical tests are used to achieve these goals, what Van Evera (1997:31-2) colorfully calls *hoop tests* and *smoking gun tests* (see also Bennett 2008:706; Collier 2011). A hoop test proposes that a given piece of evidence—namely, a specific causal-process observation—must be present for a hypothesis to be valid. Failing a hoop test eliminates a hypothesis, but passing a hoop test does not confirm a hypothesis. Smoking gun tests, by contrast, propose that if a given piece of evidence—namely, a specific CPO—is present, then the hypothesis must be valid. Passing a smoking gun test lends decisive support in favor of a hypothesis, though failing a smoking gun test does not eliminate a hypothesis.

Hoop tests and smoking gun tests can be used to evaluate hypotheses proposing that (1) certain specific unobserved events or processes occurred and (2) there is a causal connection between two or more events or processes. The first kind of hypothesis involves a descriptive inference about what actually happened in the history of a given case. The second kind of hypothesis seeks to establish causality among events or processes that are believed to have occurred within a given case.

The precise certitude of the inferences derived from hoop tests and smoking gun tests can vary. Ideally, hoop tests convincingly eliminate hypotheses, whereas smoking gun tests convincingly confirm hypotheses. Yet, in practice, one may not be able to carry out strong tests: For various reasons, which we shall explore, there may be considerable doubt. If there is such doubt, hoop tests and smoking gun tests become *straw in the wind tests*. These tests provide some evidence in favor of or against a hypothesis, but they are not decisive. They neither confirm nor eliminate the hypothesis in question (Van Evera 1997:32).

Although process tracing is a widely discussed method, its specific procedures are often not specified clearly. As Collier (2011:823) has recently noted, ‘‘too often this tool is neither well understood nor rigorously applied.’’ The present article is motivated by the need to further clarify process tracing tests, and it builds especially on the groundbreaking work of Bennett (2006, 2008; see also Bennett and Elman 2006), Collier (1993, 2011), George (1979; George and Bennett 2005; George and McKeown 1985), and Van Evera (1997) as well as relevant insights from the literature on causal mechanisms (e.g., Gerring 2010; Mayntz 2004; Waldner 2007). The article combines these ideas with tools developed by methodologists working on criteria for assessing the relative importance of necessary and sufficient conditions (Braumoeller and Goertz 2000; Goertz 2003, 2006; Mahoney 2008; Ragin 2000, 2008). By combining the literature on process tracing with the literature on the empirical assessment of the importance of necessary/sufficient conditions, the article specifies more precisely the method of process tracing and develops extensions of the main tests that are used with this method.

Among the insights gained are new ideas for evaluating the consequences of passing hoop tests and failing smoking gun tests. While scholars agree that failing a hoop test counts as decisive evidence against a hypothesis, they conclude that passing a hoop test ‘‘does not greatly increase confidence’’ in the validity of a hypothesis (Bennett 2008:706).[2] But recent writings on the relative importance of necessary conditions (Goertz 2006; Ragin

2008) provide a basis for specifying the difficulty of a given hoop test. Passing a *difficult* hoop test does lend positive support in favor of a hypothesis. Likewise, while scholars agree that passing a smoking gun test counts as decisive evidence in favor of a hypothesis, they suggest that failing a smoking gun test has little or no consequence for the validity of a hypothesis (Bennett 2008:706; Van Evera 1997:32). Yet recent work on the relative importance of sufficient conditions provides a foundation for understanding why some smoking gun tests are easier to fail than others. Failing an *easy* smoking gun test provides significant evidence that a hypothesis is not valid.

The criteria developed in this article for judging the strength of process tracing tests are intended to be applied to individual hypotheses. In practice, of course, one often seeks to test not only whether a given hypothesis is valid but also whether contradictory alternatives are *not* valid. One's assessment of the utility of an explanation depends in part on one's assessment of any rival explanations. Accordingly, the most useful CPOs will have strong probative value for both an explanation and its rivals (Bennett 2008; Hall 2006). Yet it is also true that separate process tracing tests are typically required for each individual explanation. Even when a single CPO has implications for adjudicating among multiple explanations, the generalizations that are used with that single CPO will usually be different for each explanation. The resulting process tracing tests likewise may well have varying probative value for each explanation. Given this, the really key issue involves specifying criteria for evaluating the strength of process tracing tests as applied to individual explanations. That is the issue addressed in this article.[3]

The article's focus on ideas about necessary and sufficient conditions is appropriate because process tracing tests are fundamentally built around these ideas. This is true in two ways that are worth distinguishing. First, as has been acknowledged, hoop tests and smoking guns tests are themselves defined by whether passing a test is necessary for confirming a given explanation (i.e., a hoop test) or whether passing a test is sufficient for confirming a given explanation (i.e., a smoking gun test) (Bennett 2008; Collier 2011). Second, as has not been acknowledged, the generalizations that are used with hoop tests and smoking gun tests involve necessary and sufficient conditions. The use of these generalizations is what allows process tracing tests to eliminate or confirm explanations. The various ways in which hoop tests and smoking gun tests use necessary and sufficient condition generalizations are developed systematically in this article.[4]

Finally, it bears emphasis that the criteria of necessity and sufficiency establish ideal benchmarks that are not always achieved in practice. The degree to which the generalizations used with process tracing tests achieve these benchmarks is one criterion that shapes the strength of the tests. To the extent that the tests cannot draw on generalizations about necessary or sufficient conditions, but rather must use probabilistic generalizations, they become straw in the wind tests. Straw in the wind tests point in the direction of a hypothesis being valid or not, but they can neither confirm nor eliminate it. It is also bears emphasizing at the onset that the generalizations used in process tracing tests can always be mistaken, with potentially devastating consequences for the validity of one's inference.[5]

Inferring the Existence of an Unobserved Event or Process

Assume that one hypothesizes that ''*X* was a cause of *Y* in case *Z*,'' where *X* and *Y* are particular occurrences or specific values on variables. For example, one might hypothesize that strong rural community solidarity was a cause of peasant revolution in eighteenth-century France (Skocpol 1979). Or one might hypothesize that progressive public policy was a cause of high female status in Kerala, India (Drèze and Sen 1989). A basic requirement for these hypotheses to be valid is that the cause and outcome actually occurred. Thus, strong rural community solidarity cannot be a cause of peasant revolution in France if such solidarity did not exist. And progressive public action cannot be a cause of high female status in Kerala if female status was not high in that part of India.

How can an analyst establish that a posited cause and/or an outcome actually occurred/existed in a single case? Here I consider how hoop tests and smoking gun tests are used for this descriptive purpose.

Hoop Tests

Passing a hoop test is *necessary but not sufficient* for the validity of a given hypothesis (Bennett 2008; Collier 2011; Van Evera 1997). The hypothesis must ''jump through the hoop'' to warrant further consideration. When hypothesizing that a cause or outcome occurred, therefore, the consequence of failing a hoop test is to *eliminate* the possibility that the cause or outcome took place. As we shall see, passing a hoop test will not confirm the hypothesis, though it might lend ''straw in the wind'' support for the hypothesis.

The analyst using hoop tests to explore whether a case experienced a particular cause or outcome addresses two basic questions. First, he or she asks whether the case possesses all conditions that are known to be *necessary* for the cause or outcome. These conditions are CPOs that have to be present if the cause or outcome occurred. For example, consider Skocpol's (1979) claim that a high level of rural community solidarity existed in eighteenth-century France but not in nineteenth-century Prussia (East of the Elbe). Skocpol notes that possessing some land and having some autonomy from landlords and state agents are necessary ingredients for having strong rural peasant communities (by commonsense definitions). To establish that Prussia lacked strong rural communities, therefore, she shows that it fails the hoop test: Peasants controlled at most only tiny plots and were under the close supervision of Junker landlords. By contrast, the hypothesis about France passes the hoop test: French peasants controlled substantial land as small holdings and lived in villages free from supervision by royal officials.

Second, to infer the existence of a cause or outcome, one can also ask about auxiliary traces that it would have left behind if it actually occurred. This kind of hoop test builds on the idea that causes and outcomes are often *sufficient* for certain subsequent traces that are not themselves the target of analytic interest. If the investigator can show clearly that one or more of these traces are definitely *not* present, then he or she can safely infer that the cause or outcome of interest did not take place (i.e., the hypothesis fails the hoop test). In other words, a *necessary* condition for the validity of a hypothesis about the existence of a cause or outcome is the presence of any auxiliary traces for which that cause or outcome is *sufficient*.

For example, consider Drèze and Sen's (1989) descriptive hypothesis that the state of Kerala in India is marked by relatively high female status, whereas most of India is not. Although high female status is a latent concept, the authors reason that if it is present, it would have to leave behind observable traces, such as relatively high literacy and survival rates for females. They show that India as a whole lacks these conditions, performing poorly on female literacy and life expectancy statistics. Thus, India fails the hoop test for having high female status. By contrast, the state of Kerala does well on these same indicators, such that it passes the hoop test.

Failing a hoop test is a standard way of inferring the absence of some event or process. But can passing a hoop test lend some positive evidence in favor of a hypothesis? The answer is yes. For example, the ability of Kerala to pass the hoop test discussed previously makes it far more likely that high female status exists in that part of India. The reason why is the hypothesis has passed a demanding hoop test that is difficult to pass unless it is true. Just as some hoops are smaller than others, and thus more difficult to jump through, some hoop tests are more demanding and thus harder to pass. While failing a hoop test will eliminate a hypothesis regardless of whether the test is easy or hard, passing a hoop test will lend positive support for a hypothesis in proportion to the degree that it is a difficult test.

The relative difficulty of passing a hoop test is directly related to the frequency with which the condition (i.e., the specific CPO) that is necessary for the hypothesis to be valid is present more generally (see Braumoeller and Goertz 2000; Goertz 2006; Ragin 2008). If the condition is always present, the hoop test is trivial, since the hypothesis will automatically pass. By contrast, if the condition necessary for the hypothesis to be valid is quite rare or abnormal to a given context, the hoop test will be hard to pass (see Hart and Honoré 1985). For example, it is unusual for a state in India (or other very poor countries) to perform reasonably well on female literacy and longevity. The hoop test requires Kerala to possess an abnormal condition, making it a difficult test.

In short, demanding hoop tests make reference to important (i.e., rare, abnormal) necessary conditions, and passing them increases the subjective probability that a hypothesis is correct.[6] In this sense, passing a demanding hoop test is equivalent to passing a strong straw in the wind test. Although we may not yet be certain that Kerala has high female status on the basis of passing the hoop test, the evidence is clearly pointing in that direction.

Smoking Gun Tests

Passing a smoking gun test is *sufficient but not necessary* for the validity of a given hypothesis (Bennett 2008; Collier 2011; Van Evera 1997). Thus, if one hypothesizes that an unobserved cause or outcome took place, the consequence of passing a smoking gun test is to *confirm* its existence. Failing a smoking gun test does not eliminate the possibility that the cause or outcome exists. However, this failure may count substantially against the hypothesis, depending on the difficulty of the test.

Smoking gun tests used to explore whether a case experienced a particular cause or outcome typically inquire about auxiliary traces for which the cause or outcome is a *necessary condition*. This kind of smoking gun test builds on the idea that causes and outcomes are often essential ingredients for certain subsequent traces that are not themselves the target of analytic interest. If the investigator can show clearly that one or more of these traces are present, then he or she can safely infer that the cause or outcome of interest took place. The existence of a condition for which the cause or outcome is *necessary* amounts to evidence that is *sufficient* for the validity of a hypothesis proposing that the cause or outcome exists.

This is the implicit method that is used to infer the existence of most ‘‘basic historical facts’’ that cannot be directly observed by the investigator.[7] Analysts know that certain events occurred in the past because these events leave behind traces that otherwise could not possibly exist. We have smoking gun proof that Abraham Lincoln existed: How else can we explain the hundreds of thousands of primary and secondary records depicting his life and actions? Likewise, we have smoking gun evidence that massive peasant revolts took place in eighteenth-century France. These revolts were necessary for countless traces of evidence that have persisted until today, including a vast number of sources describing features of the revolts that have been scrutinized by historical experts.

The authenticity of historical facts is questioned precisely when events or processes do not leave behind smoking gun traces. For example, some authors contend that Kerala is marked by high morbidity rates for both males and females. They derive this conclusion from self-reported data in Kerala about illnesses. Yet Drèze and Sen (1989) argue that these reports are hardly smoking gun proof that high female morbidity actually exists. They note that self-reported morbidity indicators are often extremely misleading insofar as one is interested in actual illness as opposed to issues of perception and communication. Drèze and Sen thus question the authenticity of the ‘‘fact’’ that Kerala has high female morbidity rates.

One can also carry out a smoking gun test by asking whether the case possesses any condition—or combination of conditions—that is known to be *sufficient* for the cause or outcome. If this condition or combination is present, then the cause or outcome must also be present. Yet, in general, social scientists have not discovered many regularities in which observable conditions are sufficient for a specific event or outcome. Instead, their regularities about sufficient conditions tend to link an unobserved event or variable value (e.g., high intelligence) to subsequent observable traits for which it is approximately sufficient (e.g., high scores on tests). Smoking gun tests in which analysts draw on preexisting generalizations about antecedent observable conditions that are sufficient for an unobserved event of interest are thus used less frequently in social science research.

The validity of any smoking gun test depends on the validity of the generalization that is used in the test. If there is good reason to doubt the validity of the generalization, the smoking gun test is not strong, and it cannot decisively support the hypothesis in question. One must then settle for a straw in the wind test that points in the direction of the hypothesis being valid. For example, if a given CPO is *usually* associated with the existence of a particular event, its presence certainly increases the subjective likelihood that the event exists, but there is still a decent chance that the event does not objectively exist.[8] As discussed in the following, the strength of the probabilistic generalization used in the test will dictate exactly how likely it is that an event exists if the hypothesis passes the straw in the wind test.

Finally, one might ask about the extent to which failing a legitimate smoking gun test counts against a hypothesis. The answer is that it depends on the difficulty of the test, which in turn is related to the more general commonality of the condition (i.e., the CPO) used in the test. All smoking gun tests make reference to a condition (or a combination of conditions) whose presence is sufficient for the validity of the hypothesis under investigation. However, the frequency at which this condition is present can vary. Hypotheses that fail a smoking gun test in which the condition is often present or is ''normal'' in a given context are more likely to be wrong than hypotheses that fail a smoking gun test in which the condition is only rarely present. Failing an easy smoking gun test (i.e., lacking a condition that is often present or normal) amounts to failing a strong straw in the wind test about the validity of the hypothesis.

As an example, consider Drèze and Sen's (1989:221) assessment of the hypothesis that Kerala is characterized by low calorie intake. Although good data on calorie intake are not available, the authors assert that the high incidence of ''severe undernutrition'' (for which there is good data) would nevertheless offer a smoking gun confirmation of this hypothesis. Yet Kerala is marked by comparatively low percentages of severe undernutrition. Thus, the hypothesis fails the smoking gun test. Moreover, because high incidences of severe undernutrition are common throughout India and the developing world, they imply that this failure counts quite heavily against the hypothesis.

Inferring a Causal Connection

If we agree that an antecedent event X occurs before a subsequent event Y in a particular case, how can we know if X causes Y? For example, given that peasants in eighteenth-century France exhibited strong rural community solidarity and carried out a peasant revolution, how can we further know that their rural community solidarity was a *cause* of their revolution? This section considers how hoop tests and smoking gun tests can be used to help answer these questions of causal inference.

The discussion builds on the existing literature on process tracing that stresses the importance of identifying causal mechanisms for making causal inferences. In this literature, causal mechanisms are often understood as intervening steps between an initial cause and a final outcome (Bennett 2008; George and Bennett 2005; Hall 2006; Mayntz 2004; see also Gerring 2010). The identification of one or more of these intervening steps is requisite for using process tracing tests to help infer a causal connection between two events that are believed to exist.

Hoop Tests

To assess a hypothesis about whether a known value on *X* was a cause of a known value on *Y*, a hoop test identifies a condition whose presence is *necessary* for the hypothesis to be true. The hoop test specifically asks about the presence of data concerning one or more intervening mechanisms linking *X* and *Y*. These within-case observations (i.e., CPOs) about possible mechanisms are needed to carry out hoop tests when evaluating causal hypotheses.

The form of within-case information concerning a mechanism required for a hoop test will vary depending on whether the hypothesis under consideration posits that *X* is necessary for *Y* or that *X* is sufficient for *Y* (including when *X* is a combination of factors sufficient for *Y*). Let us first assume that one has a hypothesis positing that *X* is *necessary* for *Y*. For example, Skocpol (1979) explicitly hypothesizes that in the counterfactual absence of rural community solidarity, a peasant revolution would not have taken place in eighteenth-century France. It follows logically that rural community solidarity was necessary for peasant revolution in France.[9]

To carry out a hoop test of this hypothesis, an analyst such as Skocpol can draw on existing general knowledge and identify a mechanism (*M*) that is known or has been established to be *sufficient* for the outcome. For the hypothesis to pass the hoop test, *X* must be necessary for the more proximate *M* as well as *Y*. The leverage gained by this kind of test derives from the fact that while *X* being necessary for *Y* is in doubt, the status of *M* being sufficient for *Y* and of *X* being necessary for *M* might be more readily established. These established or obtainable findings can be used to make a logical inference about whether *X* is necessary for *Y*. The inference is possible because if *X* really is necessary for *Y*, it must also be necessary for all intervening mechanisms that are sufficient for *Y*, including combinations of conditions that are sufficient for *Y* (Mahoney, Kimball, and Koivu 2009). Logically speaking, X *cannot be necessary for* Y *unless it is necessary for all intervening conditions that are sufficient for* Y.

To return to the Skocpol (1979) example, the historical literature establishes that certain specific rural rebellions that occurred in parts of France in 1789—composed centrally of struggles against seigniorial practices, especially in the north and northeast—were events sufficient (i.e., they were enough by themselves) to generate the overall outcome of ''peasant revolution'' in eighteenth-century France. Accordingly, Skocpol can use process tracing to explore whether her main causal factor—rural community solidarity—was necessary for these specific anti-seigniorial revolts that took place in 1789. She notes that the literature shows that the revolts were made possible by the assemblies and organizations that existed in the relevant peasant villages. Peasant communities with ''considerable property, community autonomy, and antiseigneurial solidarity . . . [possessed] a preexisting potential for antiseigneurial revolts'' (p. 125). In her judgment, then, the hypothesis passes a hoop test: The main causal factor appears to be necessary for a specific set of revolts sufficient to yield a peasant revolution.

The degree of support that a hypothesis about a necessary cause receives from passing a hoop test varies depending on the difficulty of that test, which in turn depends on importance of the mechanism CPO used in the test. If the mechanism used in the hoop test is an important sufficient cause (e.g., the mechanism is nearly essential for the outcome), the hoop test will be more difficult to pass and thus more supportive of the hypothesis when passed. In the Skocpol example, the mechanism is a series of anti-seigniorial revolts taking place in specific regions of France. Although these revolts may not have been fully necessary for a peasant revolution, they were sufficient, and they were the ones that most clearly generated the peasant revolution. In that sense, the hoop test is difficult: the mechanism is an important sufficient cause of peasant revolution in France. The fact that Skocpol can show that rural community solidarity was necessary for these revolts suggests that such solidarity also was necessary for the outcome of peasant revolution itself.

If one's initial hypothesis is that X is *sufficient* for Y (including when X is a combination of conditions that are jointly sufficient for Y), by contrast, a hoop test can be carried out in conjunction with any mechanism (M) that is *necessary* for the outcome. For the hypothesis to pass the hoop test, X must be sufficient for M. If X is not sufficient for M, the hypothesis fails the hoop test, and one concludes that X is not sufficient for Y. The assumption is that if X really is sufficient for Y, it must also be (according to elementary logic) sufficient for all intervening mechanisms that are necessary for Y (Mahoney et al. 2009). X *cannot be sufficient for* Y *if it is not sufficient for all intervening conditions that are necessary for* Y.

For example, Drèze and Sen (1989) consider and reject the hypothesis that progressive public policy might have been sufficient for high female status in Kerala. They make this argument by pointing out that a historically longstanding female-male ratio greater than unity was a necessary mechanism for reducing gender bias in Kerala. Progressive public policy, however, was not itself sufficient for this mechanism: The female-male birth rate had important cultural roots, including a partially matrilineal system of inheritance. Thus, the hypothesis fails the hoop test: Activist public policy could not have been sufficient for high female status in Kerala because it was not sufficient for certain intervening conditions (e.g., the female-male birth ratio) that are established as necessary for this outcome. Drèze and Sen instead argue that progressive public policy was a necessary but not sufficient condition for high female status.

Smoking Gun Tests

Smoking gun tests to assess causal connections also require the analyst to ask about the presence of mechanisms linking *X* and *Y*. Again, the specific tasks that need to be carried out vary depending on whether one hypothesizes that *X* is necessary for *Y* or that *X* is sufficient for *Y*.

If the cause is hypothesized to be *necessary* for the outcome, the analyst can run a smoking gun test by identifying a mechanism that is necessary for the outcome. The analyst then determines whether the cause itself is necessary for this mechanism. If the cause is necessary for a mechanism that is known to be necessary for the outcome, the cause itself *must* be necessary for the outcome. It would be logically impossible for the cause to be necessary for the mechanism but not the outcome (Mahoney et al. 2009). Hence, this test can provide *smoking gun* evidence in favor of the hypothesis.

The assumption behind the test is that although we may not know whether *X* is necessary for *Y*, we may know (or be able to establish) that *M* is necessary for *Y*. We can then explore whether *X* is necessary for *M*. The question of whether *X* is necessary for *M* may be more easily answered than the question of whether *X* is necessary for *Y*. This is true because, among other reasons, *X* and *M* are more closely situated to one another in time and thus can be more easily connected causally. As several analysts suggest, causal links between proximate events are often intuitively obvious, especially with necessary conditions (Abbott 1992; Goldstone 1998; Griffin 1993; Roberts 1996). Chains of linked necessary conditions provide a good opportunity for the analyst to carry out a smoking gun test: She or he can show how an initial cause was essential to put the overall sequence in motion, culminating in the outcome.

As an example, let us return once more to the hypothesis that progressive public action was necessary for high female status in Kerala, India. One way to try to confirm this hypothesis is to find an intervening mechanism that is necessary for high female status. Drèze and Sen (1998) strongly suggest that extensive female education is such a necessary mechanism. As they write, ''It would be surprising if a greater level of female education—and less gender inequality in the sharing of education—had not contributed to better prospects of a plausible life for women, both through raising the status of women and through increasing female economic power'' (p. 224). The question then becomes whether progressive public action was necessary for extensive female education in Kerala. The authors answer in the affirmative: ''Public policy [in Kerala] put much greater emphasis on general education and literacy than was the case in the rest of India, and the emphasis on female education was particularly exceptional'' (p. 223). In short, they suggest that equitable education was one necessary precursor for high female status in Kerala and that antecedent progressive public policy was necessary for this education. If true, it follows logically that progressive public policy must have been necessary for high female status in Kerala.

Finally, let us consider a hypothesis in which the cause (or a combination of causes) is believed to be *sufficient* for the outcome. To carry out a smoking gun test, one locates a more proximate mechanism that is established as sufficient for the outcome. One then shows that the cause is sufficient for this mechanism. The assumption is that if the cause is sufficient for a mechanism that is known to be sufficient for the outcome, the cause itself *must* be sufficient for the outcome. It would be logically impossible for the cause to be sufficient for the mechanism but not the outcome (Mahoney et al. 2009).

This kind of smoking gun test is routinely carried out by detectives and medical examiners. For example, a medical examiner during an autopsy may establish a proximate mechanism sufficient for death and a more remote cause sufficient for this mechanism. A proximate sufficient cause of death might be the transection of the left internal jugular vein. A sharp force injury brought on by a knife might be sufficient for this transection. The obvious conclusion that follows logically is that the knife wound is sufficient for the death.

In the social sciences, this kind of process tracing test is carried out when analysts develop arguments about a mechanism that is sufficient for the occurrence of an outcome of interest. For example, Skocpol (1979) holds that the combination of peasant revolution and state breakdown was a sufficient mechanism for a full-blown social revolution in France. In turn, she works to show how various other causal conditions (e.g., international pressures, dominant class political leverage, and peasant solidarity and autonomy) were jointly sufficient for this mechanism (Goertz and Mahoney 2005). To the extent that she can plausibly show that these causal conditions were enough to generate peasant revolution and state breakdown, she can then assert that they were also jointly sufficient for the outcome of social revolution itself.

In summary, process tracing tests draw on information concerning mechanism as a basis for causal inference. Although the tests are usually not carried out explicitly, they are often used implicitly by analysts working on comparative-historical and case study research. Appendix A provides three more examples of these tests from well-known studies that seek to explain democratization in England.

Process Tracing Versus the D-N Model of Explanation

All process tracing tests use and require—broadly speaking—a mode of reasoning in which the analyst derives logically a conclusion from a set of premises. The premises include ''facts'' from the case and one or more preexisting generalizations that can be applied to these facts. This approach shares some intriguing—though not often acknowledged—commonalities with the (in)famous deductive-nomological (D-N) model of explanation (Hempel [1942] 1965; Nagel 1961; Popper [1934] 1968). Like process tracing, the D-N model is designed to explain specific outcomes in particular cases using within-case data and relevant generalizations.

Yet there are also important differences between process-tracing tests and the D-N model of explanation (George and Bennett 2005; Roberts 1996). Sorting out these similarities and differences proves useful for elucidating further the method of process tracing.

Deductive and Inductive Explanation

In the field of logic, the concept of *deduction* has a particular meaning: A deductive argument is one in which the truth of the premises guarantees the truth of the conclusion. With the D-N model, the premises are a set of conditions and one or more ''law-like'' generalizations; the conclusion is typically an event in a specific case. The model is deductive because the event is a logical necessity, assuming the truth of the premises.

Process tracing tests also strive to be deductive in this specific sense. For example, consider Skocpol's (1979) line of reasoning that: (1) a necessary condition for rural community solidarity is the possession of significant land and some autonomy; (2) peasants in Prussia (East of the Elbe) did not possess much land or have any autonomy; and (3) therefore, rural community solidarity did not exist in Prussia. The conclusion (3) follows logically and necessarily from the premises (1) and (2). Obviously, either or both of the premises could be mistaken. But that is a separate issue from the deductive form of the argument.

An *inductive* mode of reasoning is used when the truth of the premises does not guarantee the truth of the conclusion. With the inductive-statistical (I-S) model of explanation, for example, the premises yield a probabilistic prediction about the occurrence of a specific event, even when they are all true (Hempel [1942] 1965). To be used at all effectively in the explanation of specific events, this I-S model requires that the analyst draw on a strong generalization.[10] One needs a strong generalization to be able to predict that the event is likely to occur (or not occur). Highly probabilistic generalizations that prohibit such predictions are of limited use with the I-S model (Hempel [1942] 1965; Scriven 1959; Costner and Leik 1964; Railton 1978).

Analogously, with process tracing, the absence of fully deductive arguments means that analysts carry out probabilistic tests—namely, straw in the wind tests. The strength of these tests depends in part on the strength of the generalization that is used with the test. If the generalization is highly probabilistic, the test is very weak: Passing it provides limited support and failing it provides limited disconfirmation. Yet, even if a strong test is not feasible, it may be possible to subject a hypothesis to multiple straw in the wind tests. Although a hypothesis that passes any one straw in the wind test may not be well supported, a hypothesis that passes several straw in the wind tests may generate a good deal of confidence in its validity. Likewise, while a false hypothesis might by chance pass any one straw in the wind test, the probability of a false hypothesis repeatedly passing multiple instances of such tests may be quite low. Hence, when deductive analysis is not possible, one can sometimes compensate by subjecting hypotheses to multiple inductive tests.

Nomological Statements and Necessary/Sufficient Conditions

To use the D-N model, a scholar must include one or more "law-like" statements in the premises of the argument. These statements are called *nomological* because although they are law-like, they are also possibly false. Nomological statements have a law-like form but they may or may not be true (Hempel [1942] 1965).

Although the original formulators of the D-N model did not usually discuss explicitly necessary and sufficient conditions, these are the kinds of generalizations that they had in mind as constituting nomological statements. They regarded a statement such as "all *X* are *Y*" as nomological (i.e., as a law-like regularity). Translated into the language of necessary and sufficient conditions, this statement says that *X* is sufficient for *Y* and that not *X* is necessary for not *Y*.

Given that necessary and sufficient conditions have a law-like form, it follows that the strongest process tracing tests are deductive arguments that utilize nomological statements. A strong hoop test eliminates a hypothesis because it employs a generalization that has a law-like form. The same is true about the ability of a smoking gun test to decisively support a hypothesis. Insofar as researchers using process tracing cannot draw on nomological statements, they are forced to settle for inductive straw in the wind tests.

Because the use of law-like generalizations in the social sciences is a source of controversy, it is important to make two clarifying points here (see also Roberts 1996). First, while these generalizations always pertain to more than one case, they are not universal laws that apply across all times and places. They are limited—perhaps severely limited—by scope conditions that specify the parameters within which they apply. For any given process tracing test, the key requirement is simply that the generalization apply to the kind of case under analysis.

Second, these generalizations may be relatively trivial relationships that amount to little more than platitudes. For example, the idea that possessing some autonomy and property is necessary for peasant community solidarity seems almost trivially true. But it is precisely the very high probability of the generalization that makes it especially useful for inferring whether peasant community solidarity is present in any given community. The quest for strong generalizations is one reason why researchers often study causal mechanisms that are closely situated to one another in time. The causal linkages between proximate events can frequently be covered by generalizations that are more or less truisms (Roberts 1996). They often involve basic modes of human behavior and psychology (e.g., people seek change when facing a crisis) that are so obvious that social scientists can simply take them for granted.

Covering Laws Versus Mechanisms

The real differences between the D-N model and process tracing are rooted in their alternative understandings of explanation, not their commitment to deductive arguments or their effort to find and use nomological statements.

The D-N model purports to offer valid causal explanations of specific events by subsuming them under covering laws. By contrast, while process tracing tests are often deductive and nomological, they adopt a different understanding of explanation. In particular, process tracing is built around a scientific realist understanding of explanation that emphasizes causal mechanisms.

To see these differences, let us recall the two main purposes of process tracing tests: (1) inferring the *existence* of an unobserved event or process and (2) inferring a *causal connection* between one specific event or process and another. Clearly, the first kind of process tracing test—that is, inferring the *existence* of an event—is distinct from *explaining* that event. For instance, when Skocpol uses process tracing tests to infer the existence of rural community solidarity in France, she does not purport to have explained such solidarity. Deducing the existence of something involves descriptive inference and is in no way equivalent to causal explanation. Yet advocates of the D-N model never made this distinction clearly. They assumed that an event was explained if it could be logically deduced from true premises. As Popper ([1934] 1968:38) put it, ''To give a *causal explanation* of an event means to deduce a statement which describes it, using as premises of the deduction one or more *universal laws*, together with certain singular statements, the *initial conditions*'' (emphasis in original).

Scholars who use process tracing, by contrast, reject the view that an event is explained when it can be subsumed under and predicted by a covering law model. As critics have pointed out many times, the D-N model's equation of prediction with explanation is deeply unsatisfying. To cite only one famous example: A barometer's reading may predict weather patterns, but the barometer certainly does not causally explain these patterns (Salmon 1990:46-7). As an alternative, scientific realists of various stripes maintain that to explain a phenomenon is to identify the causal mechanisms that permit and/or generate the phenomenon (e.g., Harré 1970; McMullin 1984).

This alternative model of explanation focused on causal mechanisms is more consistent with what process tracing analysts are doing when carrying out tests to establish a causal connection between two events (George and Bennett 2005). Process tracing tests that are designed for causal inference always require the analyst to locate and draw upon mechanisms. Without locating CPOs that embody information concerning mechanism, one cannot use process tracing tests to help establish that one event causes another.

In practice, the timing of the discovery of relevant causal mechanisms and the specification of the specific hypotheses being tested with process tracing can vary. In some instances, the discovery of one or more CPOs related to a causal mechanism may *precede* the formulation of a specific hypothesis and a process tracing test. For example, the analyst may discover an unexpected yet essential piece of data concerning a mechanism while learning about the case—much as a detective might stumble upon an unanticipated but decisive clue in the course of an investigation. Although the analyst was not specifically looking for the data, its discovery has immediate value because it suggests a hypothesis and a process tracing test that had not been previously conceived.

Alternatively, the analyst might first design a process tracing test and then intentionally look for a piece of data concerning a particular causal mechanism required to carry out that test. One can think of this as analogous to a detective in search of specific evidence that will help solve the mystery. From some epistemological standpoints, this is an especially strong mode of inference because the test is conceived prior to the analysis of the data (e.g., Popper [1934] 1968). The investigator specifies *in advance* the kind of evidence concerning mechanism that counts in favor of or against a hypothesis and then looks to see if those data are present.

It is important to note, however, that process tracing can yield strong results even when the test is formulated post hoc as a response to the serendipitous discovery of data. This kind of post hoc process tracing seems quite common in qualitative and case-oriented research. Like detectives, qualitative and case-oriented researchers begin their investigations with an outcome to be explained and several different theories and hunches. In the course of the investigation, they discover unanticipated and perhaps unforeseeable CPOs that nevertheless have tremendous bearing on their final conclusions. They formulate new hypotheses and carry out process tracing tests in light of these discoveries.[11] The newly found data may in fact lead the researcher to look for additional CPOs from within the case that can be used for still other process tracing tests. In this sense, when cases contain many valuable CPOs, it is possible and useful to employ the same case for developing and testing a hypothesis (see also Rueschemeyer 2003). One simply draws on different CPOs from within that case to carry out multiple process tracing tests.

Conclusion

Process tracing involves both searching for within-case data to carry out strong tests of hypotheses and using knowledge of existing generalizations to design such tests. To excel at the data collection end of process tracing, knowing the facts and details of the case is essential. Expert knowledge provides access to potentially useful pieces of evidence—CPOs—that simply may not be available to those less familiar with the case (e.g., Tansey 2007). To excel at designing good process tracing tests, by contrast, one must possess knowledge of relevant existing generalizations, perhaps established from analyses of other cases. These generalizations need not be ''universal laws,'' but they ideally will specify necessary and/or sufficient conditions whose scope conditions encompass the specific case under analysis.

By way of conclusion, it is useful to consider four questions that are often essential to the explanation of a specific outcome in a particular case. The questions suggest a series of steps for carrying out or evaluating process tracing tests.

(1) Did the outcome to be explained actually occur? Ideally, of course, one starts with a well conceptualized and carefully defined outcome. The challenge then is to demonstrate rather than assume that this outcome occurred in the specific case under analysis. With process tracing, one can try to show how a hypothesis positing the existence of the outcome can pass one or more smoking gun tests. If indeed the existence of the outcome is ''obvious,'' then the researcher should have little trouble showing how a hypothesis positing its existence easily passes smoking gun tests.

If the hypothesis cannot pass a smoking gun test, the analyst must work to establish its existence using hoop tests and straw in the wind tests. While passing an easy hoop test will not increase much one's confidence that the hypothesis is valid, the ability of a hypothesis to pass a difficult hoop test will enhance the estimated likelihood that the outcome occurred. Likewise, if the hypothesis passes several hoop tests or several different straw in the wind tests, one's confidence that outcome has occurred often will be increased. The ability to pass several difficult hoop tests can come close to approximating smoking gun support for a hypothesis.

(2) Did the causal factors hypothesized to explain this outcome actually occur? The same issues apply to each event or process that is hypothesized to have caused the outcome. These causal factors must be well conceptualized, and the analyst ideally should be able to demonstrate how one or more process tracing tests clearly establish their existence.

(3) Did the posited cause(s) actually cause the outcome? Exactly how difficult it is to answer this question depends on the specific hypothesis under consideration. With some hypotheses, the posited cause may be obviously associated with the outcome, such that the main challenge of explanation involves showing that both the cause and outcome were present. But in many other instances, the key issue when trying to explain an outcome in a specific case is whether a known antecedent event really acted as a cause.

To make these causal inferences, process tracing tests require the analyst to explore the mechanisms linking the antecedent event to the outcome. If the antecedent event is posited to have been necessary for the outcome, a smoking gun test will normally involve linking this event to one or more mechanisms that can be shown to have been necessary for the outcome. If the antecedent is a necessary condition for any mechanism that is known to be necessary for the outcome, it follows logically that the event itself was also necessary for the outcome.

Hoop tests and straw in the wind tests can also help the analyst infer a causal linkage through the identification of one or more mechanisms. To make a strong inference, the analyst can work to show that the hypothesis passes one or more difficult hoop tests. Likewise, the analyst has some basis for inferring causation when the hypothesis passes a straw in the wind test that approximates a smoking gun test. Even relatively weak straw in the wind tests can add up to strong support when they are taken together.

(4) How are rival hypotheses eliminated? For many outcomes, there are multiple competing hypotheses to explain their occurrence. A good explanation of an outcome not only builds a positive case for the main hypothesis of interest but also reveals problems with any rival hypotheses that represent alternative and contradicting explanations of that outcome.

Rival hypotheses can be eliminated with process tracing tests in several different ways. One possibility is to show that the causal event posited by the hypothesis did not exist. If a hypothesis positing the existence of given cause fails a hoop test, the implication is that the cause did not take place, such that the hypothesis cannot possibly be right.

Alternatively, a rival hypothesis can be ruled out by using information concerning causal mechanisms. With a hoop test, the researcher identifies one or more mechanisms that must be present for the hypothesis to be valid. If the mechanism(s) are not present, the hypothesis fails the hoop test and can be eliminated. It is also possible to cast doubt on a rival hypothesis by demonstrating that it fails to pass an easy smoking gun test. Along the same lines, failing a straw in the wind test counts against a rival hypothesis. This is especially true if the straw in the wind test comes close to approximating a hoop test.

These four questions provide a basis for evaluating explanations of outcomes in specific cases. How well an analyst explains an outcome depends significantly on how well he or she can answer these questions. In turn, the quality of these answers depends heavily on the strength of the process tracing tests carried out by the researcher.

Appendix A
Examples Of Process Tracing Tests For Causal Inference: Explaining Democracy In England

In this appendix, I briefly discuss three important hypotheses embedded within major explanations of democratization in England. Each of the examples involves the use of an (implicit) process tracing test with a mechanism to help make a causal inference. Although all of the examples focus on democratization in England, they concern explanations that do not directly compete with or contradict one another.

1. Moore on the Enclosure Movement

In *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World* (1966), Barrington Moore famously hypothesizes that the enclosure movement (between 1760 and 1832) in England was a necessary cause of the parliamentary democracy that developed in this country after the mid-nineteenth century (pp. 20-39). The core support for the argument involves a *smoking gun test* concerning an intervening mechanism that links the enclosure movement and the construction of parliamentary democracy. This key intervening variable is the destruction of traditional peasant society.

Moore uses cross-case comparisons to establish the generalization that democratization could not have gone forward during this world historical period in the midst of a powerful traditional peasantry. Whenever such a peasantry was present, the result was reactionary fascism (e.g., Germany, Japan), communism (e.g., Russia, China), or premodern backwardness (e.g., India). Thus, Moore contends that the removal or historical absence of a traditional peasantry is necessary for a democratic path to the modern world.

In turn, Moore uses the historical literature to establish the fact that, in England, the enclosure movement was essential to the process through which the traditional peasantry was removed. Without the enclosures, he argues, the smallholding peasant villages would have persisted.

Thus, the enclosure movement was necessary for the destruction of the traditional peasantry, which was necessary for parliamentary democracy. For Moore, these facts add up to smoking gun proof that the enclosure movement was necessary for England's democracy. Moreover, since the enclosure movement was an unusual event in the early modern world (Moore treats it as perhaps the main factor that sets England apart from the rest of the world), one can regard it as a very important necessary condition for England's democratic path.

2. Luebbert on Lib-Labism

In *Liberalism, Fascism, or Social Democracy: Social Classes and the Political Origins of Regimes in Interwar Europe* (1991), Gregory Luebbert argues that an alliance between liberal parties and the labor movement (i.e., a Lib-Lab alliance) before World War I was nearly sufficient (in the historical context of Europe) to ensure that England would develop a liberal democracy during the interwar period. One key component of this argument involves a *hoop test* with an intervening mechanism: a moderate labor movement. This mechanism is established to be necessary for liberal democracy. In turn, if the hypothesis is correct, lib-labism must be sufficient for this mechanism. A sufficient cause of an outcome must be sufficient for all intervening mechanisms that are necessary for the outcome.

In this test, Luebbert (implicitly) employs the generalization that a moderate labor movement was necessary for liberal democracy during the interwar period in Europe. The generalization derives from an analysis of other European cases where strong labor movements were present. In these cases, ''labor peace and discipline, and by correlate the stability of the political order . . . would require a fundamental break with the liberal model'' (p. 10). Only if labor was docile and without class consciousness could liberal democracy prevail.

Given this generalization, the hoop test involves showing that a Lib-Lab alliance was sufficient for a moderate labor movement (in the context of England). Luebbert summarizes the evidence for this conclusion as follows: ''The message to be drawn from the historical material is unambiguous: very few British workers had much use for class politics precisely because Lib-Labism gave them a measure of confidence, however small, that was sufficient to undermine a more comprehensive vision'' (p. 25).

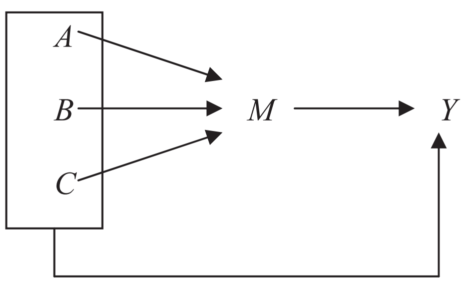
It is important to recognize that the form of this argument is a hoop test, not a smoking gun test. Thus, the ability of the hypothesis to pass the test does not provide decisive evidence that the hypothesis is correct (though failing the test would have eliminated it). Nevertheless, the test is convincing in that it draws on a generalization that Luebbert rigorously develops through systematic cross-case comparisons elsewhere in the book. It provides good straw in the wind support for the claim that lib-labism was sufficient (in the context of England) for a liberal democratic regime outcome during the interwar period.

3. Downing on Medieval Constitutionalism

In *The Military Revolution and Political Change: Origins of Democracy and Autocracy in Early Modern Europe* (1992), Brian Downing argues that the simultaneous presence of three conditions during the medieval period in England (and in Europe more generally) were necessary for the early emergence and consolidation of liberal democracy: (*A*) a rough balance of power between crown and nobility, (*B*) decentralized military systems, and (*C*) peasant property rights (pp. 19-26). He argues that if any one of these conditions had been missing, England could not have embarked on its early path toward democracy.

This argument is developed by showing how the conjunction of these three conditions was necessary for a more proximate mechanism: western constitutionalism. Downing defines western constitutionalism by the presence of local government, parliamentary bodies, and the rule of law. Although western constitutionalism was not sufficient for early democracy, it provided an essential ingredient for that outcome. Downing develops this generalization by contrasting Western Europe to various other regions of the world where democracy did not develop.

The structure of Downing's argument is thus as follows:



Here A is balance of power between crown and nobility; B is decentralized military systems; C is peasant property rights; M is western constitutionalism; Y is liberal democracy; and $\rightarrow$ stands for a hypothesized *necessary condition* relationship.

The logic of the process tracing test involves the fact that Downing is able to make a strong claim that M is necessary for Y using cross-case evidence. In turn, given this generalization, he works to convince us that A, B, and C are each necessary for M. He is aided in this effort by the fact that A, B, and C are temporally proximate to M, such that it easier to show that the counterfactual absence of any one of them would have eliminated M. Once he has persuaded us that A, B, and C are necessary for M, he then logically reasons that these three factors *must* also be necessary for Y, given that M is established to be necessary for Y.

Overall, then, the form of the test parallels the Moore example above. Downing has carried out a smoking gun test of the hypothesis that three conditions in the medieval period were necessary for liberal democracy in England.

Acknowledgments

For helpful comments on previous drafts, I would like thank Jennifer Cyr, Colin Elman, John Gerring, Gary Goertz, Erin Kimball, Matthew Lange, Isaac Reed, Carsten Schneider David Waldner, Christopher Winship, and the anonymous reviewers at *SMR*. I am indebted to Andrew Bennett and David Collier for many conversations about the ideas in this article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Process tracing can also be used for the purposes of hypothesis formulation and theory development. My focus in this article, however, concerns process tracing as a method of hypothesis testing. I am specifically concerned with specifying the logic of process tracing tests.
2. Or, as Van Evera (1997:31) puts it, ‘‘passage of the [hoop] test still leaves the theory in limbo.’’
3. Another kind of test discussed in the literature is a *doubly decisive test* (Van Evera 1997:32). With this test, the analyst looks for a causal-process observation (CPO) that decisively supports a particular explanation and eliminates all other explanations. The CPO is not only sufficient for the validity of a given explanation but also sufficient for the invalidity of all other explanations. Yet doubly decisive tests require the analyst to apply separate process tracing tests to each individual explanation. While a single CPO may be used with these tests, it is employed in conjunction with different generalizations in the

assessment of each individual explanation. Accordingly, with doubly decisive tests, separate process tracing tests must be carried out for each explanation.

4. A focus on necessary and sufficient conditions is not inconsistent with proposals that link process tracing to a Bayesian-like mode of inference (Bennett 2006, 2008; see also McKeown 1999). Scholars who view process tracing in Bayesian-like terms assume that researchers derive their conclusions from the use of process tracing tests. Understanding the logic of these process tracing tests is independent of the question of whether process tracing should be viewed as roughly analogous to Bayesian inference.
5. Even when a generalization is valid, problems can arise in the effort to connect the facts of a specific case to the generalization. For these and other reasons, the formulation of strong process tracing tests often depends on the creativity and imagination of the researcher.
6. Important necessary conditions come close to being sufficient conditions (Goertz 2006; Mahoney 2008; Ragin 2008). This is why passing a hoop test that involves an important necessary condition provides positive support in favor of the hypothesis.
7. Philosophers sometimes note that present traces overdetermine the existence of a past event, by which they mean that several present-day traces are each individually sufficient to infer the existence of a past event (Cleland 2002; Lewis 1986).
8. Here it is perhaps worth noting that noncausal correlations can be used with straw in the wind tests when inferring the descriptive existence of a cause or outcome. The presence of factors strongly correlated with an unobserved event make it more likely (subjectively) that the event occurred even if these factors are not causally related to the event.
9. On the inherent link between necessary condition hypotheses and counterfactuals, see Lewis (1986) and Goertz and Starr (2003).
10. In the social sciences, there is a well developed methodology for working with conditions that are probabilistically necessary and probabilistically sufficient for outcomes (e.g., Ragin 2000, 2008).
11. This aspect of qualitative research might be seen as contrasting with quantitative research, where investigators are sometimes encouraged to design a statistical model before the analysis of any data. In practice, though, various refinements to statistical models in light of the data routinely occur in quantitative research as well (Collier, Brady, and Seawright 2010).

References

Abbott, Andrew. 1992. ‘‘From Causes to Events: Notes on Narrative Positivism.’’ *Sociological Methods and Research* 20:428-55.

Bennett, Andrew. 2006. ‘‘Stirring the Frequentist Pot With a Dash Bayes.’’ *Political Analysis* 14:339-44.

Bennett, Andrew. 2008. ''Process Tracing: A Bayesian Perspective.'' Pp. 217-70 in *The Oxford Handbook of Political Methodology*, edited by J. Box-Steffensmeier, H. E. Brady, and D. Collier. Oxford: Oxford University Press.

Bennett, Andrew and Colin Elman. 2006. ''Recent Developments in Case Study Methods.'' *Annual Review of Political Science* 9:455-76.

Braumoeller, Bear F. and Gary Goertz. 2000. ''The Methodology of Necessary Conditions.'' *American Journal of Political Science* 44:844-58.

Cleland, Carol E. 2002. ''Differences Between Historical Science and Experimental Science.'' *Philosophy of Science* 69:474-96.

Collier, David. 1993. ''The Comparative Method.'' Pp. 105-19 in *Political Science: The State of the Discipline II*, edited by A. Finifter. Washington, DC: American Political Science Association.

Collier, David. 2011. ''Understanding Process Tracing.'' *PS: Political Science and Politics* 44:823-30.

Collier, David, Henry E. Brady, and Jason Seawright. 2010. ''Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology.'' Pp. 161-99 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. 2nd ed., edited by H. E. Brady and D. Collier. Lanham: Rowman and Littlefield.

Costner, Herbert L. and Robert K. Leik. 1964. ''Deductions from 'Axiomatic Theory.''' *American Sociological Review* 29:819-35.

Downing, Brian M. 1992. *The Military Revolution and Political Change: Origins of Democracy and Autocracy in Early Modern Europe*. Princeton: Princeton University Press.

Drèze, Jean and Amartya Sen. 1989. *Hunger and Public Action*. Oxford: Oxford University Press.

George, Alexander L. 1979. ''Case Studies and Theory Development: The Method of Structured, Focused Comparison.'' Pp. 43-68 in *Diplomacy: New Approaches in History, Theory, and Policy*, edited by P. G. Lauren. New York: The Free Press.

George, Alexander L. and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

George, Alexander L. and Timothy J. McKeown. 1985. ''Case Studies and Theories of Organizational Decision Making.'' Pp. 29-41 in *Advances in Information Processing in Organizations*. Vol. 2. Santa Barbara, CA: JAI Press.

Gerring, John. 2010. ''Causal Mechanisms: Yes, but . . . '' *Comparative Political Studies* 43:1499-526.

Goertz, Gary. 2003. ''Cause, Correlation, and Necessary Conditions.'' Pp. 47-64 in *Necessary Conditions: Theory, Methodology, and Applications*, edited by G. Goertz and H. Starr. Lanham: Rowman and Littlefield.

Goertz, Gary. 2006. ''Assessing the Trivialness, Relevance, and Relative Importance of Necessary or Sufficient Conditions in Social Science.'' *Studies in Comparative International Development* 41:88-109.

Goertz, Gary and James Mahoney. 2005. ‘‘Two-level Theories and Fuzzy-set Analysis.’’ *Sociological Methods and Research* 33:497-538.

Goertz, Gary and Harvey Starr. 2003. ‘‘Introduction: Necessary Condition Logics, Research Design, and Theory.’’ Pp. 1-24 in *Necessary Conditions: Theory, Methodology, and Applications*, edited by G. Goertz and H. Starr. Lanham: Rowman and Littlefield.

Goldstone, Jack A. 1998. ‘‘Initial Conditions, General Laws, Path Dependence, and Explanation in Historical Sociology.’’ *American Journal of Sociology* 104:829-45.

Griffin, Larry J. 1993. ‘‘Narrative, Event-structure, and Causal Interpretation in Historical Sociology.’’ *American Journal of Sociology* 98:1094-133.

Hall, Peter A. 2006. ‘‘Systematic Process Analysis: When and How to Use It.’’ *European Management Review* 3:24-31.

Harré, Rom. 1970. *The Principles of Scientific Thinking*. Chicago: University of Chicago Press.

Hart, H. L. A. and Tony Honoré. 1985. *Causation in the Law*. 2nd ed. Oxford: Oxford University Press.

Hempel, Carl G. [1942] 1965. *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. New York: Free Press.

Lewis, David. 1986. *Philosophical Papers*. Oxford: Oxford University Press.

Luebbert, Gregory M. 1991. *Liberalism, Fascism, or Social Democracy: Social Classes and the Political Origins of Regimes in Interwar Europe*. New York: Oxford University Press.

Mahoney, James. 2008. ‘‘Toward a Unified Theory of Causality.’’ *Comparative Political Studies* 41:412-36.

Mahoney, James, Erin Kimball, and Kendra Koivu. 2009. ‘‘The Logic of Historical Explanation in the Social Sciences.’’ *Comparative Political Studies* 42:114-46.

Mayntz, Renate. 2004. ‘‘Mechanisms in the Analysis of Social Macro-phenomena.’’ *Philosophy in the Social Sciences* 34:237-54.

McKeown, Timothy J. 1999. ‘‘Case Studies and the Statistical Worldview: Review of King, Keohane, and Verba’s *Designing Social Inquiry*.’’ *International Organization* 53:161-90.

McMullin, Ernan. 1984. ‘‘Two Ideals of Explanation in Natural Science.’’ *Midwest Studies in Philosophy* 9:205-20.

Moore, Barrington Jr. 1966. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.

Nagel, Ernst. 1961. *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace, and World.

Popper, Karl. [1934] 1968. *The Logic of Scientific Discovery*. New York: Harper and Row.

Ragin, Charles C. 2000. *Fuzzy-set Social Science*. Chicago: University of Chicago Press.

Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.
Railton, Peter. 1978. ''A Deductive-nomological Model of Probabilistic Explanation.'' *Philosophy of Science* 45:206-26.
Roberts, Clayton. 1996. *The Logic of Historical Explanation*. University Park: Penn State University Press.
Rueschemeyer, Dietrich. 2003. ''Can One or a Few Cases Yield Theoretical Gains?'' Pp. 305-36 in *Comparative Historical Analysis in the Social Sciences*, edited by J. Mahoney and D. Rueschemeyer. New York: Cambridge University Press.
Salmon, Wesley. 1990. *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press.
Scriven, Michael. 1959. ''Truisms as the Grounds for Historical Explanations.'' Pp. 443-75 in *Theories of History*, edited by P. Gardiner. Glencoe: Free Press.
Skocpol, Theda. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.
Tansey, Oisín. 2007. ''Process Tracing and Elite Interviewing: A Case for Non-probability Sampling.'' *PS: Political Science and Politics* 40:765-72.
Van Evera, Stephen. 1997. *Guide to Methods for Students of Political Science*. Ithaca, NY: Cornell University Press.
Waldner, David. 2007. ''Transforming Inferences into Explanations: Lessons from the Study of Mass Extinctions.'' Pp. 145-176 in *Theory and Evidence in Comparative Politics and International Relations*, edited by R. N. Lebow and M. I. Lichback. New York: Palgrave-MacMillan.

Bio

James Mahoney is Fitzgerald Professor of Economic History in the Departments of Political Science and Sociology at Northwestern University.