

Process Tracing and the Problem of Evaluating Causal Relationships Within a Case

Philosophy of the Social Sciences

2026, Vol. 0(0) 1–36

© The Author(s) 2026



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/00483931261441189

journals.sagepub.com/home/posYafeng Wang¹ 

Abstract

Process tracing method in the social sciences seek to assess hypothesized causal mechanisms in individual cases, but practitioners face a problem: How to use within-case evidence to evaluate singular causal relationships within hypothesized mechanism. This paper presents a partial solution to this problem by treating detailed features of the outcomes as observable consequences of the hypothesized causal relationships, given suitable auxiliary assumptions. This approach is then illustrated through analysis of a process-tracing study by Vesla Weaver. The paper concludes by examining how the feature-based approach relates to other frameworks for assessing singular causal claims.

Keywords

process tracing, singular causation, causal inference, causal mechanism, features

Received: 9 January 2026

¹Northeastern University, Boston, MA, USA

Corresponding author:

Yafeng Wang, Department of Philosophy and Religion, Northeastern University, 360 Huntington Avenue, Boston, MA 02115, USA.

Email: ya.wang@northeastern.edu

Introduction

Why did the Great Recession happen? Why did the Cold War end peacefully? Why did the U.S. incarceration rate increase so dramatically in the last four decades? Puzzles such as these demand causal explanations, and many social scientists are converging to the view that the type of causal explanations needed for these puzzles includes references to hypothesized causal mechanisms. But such a view raises further questions about what causal mechanisms are, and how competing causal explanations referencing different causal mechanisms can be evaluated using empirical observations.

In recent decades, a method for testing causal mechanisms in case studies called “process tracing” has been frequently discussed in the social science methodology literature. Proponents of this method argue that it is a “qualitative” causal inference method different from quantitative methods based on the statistical analysis of group comparison studies, and that it works by drawing detailed observations from individual cases to support or overturn a hypothesized causal mechanism (Beach and Pedersen 2013 [2019]; Bennett and Checkel 2015; Mahoney 2010).

One serious problem that the standard process tracing methodology faces concerns how observations made within an actual case can help evaluate singular causal claims postulated by a hypothetical causal mechanism, which seems to require counterfactual evidence (Runhardt 2015). I call this problem “the problem of evaluating causal relationships within a case.”

In this paper, I offer a partial solution by arguing that singular causal relationships leave observable traces in the features of outcomes. While different causes might produce outcomes of the same general type, the outcomes they produce typically have different characteristic features. Process tracers can assess singular causal claims by testing whether observed outcome features match what different causal hypotheses predict. This is a partial solution because it focuses on outcome features as one important type of within-case evidence, without claiming this exhausts all possible methods for assessing singular causal relationships.

This paper is organized as follows. First, I describe how process tracers conceptualize “causal mechanisms,” and how they are expected to use within-case evidence to evaluate whether an instance of a causal mechanism is present in a case. Second, I describe the problem that it is unclear how to use within-case observations to evaluate those singular causal relationships belonging to a causal mechanism. Third, I provide a partial solution to this problem based on the following idea: given suitable auxiliary assumptions, expected features of the outcome can be regarded as observable consequences of hypothesized singular causal relationships. Fourth, I use a process tracing

study by sociologist Vesla Weaver to illustrate how the features of an outcome can help evaluate singular causal relationships involving that outcome. Finally, I examine alternative frameworks for assessing singular causal claims and explore how the feature-based approach complements or contrasts with these alternatives. In doing so, I also address questions about the assumptions process tracers make regarding the nature of social phenomena and the role of social meanings and interpretation in causal explanation.

Background

Process tracing is a research method commonly used in social science case study research. A case study is an intensive study of a single case that draws on within-case evidence. A case is a temporally and spatially bounded instance of a type of phenomenon (Bennett and Checkel 2015; Gerring 2007 [2017]).¹ For instance, the French Revolution is a case of a revolution. Within-case evidence is evidence from within the boundaries that define the case itself, whether those boundaries are temporal, spatial, or the specific phenomenon being studied (Bennett and Checkel 2015, 8). As a case study method, process tracing uses within-case evidence to trace causal mechanisms within a case. According to Beach and Pedersen (2013 [2019], 1–2), the “process” being traced is a causal mechanism that links a cause (or set of causes) with an outcome. The purposes of these inferences are “either developing or testing hypotheses about causal mechanisms that might causally explain the case” (Bennett and Checkel 2015, 7).

Process tracing serves multiple analytical purposes, one of which is theory testing (Beach and Pedersen 2013 [2019]; Bennett and Checkel 2015). To test theories, researchers examine whether the causal mechanisms proposed by existing theories operate as expected within specific cases (Bennett 2008; Goertz and Mahoney 2012). Consider the following research context: Statistical analyses have shown correlations between variables (e.g., factors *X* and *Y*) across a large population of cases, and researchers draw from a theory to explain these correlations by proposing specific causal mechanisms that might connect *X* and *Y*. In this context, process tracing allows researchers to test the hypothesized causal mechanisms by examining their operation in carefully selected individual cases.² If the test results indicate that hypothetical mechanisms operate as the theory suggests in the selected cases, they support the theory. If the test results indicate otherwise, they suggest that the theory may need modification or that alternative causal mechanisms were responsible for the observed outcomes.

Process tracing can also be used for theory building when researchers lack clear theoretical guidance about potential causal mechanisms. This variant is

often used when researchers know (or suspect) a correlation exists between cause and outcome but are uncertain about the mechanisms connecting them. It starts with empirical material and works inductively to build a plausible causal mechanism. The process involves casting a wide net to gather empirical material before knowing exactly what causal story it tells, then looking for patterns that suggest how a particular causal mechanism operated (Beach and Pedersen 2013 [2019], 11). The goal is to build up a plausible mechanism that can potentially be generalized beyond the single case (Beach and Pedersen 2013 [2019], 10).

Another use case of process tracing provides a comprehensive causal account of a specific historical outcome. Beach and Pedersen (2013 [2019], 11) call this variant “explaining outcome” process tracing. Suppose researchers are investigating a case with an interesting and puzzling outcome. The outcome is multifaceted and has many relevant features that call for explanation. Explaining-outcome process tracing can be used to argue that a causal mechanism is present in the case, that it can account for many important features of the outcome (or it can account for them better than alternative accounts).

Two questions arise from the preceding discussions. First, how do process tracers conceptualize “causal mechanisms”? Second, how do process tracers use evidence to evaluate whether (an instance of) a causal mechanism is present in a case? To address these questions, I survey existing accounts of process tracing and arrive at two claims. First, despite varying definitions of causal mechanisms in the literature, process tracers generally agree that a causal mechanism instance within a case includes chains of intermediate, singular causal links connecting some causal factor to an outcome. Second, process tracers also tend to agree that the main form of evidence used to evaluate whether a causal mechanism instance is present in a case is within-case evidence, or more specifically, the correspondence between observable implications of hypothetical causal mechanisms and actual observations within the case.

The concept of causal mechanisms lies at the heart of process tracing methodology, yet the social science literature on process tracing lacks a consistent definition for this concept. To take just a few examples: Bennett and Checkel (2015, 12) define causal mechanisms as “ultimately unobservable physical, social, or psychological processes through which agents with causal capacities operate, but only in specific contexts or conditions, to transfer energy, information, or matter to other entities.” Goertz and Mahoney (2012, 100) define them as “the intervening processes through which causes exert their effects.” Beach and Pedersen (2013 [2019], 38) define mechanisms as

“systems of interlocking parts that transmit powers or forces between a cause (or a set of causes) to an outcome.”

Despite these definitional differences, [Clarke \(2023, 308\)](#) argues that process tracers at least agree on one point: “Namely, to describe a mechanism one must at very least describe a chain of causes that end in the outcome in question: A causes B, B causes C, C causes D, D causes E.” I agree with Clarke’s assessment, and I will support it by examining two conceptions of causal mechanisms discussed in [Beach and Pedersen \(2013 \[2019\]\)](#): The minimalist conception and the systems conception ([Beach and Pedersen 2013 \[2019\]](#), 3).

The minimalist conception treats causal mechanisms as intervening factors between a cause and the outcome, distinguished by their temporal position. In the representation $X \rightarrow M \rightarrow Y$, the letters refer to events or specific values on variables, with X being the cause, M the causal mechanism, and Y the outcome ([Beach and Pedersen 2013 \[2019\]](#), 35–36). Sometimes the causal mechanism M is labeled an “intervening variable” ([Waldner 2015](#), 131). However, if M is an instance of a causal mechanism within a single case, a more accurate description of M would be “particular events or specific values on variables” ([Mahoney 2015](#), 206) that stand between a cause and outcome in time. Moreover, the minimalist conception does not unpack the intervening factor M in detail, describing it instead using simple terms or one-liners without explaining how these mechanisms work ([Beach 2021](#), 8905; [Beach and Pedersen 2013 \[2019\]](#), 35).

[Beach and Pedersen \(2013 \[2019\]\)](#) used Nina Tannenwald’s article on nuclear taboos and U.S. nuclear weapons decision-making as an example ([Tannenwald 1999](#)). According to them, Tannenwald argues that there is a causal mechanism connecting a cause (norms or “taboos” against the use of nuclear weapons) and an outcome (U.S. decision maker’s avoidance of using these weapons), and yet the mechanism is only described briefly using one-liners such as: “constraints imposed by individual decision makers’ personal moral convictions or domestic or world opinion” ([Beach and Pedersen 2013 \[2019\]](#), 36). In their view, such superficial descriptions of the intervening causal mechanism M are not sufficiently detailed to help us understand how the existence of the unclear taboo imposes constraints that shape the behaviors of decision makers ([Beach and Pedersen 2013 \[2019\]](#), 36).

Because of the limitations of the minimalist conception, [Beach and Pedersen \(2013 \[2019\]](#), 37–38) prefer “the systems conception” of (causal) mechanisms, which defines mechanisms as “systems of interlocking parts that transmit powers or forces between a cause (or a set of causes) to an outcome.” [Beach and Pedersen \(2013 \[2019\]\)](#) further suggest that each of the parts of the mechanism should be described in terms of entities that engage in activities.

Entities in the social sciences include micro-level or macro-level actors, which have properties and capacities that enable them to engage in activities that can impact other actors in the process. Activities are what social actors do, and can take the form of speech acts, voting, paying bribes, etc. (Beach 2021, 8909).

Using graphical notations, the systems conception of causal mechanisms can be represented as:

$$X \rightarrow \text{Part 1 (e1 * a1)} \rightarrow \text{Part 2 (e2 * a2)} \rightarrow \dots \rightarrow Y$$

with X being a cause, Y being the outcome, “e” being entities, and “a” being activities. This representation can be read as the cause triggers a mechanism that moves through part 1, which is composed of entity 1 engaging in activity 1, and so forth, until the outcome (Beach and Pedersen 2013 [2019], 70). Importantly, the arrows in the directed graph are meant to represent causal relationships. Unpacking causal mechanisms into causally connected parts is useful because it helps researchers understand how the mechanism works and find within-case evidence for each part of the process.

Consider a toy example from Punton and Welle (2015). Suppose we want to understand how the universal health care campaign in Ghana (X) causally contributed to the availability of free universal health care in Ghana (Y). A possible causal mechanism connecting X and Y could be the following: the universal health care campaign (X) caused civil society (entity 1) to conduct coordinated advocacy for free universal healthcare (activity 1), which causes the public (entity 2) to become aware of the limitations of current health care financing (activity 2), which cause the public (entity 2) to demand free universal health care from government actors (activity 3), which causes government actors (entity 3) to increasingly support free universal healthcare (activity 4), which causes government (entity 4) to amend policies and processes to move toward free universal health care (activity 5), which finally causes the availability of free universal healthcare (Y) (Punton and Welle 2015, 3). Even though this is a simplistic example and is not meant to fully capture how Ghana came to have free universal health care, the structure of the example does help illustrate how the systems approach conceptualizes causal mechanisms.

In short, the main difference between the minimalist conception and the systems conception of causal mechanisms is that the latter unpacks the causal mechanism M mediating between the cause X and the outcome Y as a series of intermediate stages, each of which consists of entities engaging in various activities, and each stage is connected to the preceding and the succeeding stages via causal relationships. The systems conception of causal mechanisms is explicitly committed to specifying chains of intermediate causal

relationships. To the extent that the minimalist conception conceptualizes the mediating mechanism *M* as causal, it is at least implicitly committed to some intermediate causal relationships connecting the cause *X* and the outcome *Y*. Moreover, on either account, when specifying an instance of a causal mechanism within a single case, the intermediate causal relationships are singular causal relationships.

Moving on to the second question: How do process tracers use evidence to evaluate whether an instance of a causal mechanism is present within a case? Standard accounts of process tracing in the social science literature suggest that the main form of evidence used to test causal mechanisms in a case is within-case evidence: That is, evidence from within the spatial, temporal, and topical boundary of the case.

Process tracers argue that a causal mechanism instance within a case has observable implications, which are alternatively described as “empirical fingerprints” or “traces.” The basic idea is that if each part of a hypothesized causal mechanism were present within a case, the activities associated with the operation of the mechanism would leave behind observable traces, similar to how activities associated with a crime would leave behind traces in a crime scene (Beach and Pedersen 2013 [2019], 4; Bennett 2008, 705; Bennett and Checkel 2015, 30). In process tracing research, process tracers could operationalize their hypothetical causal mechanism by identifying what observable evidence would indicate its presence in a given case. They then collect within-case evidence to determine whether the predicted observable evidence is present in the case (Beach and Pedersen 2013 [2019], 9–10). When the predicted evidence is found, process tracers can reasonably infer that the hypothesized mechanism operated as expected.

In summary, regarding the second question, process tracers in the social sciences characterize the main form of evidence they rely on in terms of the correspondence between the observable implications that might have been left by the operation of a hypothetical mechanism and the actual empirical observations found in a case.

A Problem

So far, I have argued that process tracing helps evaluate whether an instance of a causal mechanism (which contains intermediate, singular causal relationships) is present within a given case. Moreover, according to the standard accounts in the social science literature, process tracing is supposed to rely heavily on within-case evidence, e.g., observable traces left behind by the operation of each part of the causal mechanism. In this section, I raise a problem for process tracing, namely, it is unclear how to use within-case

evidence to evaluate singular causal relationships within a case. Variants of this problem have been raised in the philosophical literature, particularly by [Runhardt \(2015\)](#) and [Clarke \(2023\)](#). The problem poses a challenge to process tracers because, if unresolved, it suggests that process tracing fails to fulfill a key part of its promise.

An instance of a causal mechanism within a case consists of intermediate parts or stages connected by singular causal relationships. It is clear, at least in principle, how to use within-case evidence to evaluate the presence or absence of intermediate stages or events. To illustrate this, consider again the example of universal healthcare in Ghana given by [Punton and Welle \(2015\)](#). Two of the intermediate stages of the hypothetical causal mechanism are: The public demands free universal health care from government actors, and government actors increasingly support free universal health care. In principle, these events could leave behind observable traces: For instance, surveys and interviews conducted by researchers or journalists, records of town halls of political representatives, and records of the public and private statements of important government actors can all be used to evaluate whether these two hypothetical events have indeed occurred in the case of Ghana. Perhaps some of the records or traces are not in fact available, but at least in principle, it is not difficult to conceive how within-case evidence could be brought to bear on the presence or absence of these intermediate singular events.

In contrast, it appears more challenging to use within-case evidence to assess the presence or absence of intermediate, singular causal relationships. Singular causal relationships are often associated with counterfactual dependence; for instance, the singular causal relationship “C causes E” is commonly associated with the counterfactual “if C had not occurred, E would not have occurred.” How could process tracers bring observations from within an actual case to bear on the relevant counterfactual dependencies? What constitutes the “empirical fingerprints,” “traces,” or “observable implications” of singular causal relationships within an actual case? I call this problem “the problem of evaluating causal relationships within a case.” Variants of this problem have been raised in the philosophical literature on process tracing, notably by [Runhardt \(2015\)](#) and [Clarke \(2023\)](#). I suggest, however, that neither work provides a fully satisfactory resolution of this problem.

[Runhardt \(2015\)](#) argues that to test whether a causal mechanism exists between X and Y in individual cases, not only should process tracers specify the chains or networks of intermediate events, but they must investigate whether each link of the chains (or networks) of events is genuinely causal. Drawing on [Woodward \(2003\)](#)’s interventionist theory of causation, she contends that a link between two intermediate events Z_i and Z_j is genuinely causal only if the following condition holds: If a potential intervention had

prevented Zi from happening, then event Zj would not have occurred either (Runhardt 2015, 1297).

Runhardt (2015, 1297) further argues that standard methodological recommendations by process tracers are insufficient because they fail to test whether the hypothesized intermediate links are truly causal. She uses Bakke's (2013) study of the Second Chechen War as an example to illustrate this problem. Bakke (2013) establishes that schools and training camps were built between the arrival of transnational insurgents and the use of radical tactics such as suicide bombing, and she collects observable consequences of the presence of the intervening factors. However, Bakke does not look for evidence concerning what would happen if an intervention were made, e.g., if transnational insurgents were prevented from arriving, or they were prevented from building schools and training camps. Without the relevant counterfactual evidence, Runhardt (2015) argues, Bakke fails to show these camps actually caused increased use of radical tactics. In short, observing the deductive consequences (or "causal process observations") of mechanisms is not sufficient for showing that the links in the chains or networks of mediating events are genuinely causal (Runhardt 2015, 1305).

How could process tracers support or undermine the counterfactual dependencies relevant to singular causal claims, then? Runhardt (2015) suggests that while intervening on a factor Zi in a single case study is either impossible or undesirable, evidence of such interventions can come from comparisons with other, sufficiently similar case studies that lack Zi. She coins the term "natural interventions" for these cross-case comparisons (Runhardt 2021, 433), modeling it after the research method known as natural experiments. For instance, to apply Runhardt's methodological recommendation to Bakke's study, process tracers would need to find a set of conflicts that are sufficiently similar to the Second Chechen War, but where transnational insurgents are not present. Moreover, process tracers would need to specify what they mean by "sufficiently similar" (Runhardt 2015, 1305).

I argue, however, that the methodological recommendation given by Runhardt (2015) has two shortcomings. The first shortcoming is that it is practically difficult to find comparison cases that are sufficiently similar to qualify as natural interventions. Runhardt (2015) acknowledges that in social science case studies, most of the processes being traced are complex and include unique or idiosyncratic aspects (Runhardt 2015, 1306). Consequently, it is likely that one can always find causally relevant differences between the original case and the comparison cases, so that the latter fail to qualify as natural interventions.

The second shortcoming is that Runhardt (2015) seems to imply that the standard process tracing methodology³ is misguided. According to the

standard process tracing methodology, the primary strength of process tracing lies in its ability to extract and leverage rich and heterogeneous details from individual cases (Mahoney 2010, 124). This characteristic of process tracing differentiates it from “large-N” quantitative causal inference methods that rely on comparisons of a large number of cases. It also explains why the standard methodology focuses on providing guidance about how to identify observational implications of causal mechanisms within a case.

Runhardt (2015) notes that process tracing methodologists are not focused on finding evidence of (natural) interventions. Instead, they seek observable implications of causal mechanisms within a case. She concludes the standard process tracing methodology is “lacking” (Runhardt 2015, 1305). I interpret this as saying that the standard process tracing methodology is misguided if one accepts the interventionist theory of causation. Below is a line of reasoning that could be used to justify (my interpretation of) Runhardt’s conclusion:

- (1) Suppose we accept the interventionist theory of (singular) causation.⁴
- (2) If so, then a singular causal relationship between events X and Y is analyzed as an (interventionist) counterfactual: If X were absent (due to interventions), then Y would be absent.
- (3) If so, then the *only* type of evidence relevant to evaluating the singular causal relationship between X and Y consists of comparisons with other sufficiently similar cases where X is absent due to “natural” interventions.
- (4) Given that, it is hard to see how observations within a single case—in the absence of explicit cross-case comparisons—can be useful for evaluating singular causal relationships.
- (5) It follows that standard process tracing methodology is misguided in assuming that within-case observations can help evaluate singular causal relationships posited by hypothetical causal mechanisms.

The inference from steps 2 to 3 can be challenged, however. Perhaps comparisons with other sufficiently similar cases where X is absent (due to “natural” interventions) are *one* type of relevant evidence for evaluating the actual causal relationship between X and Y. But it might not be the *only* type of relevant evidence. Singular causal relationships might also have some observational implications within a single case. This would mean that *some* within-case observations are relevant to evaluate singular causal relationships. If so, then the standard process methodology would not be *misguided*, only *incomplete*. It could be made more complete by specifying what types of

observational details within a single case can be relevant to evaluating singular causal relationships.

In short, [Runhardt \(2015\)](#)'s contention—standard process tracing methodology is misguided if we accept the interventionist theory of causation—is questionable.⁵ The questions become: Can process tracers preserve the key insight of standard process tracing methodology—the capacity of process tracing to leverage rich details within individual cases—while developing tests of the intermediate causal links? Do singular causal relationships generate observable implications within cases? If so, what forms do these observable implications take?

[Clarke \(2023\)](#) raises a similar question about how process tracers are supposed to test intermediate causal links of a causal mechanism using within-case evidence, although his way of framing the question involves describing an assumption *not* shared by process tracing. [Clarke \(2023\)](#) points out that, unlike process tracing, quantitative research methods in political science typically require collecting comparable data from a large number of cases and organizing them into a rectangular dataset. Moreover, these quantitative methods rely on an assumption that Clarke calls “unit homogeneity” ([Clarke 2023](#), 317).

To illustrate this assumption, suppose researchers are interested in the extent to which natural resource wealth R and ethnic diversity D contributed to the duration Y of the civil war. What a quantitative researcher would do is to collect data on a larger population of cases; for example, all civil wars since 1945. According to Clarke, the researcher would further assume that (1) in each case, the duration of the civil war is a function of resource wealth and ethnic diversity; and (2) the *same* mathematical function that describes the dependence of the duration of a civil war on resource wealth and ethnic diversity is shared by all the cases in the dataset. That is, all the civil wars post 1945 exhibit “unit homogeneity” ([Clarke 2023](#), 318). This assumption of unit homogeneity, across a large population of cases, is essential to the reliability of the econometric method ([Clarke 2023](#), 319).

In contrast, process tracing does not rely on the assumption of unit homogeneity. Clarke states: “A study does not count as process tracing if it requires unit homogeneity for any of the intermediate links in the chain, that is, the existence of a largish population of cases for which there is a single mathematical function that describes the propensities governing factor D , for example, given variation in C and some other variables” ([Clarke 2023](#), 319). [Clarke \(2023\)](#) does not discuss in depth why he thinks process tracing does not require unit homogeneity, but his basic idea seems to be that there is a difference in the form of evidence relied upon by process tracing versus by quantitative methods. Quantitative methods typically rely on large,

rectangular datasets, and the assumption of unit homogeneity is essential for inference based on this type of data. Process tracing, in contrast, focuses on leveraging numerous heterogeneous, “noncomparable” observations from within individual cases, and the assumption of unit homogeneity is neither necessary nor sufficient for this type of evidence (Clarke 2023, 317).

Clarke then raises the question: “If an intermediate causal hypothesis is not to be tested via standard quantitative methods, then how else is it to be tested?” (Clarke 2023, 321). He then argues that existing recommendations made by the standard process tracing methodology do not help answer this question. For instance, the so-called “causal process observations” do not help answer the question about how intermediate causal hypotheses should be tested without assuming unit homogeneity. One interpretation of “causal process observations” of causal mechanism is that these are “diagnostic evidence” of causal mechanisms, but Clarke argues that all evidence regarding a hypothesis is diagnostic. Therefore, “diagnostic evidence” does not tell us what type of evidence specifically can be used to test intermediate causal links. Another interpretation of “causal process observations” is “pattern matching.” According to some schools of philosophy of science (e.g., hypothetical deductivism), however, all scientific inference can be viewed as pattern matching (i.e., hypothesis entails the patterns in the data). Therefore, this interpretation is again vacuous. On Clarke’s view, the most obvious interpretations of “causal process observations” are all vacuous and fail to tell us what is distinctive about process tracing as a method of causal inference (Clarke 2023, 322). He concludes that “what we urgently need is a taxonomy of various different ways in which process tracers are supposedly able to test intermediate causal hypotheses without relying on unit homogeneity” (Clarke 2023, 323).

In summary, both Runhardt (2015) and Clarke (2023) highlight a challenge: It is difficult to see how process tracers can test intermediate, singular causal relationships within a case using only within-case evidence. Both accounts emphasize that without either strong assumptions about cross-case similarity (e.g., natural interventions or unit homogeneity) or a clear approach to leveraging within-case observations, standard process tracing methodology appears to be incomplete. This raises a question crucial for the methodological validity of process tracing: Can within-case evidence meaningfully test intermediate causal relationships, and if so, how?

A Partial Solution

In this section, I present a partial solution to the problem of evaluating causal relationships within individual cases. I argue that, given suitable auxiliary

assumptions, hypothesized singular causal relationships do have “empirical fingerprints” or observable implications within a case. Process tracers can use the alignment or misalignment between these observable implications and within-case observations to test the proposed singular causal relationships. My key claim is that one type of observable consequences of singular causal relationships in a case are detailed characteristics or “features” of the outcome. This answer represents a partial rather than complete solution to the problem, because it leaves open the possibility that singular causal relationships may have additional types of within-case observable implications beyond outcome features. However, this partial solution shows that it is methodologically feasible to use within-case observations to test hypothesized singular causal relationships, and it identifies a relevant type of observations for such tests.

The causal relata of singular causal relationships are specific, particular events.⁶ Unlike abstract event types, particular events have many detailed characteristics, which I call “features.” Social science researchers are often interested in large-scale, multifaceted events that possess numerous features. For instance, the 2008 financial crisis, the Arab Spring, and the Brexit referendum of 2016 are all complex events with abundant and diverse features.

A familiar type of feature is temporal characteristics of an event. Temporal features of an include:

- (1) The duration of the event, that is, how much time lapses between the start and the end of the event (Grzymala-Busse 2011, 1277). An example of duration is how much time lapses before a new state institution is established.
- (2) The tempo of an event, which refers to the frequency of the “sub-events” in a larger event. Examples include the number of bills issued by the legislature per session, or the number of deaths per year of civil war or epidemic (Grzymala-Busse 2011, 1282).
- (3) The acceleration of an event, which is the rate of change in tempo. Many political events and processes exhibit changing tempos: They speed up and slow down at given points (Grzymala-Busse 2011, 1286).
- (4) The timing of an event, which consists of the placement of a given event on a timeline and in a larger context (Grzymala-Busse 2011, 1288).

Temporal features of outcomes can serve as diagnostic evidence for identifying contributing causal mechanisms. Grzymala-Busse (2011, 1272) demonstrates this logic by arguing that the tempo of institutional development can provide information about its causes: Rapid institutional building suggests

elite-driven mechanisms rather than popular deliberation, since mass consultation and negotiation requires an extended timeframe that would preclude swift implementation.

Bennett's (2008) analysis of the peaceful end of the Cold War also uses temporal features of the outcome. He surveys three prominent explanations for the nonuse of force in 1989: the realist explanation, which emphasizes the changing material balance of power; the domestic politics explanation, which focuses on the changing nature of the Soviet Union's ruling coalition; and the ideational explanation, which highlights the lessons Soviet leaders drew from their recent unsuccessful military intervention in Afghanistan (Bennett 2008, 715). Bennett (2008) argues that the specific chronology of Soviet foreign policy shifts supports the ideational explanation, because the timing of the policy change aligns with when Soviet leaders would have processed and internalized lessons from their unsuccessful military interventions in Afghanistan and other contexts (Bennett 2008, 716).

The form of evidence based on outcome features can be generalized as follows:

General format: "If a hypothesis regarding a singular causal relationship between factor X and outcome Y is true, and if various auxiliary assumptions also hold, **then** we should have certain expectations about the features of the outcome Y."

The auxiliary assumptions typically include features of X and type-level background information, and the expectations that we should have correspond to our rational degrees of belief regarding the features of the outcome Y. This general format can be instantiated in multiple ways, for instance:

Version 1: If X causes Y, and if various auxiliary assumptions hold, then we would expect Y to possess feature F.

Version 2: If X does not cause Y, and if various auxiliary assumptions hold, then we would expect Y not to possess F.

Both versions reason from causal hypotheses and background information to concrete, testable predictions about observable features. Taken together (when holding fixed suitable auxiliary assumptions), they imply that the causal hypothesis "X causes Y" raises the likelihood of observing feature F in outcome Y.

To illustrate tests based on outcome features, I offer two examples drawn from my background information. While the specific claims in these examples could be contested if my background information proves inaccurate, their primary function is to demonstrate the form or structure of such tests.

First Example: If the Great Recession was a cause of the recent electoral success of populist political parties across European nations, then we would expect to see the following features in the outcome: (1) Populist parties gaining significant electoral ground shortly after 2008–2009 and continuing through the recovery period; (2) Stronger populist performance in regions that experienced the most severe economic impacts from the recession (highest unemployment, largest drops in GDP, most austerity measures); (3) Populist parties performing particularly well among voters most economically affected by the recession (unemployed, working-class, those who lost homes or savings).

Second Example: If the partisan gerrymandering in Pennsylvania's 2011 congressional redistricting affected the 2012 U.S. House election results, then we would expect to see the following features in those election results: (1) There is a significant mismatch between the statewide popular vote share and the proportion of seats won by each party; (2) One party's votes being "wasted" at much higher rates than the other party's votes, either through excessive concentration in safe districts or spreading too thin across competitive districts; and (3) The party in charge of redistricting wins a disproportionate number of seats through oddly shaped districts that unnaturally split communities or combine disparate geographic areas.

The two examples are based on my background information (not explicitly stated in the examples). Example 1 draws on my understanding of how economic crises typically generate spatially and demographically differentiated patterns of political behavior. For instance, if economic hardship drives populist support, then areas hit hardest economically should show the strongest populist performance. If economic distress motivates populist voting, then the most economically affected individuals should be most likely to vote populist.

Similarly, Example 2 derives from type-level background information of how partisan gerrymandering affects various features of democratic representation. According to Samuel Wang (2016, 1266), a central principle of partisan gerrymandering is that it concentrates voters on a district-by-district basis such that both sides' wins are reliable, but the redistricting party's victory margins are smaller than those of the opposing party and are thereby used more efficiently. Significant mismatch between representation and popular support, evidence of strategic "wasting" the opposing party's votes, and unnatural shapes of gerrymandered districts all follow from this central principle.

My account of process tracing tests based on outcome features has several virtues. First, it addresses the problem of evaluating causal relationships within individual cases by specifying one type of observable consequences of

singular causal relationships. This shows that using within-case observations to test hypothesized singular causal relationships is at least methodologically feasible. In doing so, my account also reconciles the challenge raised by [Runhardt \(2015\)](#) and [Clarke \(2023\)](#) with the standard process tracing methodology. According to this challenge, it is difficult to see how the standard recommendations of using within-case evidence and causal process observations can help evaluate singular causal relationships within a case. My account demonstrates one way in which this can be done, meaning that the primary strength of process tracing—its capacity to leverage heterogeneous details within single cases—can be preserved when testing singular causal relationships.

Second, my account is compatible with different interpretations of how process tracing evidence should be evaluated, including both Bayesian and non-Bayesian approaches. For those who adopt the Bayesian interpretation of process tracing ([Bennett 2008](#); [Fairfield and Charman 2017](#)), my account naturally accommodates this framework: Tests based on outcome features can be framed in terms of the likelihood of feature-level evidence given a causal hypothesis, where more unlikely or surprising features provide stronger support for the hypothesis that predicts them ([Bennett 2008](#), 709). However, my account does not require Bayesian formalization. The basic logic that different causes produce outcomes with different characteristic features, and that observing these features provides evidence for causal claims, can be understood and applied without probabilistic formulation. This flexibility allows my account to be used by process tracers with different epistemological commitments regarding how evidence should be interpreted.

Third, my account also fits well with the idea that singular causal claims are rich in their counterfactual content, and that the basic “not-not” counterfactual dependence—the simple “if C had not occurred, then E would not have occurred”—typically does not exhaust the explanatory information associated with singular causal relationships. To see this point more clearly, I will discuss a few works in the philosophical literature on causation and the relationship of my account to them.

[Lewis \(2000\)](#) goes beyond the not-not counterfactual dependence by introducing the concept of “influence,” which is a pattern where variations in the details of the cause systematically relate to variations in the effect. In his words: “C influences E if and only if there is a substantial range C1, C2 ... of different not-too-distant alterations of C (including the actual alteration of C) and there is a range E1, E2 ... of alterations of E, at least some of which differ, such that if C1 had occurred, E1 would have occurred, and if C2 had occurred, E2 would have occurred, and so on” ([Lewis 2000](#), 190).

The intuition behind influence can be illustrated using an example. Suppose Billy throws a rock that shatters a bottle. Beyond simple counterfactual dependence (no throw, no shattering), influence captures how variations in the throw systematically affect variations in the shattering. If Billy had thrown earlier or later, the bottle would have shattered correspondingly earlier or later. Different trajectories or momentum would produce different shattering patterns, while throws that missed entirely would prevent shattering altogether (Woodward 2003, 218). Lewis notes that influence is a “pattern of dependence of how, when, and whether upon how, when, and whether,” and that it “admits of degree in a rough and multidimensional way” (Lewis 2000, 190).

Lewis (2000) defines causation as the ancestral of (sufficient) influence. Woodward (2003), in contrast, does not define causation through influence like Lewis does. Nevertheless, Woodward agrees with Lewis that singular causal claims typically (though not in every case) involve commitments to influence relationships as Lewis defined them (Woodward 2003, 219). According to Woodward, in situations in which no causal overdetermination is present, singular causal claims are associated with not-not counterfactuals: “if C had not occurred, then E would not have occurred” (Woodward 2003, 211). However, Woodward notes that causal claims associated only with such simple counterfactuals are often explanatorily shallow. In contrast, singular causal claims that are also associated with influence-type counterfactuals express more fine-grained and detailed patterns of counterfactual dependence, and therefore carry more explanatory information (Woodward 2003, 219).

Yafeng Wang (2022) further argues that fine-grained patterns of counterfactual dependence among features of events can help identify the actual causes of engineering failures such as plane crashes. Wang (2022) introduces the concept of “features” of events, which he defines as the detailed properties and spatiotemporal characteristics of these events. He further formulates a few conditional statements that capture the dependence of the features of an outcome on the features of one of its causes. One such statement is “feature necessity,” which states that “If an event C is a cause of an event E, then, given that E has certain features, C must have certain corresponding features” (Wang 2022, 104). Another statement is “feature sufficiency,” which states that “If an event C is a cause of an event E, then, given that C has certain features, E must have certain corresponding features” (Wang 2022, 104). Wang (2022) focuses primarily on feature necessity due to its central role in what he terms “descriptive enrichment,” which means using rich outcome details to infer the specific characteristics that its causes must have possessed. Still, he acknowledges briefly that feature sufficiency “could be used in a hypothetical-

deductive manner to test specific hypotheses about the features of past events” (Wang 2022, 104).

My account of process tracing tests based on outcome features is inspired by the central idea shared by Lewis (2000), Woodward (2003), and Wang (2022), namely that singular causal relationships between events C and E are often associated with a rich array of counterfactual dependencies between the detailed characteristics of C and E. I borrowed the term “feature” from Wang (2022), and the specific form of my process tracing tests has affinity with Wang’s (2022) formulation of feature sufficiency. However, in my account, observations about the features of the cause are included as part of the auxiliary assumptions needed to derive expected features of the outcome.

The idea that singular causal claims are associated with fine-grained and detailed patterns of counterfactual dependence aligns perfectly with process tracing’s central strategy. Process tracing focuses on identifying and leveraging fine-grained details to establish causal mechanisms in individual cases. To evaluate whether a singular causal relationship exists between events X and Y, researchers often must describe X and Y at much finer levels of detail and examine their individual features. Moreover, process tracers may initially possess more knowledge about type-level dependencies at the feature level than about the relationship between singular events X and Y, which makes feature-level analysis a valuable pathway for evaluating singular causal claims at the event level.

This completes my partial solution to the problem of evaluating causal relationships within a case. Next, I will describe an example of process tracing research to illustrate how the features of an outcome can be used to evaluate a singular causal relationship involving that outcome.

A Case Study

In this section, I examine an instance of process tracing research exemplified in the paper “Frontlash: Race and the Development of Punitive Crime Policy” by sociologist Vesla Weaver. Weaver seeks to explain the changes in the U.S. crime policies in the late 1960s and early 1970s. More precisely, her explanandum is a series of U.S. legislation containing punitive crime policies from roughly 1965 to 1972, including major bills such as the Law Enforcement Assistance Act of 1965, the Safe Street Act of 1968, and the Controlled Substances Act of 1970 (Weaver 2007, 231).

Weaver identifies puzzling features of this outcome that need explaining. First, the crime policy changes followed immediately after major breakthroughs in the Civil Right Movement, including the Civil Right Act of 1964 and the Voting Right Act of 1965. This raises questions about why crime

became politicized in the 1960s rather than earlier, and why race came to matter when it did (Weaver 2007, 232). Second, the policies contained peculiar details, including disproportionate targeting of riots and civil disorders for increased punishment and significantly increased federal funding to state law enforcement. Third, this series of policies initiated a dramatic increase in imprisonment rates, especially for African Americans and Hispanics.

Weaver (2007) proposes a causal mechanism that purports to explain some of these puzzling features and uses process tracing to provide evidence for it. Weaver's proposed causal mechanism for the U.S. crime policy changes in the late 1960s—which I will call “the frontlash mechanism”—highlights the causal role of politically conservative leaders regarding racial issues (henceforth “conservative elites” or “conservatives”). The mechanism identifies conservative elites' strategic planning to use crime as a way of regaining racial control in reaction to Civil Rights Movement successes as the key causal factor.

Weaver defines “frontlash” as “the process by which losers in a conflict become the architects of a new program, manipulating the issue space and altering the dimension of the conflict in an effort to regain their command of the agenda” (Weaver 2007, 230). She conceptualizes frontlash as a mechanism consisting of several key stages. The process begins when two political groups clash over a particular agenda issue and one side suffers defeat. However, rather than fading away or abandoning their ideology, the losing side seeks ways to regain its former power. Instead of defending the previous status quo, the defeated group strategically seeks to “shift the locus of attack” by finding entirely new issues to champion and developing fresh policy proposals around them (Weaver 2007, 236). This allows them to build momentum without directly challenging the established norms that emerged from their previous loss. The losing group then works to shape public understanding of their new issue. External crises or major political changes can make the public more receptive to their message. If their new issue gains sufficient traction and public attention, frontlash moves into what Weaver calls “issue capture,” which is when the previously losing group gains a monopoly on understanding their new issue, making opposition to them appear politically dangerous (Weaver 2007, 236). At this point, the original winners find themselves in “strategic pursuit,” positioning closer to their opponent's position despite it contradicting their preferences. Once issue capture is complete, the once-defeated group can establish lasting policies around their new issue that serve their interests, having successfully pushed their rival's priority issue off center stage and replaced it with their own preferred issue (Weaver 2007, 236).

Weaver argues that an instance of this frontlash mechanism operated in the 1960s U.S. crime policy changes. After the 1964 Civil Rights Act established protections for racial equality, conservative elites who opposed these changes shifted the political conversation from civil rights to crime and law enforcement. Two major events in the mid-to-late 1960s—rising crime rates and urban riots—provided them the perfect opportunity to advance their case. Conservative politicians and policymakers deliberately linked civil rights activism with criminal behavior, arguing that racial protests and crimes were equivalent. By framing civil rights demonstrations as threats to public safety rather than legitimate political grievances, they transformed the debate from one about equality to one about maintaining order (Weaver 2007, 236). This strategy proved effective. Conservative elites successfully convinced the public that being “tough on crime” was more important than advancing civil rights, and they created a narrative that portrayed Black activism as dangerous to society. Under pressure from this messaging, liberal politicians began adopting conservative talking points about law and order to remain politically viable. Eventually, both parties converged around a “tough on crime” policy approach. Conservative elites had successfully used the crime issue to advance their racial agenda without openly violating the new norms of racial equality established by the Civil Rights Act (Weaver 2007, 237).

I reconstruct this frontlash mechanism as a causal graph in Figure 1 (below), where the terminal node H represents the outcome to be explained—namely, the passage of punitive crime policy legislations in the late 1960s. This causal mechanism consists of both the various causal factors (represented as nodes) and singular causal relationships among these factors (represented as directed arrows).

Alongside her frontlash explanation, Weaver briefly considers two alternative explanations for 1960s U.S. crime policy changes. The first, which may be called the “crime reduction” explanation, argues that crime rates for both violent and property crimes had been increasing throughout the 1960s, and the punitive crime policies enacted during that time were political attempts to address this real problem of increasing crime rates (Weaver 2007, 233).

The second alternative explanation is labeled “backlash” by Weaver, defined as “the politically and electorally expressed public resentment that arises from perceived racial advance, intervention, or excess” (Weaver 2007, 237). According to this explanation, perceived racial advance in the Civil Rights Movement, particularly civil rights-related riots and demonstrations, caused resentment among a significant portion of the U.S. population in the 1960s. This public resentment was expressed politically and electorally, and political elites tried to appease it by adopting harsher punitive policies to win votes from this segment of the public (Weaver 2007, 238). Weaver notes that

the key distinction between frontlash and backlash is that backlash is reactive, and the main drivers of political change are the masses. In contrast, Frontlash is “preemptive, innovative, proactive, and, above all, strategic” (Weaver 2007, 238).

Weaver’s goal is to use process tracing to argue in favor of the frontlash mechanism over the alternatives. Based on my reconstruction of Weaver’s arguments, she uses features of the outcome to argue for the causal relations postulated in the frontlash mechanism and against causal relations postulated by alternative explanations.

Consider edge 7 in Figure 1, which goes from node B (conservatives formed a strategic plan of mobilizing the crime issue to regain racial control) to node G (conservatives pushed for more punitive policies in a series of congressional bills). Node G is observable; node B, in contrast, is hypothetical and unobservable. Weaver’s argument is that edge 7 provides the best explanation for some features of the outcome (node G), and that in the absence of edge 7, these features would be difficult to explain.

Weaver evaluates edge 7 in the context of the two competing hypotheses. Although the overall causal structures postulated by the competing hypotheses

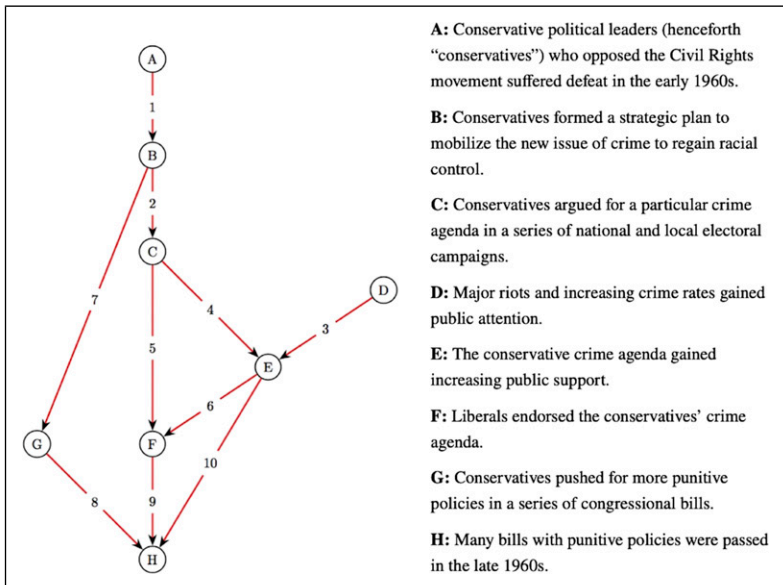


Figure 1. Causal graph of the frontlash mechanism leading to punitive crime policy legislation in the late 1960s.

are not made explicit in her paper, both clearly contain node G but postulate different causes of it. According to the crime reduction hypothesis, increasing crime rates throughout the 1960s caused conservatives to push for increased punitive policies. According to the backlash hypothesis, resentment among a significant proportion of the U.S. population toward civil rights-related demonstrations and riots caused conservatives to push for these punitive crime policies. Neither alternative hypothesis postulates node B or edge 7, and Weaver argues that the absence of edge 7 renders these alternative hypotheses incapable of explaining some relevant features of the proposed punitive policies.

The general shape of Weaver's arguments for edge 7 follows this pattern:

- (1) She cites some feature E of the node G (conservatives pushing for more punitive policies in a series of congressional bills).
- (2) She argues that if node B (the strategic planning of the conservative elites to mobilize the crime issue for racial control) was a cause of node G, then we would expect node G to have this feature E.
- (3) She argues that if node B was not a cause of node G (but some other nodes postulated by alternative explanations were), we would expect node G to not have this feature E.

In her arguments, Weaver uses three main types of features of node G. The first type of feature concerns the *content* of crime policies advocated by conservative elites. The second type of feature focuses on the *timing* of the advocacy of these crime policies. The third type of feature centers on the *identities* of the conservative elites who pushed for these crime policies.

To begin with, the argument based on the content of the crime policies pushed by conservative elites goes as follows: Compared to other motivations proposed by alternative explanations, frontlash is elite-driven and uniquely preemptive, innovative, and strategic. Given this feature of node B, we would expect node G to have the following content-related feature: many crime-related policies would include content strategically designed for racial control rather than crime control. Moreover, this type of feature can be found within the case: Many of the provisions introduced were indeed strategically designed for racial control (Weaver 2007, 255).

First, conservative elites introduced numerous laws to punish riots and civil rights unrest, even though this type of crime was not a major component of the total crime rate increases. As Weaver points out, "crime was rising fastest in the rural areas, the category of crime increasing most was property-based offenses" (Weaver 2007, 262). Between 1965 and 1969, lawmakers introduced nearly 100 pieces of federal legislation targeting riot participation,

transforming civil unrest from a local matter into a serious federal crime. Two examples illustrate this crackdown: The 1968 Omnibus Crime Control and Safe Streets Act made crossing state lines to participate in riots a federal felony, while the 1967 District of Columbia Crime Bill imposed penalties of up to 5 years in prison and \$10,000 fines for riot-related activities. Moreover, anti-riot provisions began being attached to completely unrelated legislation, including civil rights bills themselves. The 1968 open housing law contained an anti-riot provision sponsored by Strom Thurmond, nicknamed the “H. Rap Brown bill” after the Black Power leader (Weaver 2007, 250). Weaver argues that if reducing crime were legislators’ only motivation, and support for crime-related policies were unrelated to race, we would expect conservatives to push for stricter punishments for all types of violence and criminal activity, which was not the case: crime legislation was disproportionately targeted toward riots and civil unrest (Weaver 2007, 258).

Second, conservative elites also strategically changed how federal law enforcement funds would be distributed. As a case study, Weaver traced the negotiation, amendments, and passage of the Safe Streets Bill of 1968. The Safe Streets Act was one of the major crime legislations of the late 1960s, regarded by some historians as the “most extensive federal anti-crime measure in the nation’s history” (Weaver 2007, 257). Conservative legislators added amendments to this legislation to ensure that disbursement of federal law enforcement funds was not based on compliance with Civil Rights Act conditions, which gave the federal government power to withhold funds from racially discriminatory agencies (Weaver 2007, 255).

Moreover, conservative elites changed federally administered local grants into state block grants. Originally, the Johnson administration designed the program to send grants directly to cities and local agencies with populations over 50,000. However, House Republicans successfully amended the bill to give grants to states instead. This seemingly technical change had profound racial implications: direct city funding would have gone to agencies often controlled by liberal Democrats and Black officials in urban governments, while state-level distribution gave racially conservative state governors discretion over money usage (Weaver 2007, 255). The Safe Streets Act was the first major legislation to implement state block grants. As a result, the federal government invested millions of dollars in aid to state law enforcement agencies, many of which were racially discriminatory in the 1960s: States such as Alabama and Mississippi still prohibited Black people from being employed as state police (Weaver 2007, 255). Weaver remarks that “the above are powerful examples of how the conservative coalition used a pet project—crime—to advance the cornerstones of their prior agenda on civil rights” (Weaver 2007, 256).

Weaver's main point is that neither the cause proposed by the crime reduction hypothesis nor the cause proposed by the backlash hypothesis can account for the presence of these funding-related provisions in the Safe Streets Bill. The provisions regarding block grants were far too subtle and technical to result from public backlash against civil rights demonstrations. Average citizens unfamiliar with criminal justice administration would have no reason to know or care about the distinction between local grants and block grants. This means that these changes are more likely to result from strategic elite maneuvers rather than responses to popular resentment.

The second type of feature Weaver uses in her arguments concerns the timing of these crime-related policies. Weaver argues that the timing of crime policy development uniquely matches the frontlash mechanism, which predicts a relatively short lag between political defeat (the catalyst) and strategic issue mobilization. This timing pattern distinguishes frontlash from other explanations for the late 1960s crime crackdown. If rising crime rates alone drove policy changes, we would expect consistent attention to crime whenever rates increased. However, Weaver notes that "crime did not matter to conservatives until it was clear that it could be used as new currency to reestablish political advantage" (Weaver 2007, 262–63). Crime had been rising for nearly a decade before becoming politically salient. Even during the 1960 Nixon campaign, when crime had been rising substantially, it was not featured in campaign speeches. While previous national leaders had used crime as a campaign issue, none of these efforts produced sustained national crime programs (Weaver 2007, 240). This suggests that rising crime rates required additional political motivation to become a lasting policy priority—something that only emerged after the civil rights victories of the mid-1960s.

The backlash hypothesis suggests that public hostility to crime drove legislative action, but timing evidence contradicts this suggested sequence. Public opinion data shows that crime remained a low-salience issue until 1966, after it had already become a key campaign theme and after initial legislative proposals by conservative elites were enacted. Judging from the timing, rather than legislators responding to public demands, elite initiatives appear to have shaped public opinion rather than the reverse (Weaver 2007, 263–64).

The third type of feature concerns the identity of agents who pushed for crime policies. Weaver shows that the conservative coalition driving punitive crime policies after 1964 consisted of the same group of racially conservative Republican and Southern Democratic politicians and lawmakers who had previously championed segregation and opposed federal civil rights legislation. Moreover, they employed virtually identical characterizations of crime

and lawlessness in both contexts, which reveals strategic continuity between their pre-1964 segregationist agenda and their post-1964 crime agenda.

The conservative coalition included senators from Mississippi, South Carolina, Virginia, Alabama, Arkansas, and Florida, with figures like John McClellan, Richard Russell, and Strom Thurmond playing central roles. McClellan proved particularly influential due to his positions as chairman of both the Senate Subcommittee on Criminal Procedure and the Appropriations Subcommittee, which gave him extraordinary leverage over crime legislation (Weaver 2007, 243).

Before 1964, these same politicians used crime arguments to oppose civil rights. For instance, Senator Richard Russell argued in 1960 against integration by claiming “the extremely high incidence of crimes of violence among members of the Negro race is one of the major reasons why the great majority of the white people of the South are irrevocably opposed to efforts to bring about enforced association of the races” (Weaver 2007, 241). Instead of acknowledging that civil disobedience was due to legitimate grievance, they characterized civil disobedience as inherently criminal acts. After civil rights victories, these politicians simply reframed their arguments. Instead of predicting that civil rights would cause crime, they now claimed that civil rights had caused crime. The conservative strategy continued to collapse distinctions between nonviolent protest and riots, presenting all forms of Black political action as criminal behavior (Weaver 2007, 249).

In summary, the convergence of these three types of features in the outcome (conservatives pushing for more punitive policies in a series of congressional bills) provides evidence for Weaver’s causal hypothesis—that strategic planning by conservative elites to mobilize crime as a tool for regaining control over racial discourse was a key driver of this outcome. Alternative causal mechanisms would have difficulty accounting for this specific combination of features. This example illustrates how process tracers can employ feature-based tests to evaluate singular causal relationships and assess the plausibility of proposed causal mechanisms within a case.

Toward a More Complete Solution

The preceding sections presented the problem of evaluating singular causal relationships within cases and proposed a partial solution based on examining outcome features. In this section, I sketch what a more complete solution might look like and how the proposed partial solution contributes to it. I begin by discussing Hay’s (2016) critique of process tracing and argue that his critique is compatible with the project of developing a methodological toolkit for evaluating singular causal relationships in the social sciences. I then

consider two additional frameworks for assessing singular causal claims in the social sciences: [Shan and Williamson's \(2023\)](#) evidential pluralism and interpretivist approaches to process tracing. For each, I examine how they help assess singular causal claims and discuss their relationship to the feature-based approach proposed earlier. In discussing interpretivist approaches, I also discuss the assumptions process tracers make about the nature of social phenomena and the role of social meanings and interpretation in causal explanation.

In the third section, I defined the problem of evaluating singular causal relationships within cases: How can process tracers use within-case evidence to assess whether intermediate causal links in a hypothesized mechanism actually hold in a case? A more complete solution to this problem would consist of a toolkit of diverse methods and strategies for evaluating such singular causal claims by leveraging within-case evidence. It would make explicit both methods that social scientists have already employed (often implicitly) and methods they could potentially use. The feature-based approach proposed in this paper would be one tool among many that process tracers can deploy.

Developing such a methodological toolkit is compatible with an influential critique of process tracing by Colin Hay. [Hay \(2016, 500\)](#) argues that process tracing “is not a methodology but an ambition.” He acknowledges that there is considerable value in reflecting systematically on how one might best identify, track, and trace processes. However, calling process tracing a methodology presumes the challenge of identifying and tracing causal process is solved when in fact, it remains very difficult ([Hay 2016, 500](#)). What Hay objects to is not systematic methodological reflection on how to assess singular causal claims, but the presumption that challenges of tracing causal processes are already solved, and that there exists a singular, definitive method or solution ([Hay 2016, 501](#)). By presenting the feature-based approach as a partial solution, this paper adopts a similar stance of methodological pluralism. In what follows, I examine two additional frameworks for assessing singular causal claims in the social sciences and explore how they relate to the feature-based approach.

The first framework I examine is [Shan and Williamson \(2023\)](#)'s account of evidential pluralism. According to their account, establishing causal claims in the social sciences, including single-case or singular causal claims, normally requires establishing both that cause and effect are appropriately correlated and that a mechanism connects them ([Shan and Williamson 2023, 3](#)). For singular causal claims specifically, evidence of “single-case correlation”⁷ typically⁸ comes from counterfactual analysis based on established mechanisms together with detailed contextual facts ([Shan and Williamson 2023,](#)

127). This type of counterfactual analysis has been used by some historians to establish counterfactual dependence between historical events (Shan and Williamson 2023, 127n6). Evidence of a specific mechanism in a case can be provided by process tracing, which identifies and traces the causal process linking cause to effect. Once the mechanism is established, process tracing can also indirectly provide evidence for single-case correlation through counterfactual analysis applied to the mechanism (Shan and Williamson 2023, 128).

I argue that Shan and Williamson's evidential pluralism offers only a sketch of a solution to the general problem of establishing singular causal claims in the social sciences, rather than a fully developed solution. Evidential pluralism tells us that the role of process tracing is to provide evidence of mechanisms, but it does not explain how process tracing itself establishes the intermediate singular causal relationships that constitute the operation of a mechanism within a case. Suppose we want to show that event X caused event Y in a particular case. Evidential pluralism tells us to use process tracing to identify the causal mechanism linking X and Y. Yet tracing such a mechanism requires establishing a sequence of causally connected stages between X and Y, where each intermediate link is itself a singular causal relationship. How, then, are these intermediate causal relationships to be established? If the answer is that we must again rely on process tracing, the account risks generating a regress. Therefore, we need an account of how singular causal relationships can be evaluated that goes beyond the claim that process tracing helps trace the mechanism connecting cause and effect.

Shan and Williamson (2023, 153) also acknowledge that more needs to be said about how counterfactual analysis based on mechanism hypotheses supports single-case correlation, and I agree. Relying on counterfactuals of the form "if X had not happened, Y would not have happened" to support "X causes Y" in the social sciences faces multiple challenges, so it is important to specify the conditions under which counterfactual dependence supports singular causal claims.

One type of challenge concerns the prevalence of preemption in the social world. According to Rellihan (2025), human agents are rational and purposeful: when one path to a goal is blocked, they seek alternatives. Historical outcomes frequently have multiple potential paths, and if the actual path had been blocked, rational agents would likely have pursued alternatives leading to similar outcomes. The actual path taken preempts the alternatives: it is the actual cause even though alternatives existed (Rellihan 2025, 11–12). This creates problems for using not-not counterfactual dependence to establish causation: If preemption is pervasive in the social world as Rellihan claims,

then many singular causal relationships can hold without the corresponding not-not counterfactuals.

A second challenge concerns the prevalence of backtracking in historians' counterfactual reasoning. Historians who reason counterfactually often want to ensure that the counterfactual antecedent is historically realistic or plausible, which requires providing plausible backstories for how counterfactual antecedents would have come about (Reiss 2009, 721). But backtracking counterfactuals can lead to incorrect causal judgments: If we change one event X by changing its causes Z, and those causes Z also affected the outcome Y through independent pathways, the counterfactual may be true even though the changed event X did not cause the outcome Y (Reiss 2009, 722). These challenges suggest that the precise conditions under which not-not counterfactuals support singular causal claims and how those counterfactuals are themselves established require further examination.

The feature-based approach proposed in this paper complements evidential pluralism in two ways. First, it describes a concrete method for evaluating intermediate causal links within mechanisms. This grounds a methodological assumption that evidential pluralism relies upon, namely that process tracing can provide evidence for specific mechanisms in individual cases. Second, even though many singular causal claims are associated with not-not counterfactuals of the form "if X had not occurred, Y would not have occurred," evidence for "single-case correlation" need not come solely from establishing not-not counterfactuals. As argued in the fourth section, singular causal claims are often associated with other types of counterfactuals too, including influence-type counterfactuals or the dependence of the features of the outcome on the features of the cause. Evidence for these alternative types of counterfactuals can provide an alternative evidentiary route to "single-case correlation" when simple not-not counterfactual dependence is difficult or impossible to establish.

The second framework I examine is a family of approaches to process tracing that are inspired by a social science research paradigm called interpretivism. To understand interpretivism, it is helpful to contrast it with positivism, another research paradigm for investigating social phenomena in the social sciences. The core difference lies in how they treat social versus natural phenomena. Positivism treats the study of social and natural worlds as essentially similar; there is "little difference between studying the natural and social worlds" (Kaas et al. 2025, 606). Positivism takes reality to exist independently of human perception and focuses on observable data that can be measured and tested to discover causal relationships (Fischer 2025, 374n2). The emphasis is on "experience-distant evidence," or observables that can be

documented and verified independently of how actors themselves understand them (Kaas et al. 2025, 624).

In contrast, interpretivism insists that social phenomena are distinctive because they are constituted by the meanings attributed to them. Studying the social world differs fundamentally from studying nature because merely observing physical manifestations of action tells us what happened, but not why; we must also understand how social actors make sense of their interactions with other social actors (Kaas et al. 2025, 606). This emphasis on meaning leads to a focus on local causality in specific contexts: social causality is established locally because meaningful contexts give social practices their social effectiveness and generative power (Pouliot 2015, 237).

Several interpretivist approaches to process tracing have emerged in the literature, including Social Process Tracing (Kaas et al. 2025), Practice Tracing (Pouliot 2015), and Interpretive Process Tracing (Norman 2021), among others. These interpretivist approaches to process tracing share several distinctive features that set them apart from positivist variants. First, interpretivist approaches insist that understanding causal processes requires grasping the meanings, intentions, and social contexts that make actions causally effective, not merely observing physical actions. This manifests in their use of “experience-near” evidence (Kaas et al. 2025, 625). Second, interpretivist approaches share commitments to interpretive methods accessing tacit knowledge and taken-for-granted meanings, including ethnographic fieldwork and specialized interviewing that can indirectly probe tacit practical and cultural assumptions. Third, interpretivist approaches emphasize local or single-case causality over cross-case generalizations. This stems from recognizing that identical physical actions can have completely different causal effects depending on their meaning in different contexts. For example, military exercises produce cooperation between allies but threat responses between rivals because they count as different things in these contexts (Pouliot 2015, 242).

To illustrate how interpretivist approaches assess singular causal relationships, consider Kaas et al.’s Social Process Tracing (SPT) method.⁹ Drawing on Cartwright (2021), SPT assesses singular causal relationships between one actor’s action and another’s response using “if-then reasoning” grounded in causal principles (Kaas et al. 2025, 618). The logic is: If action X occurs, and if necessary support factors are present while derailers are absent, then, according to causal principle CP, response Y will follow. Causal principles are middle or low-level “tendency principles” describing widespread but not universal dispositions of individuals or institutions that require the right context to operate (Cartwright 2021, 13111). These principles can be

drawn from any social theory that explains why actors respond to one another in particular ways (Kaas et al. 2025, 626).

For example, the principle “When actors perceive an action as legitimate protest against injustice within their shared understanding of social norms, they are disposed to join or support that action” identifies support factors: Actors perceive injustice, understand the action as protest within their social context, and this understanding must overcome competing dispositions (“derailers”) (Kaas et al. 2025, 621). If process tracers discover that the support factors necessary for this principle to operate are present in a case while the derailers are absent, this provides evidence for the singular causal relationship (Cartwright 2021, 13116). For instance, if process tracers found through experience-near methods that Venezuelan barrio residents interpreted looting as a protest against government-created food scarcity rather than theft, they could use the aforementioned principle to establish a singular causal relationship between government policies creating food scarcity and the collective looting that followed (Kaas et al. 2025, 621). This approach not only establishes that a causal relationship exists but also explains why it holds.

What is the relationship between the feature-based approach and interpretivist approaches to process tracing? The approaches appear compatible in some respects. Examining outcome features could provide evidence for the meanings that interpretivists emphasize. Social Process Tracing recognizes that different causal principles generate different observable implications. For instance, grievance-motivated looting would produce different observable patterns from greed-driven looting (Kaas et al. 2025, 621). Feature analysis builds on this idea by identifying distinctive signatures that different causes leave in outcomes.

However, whether these approaches can be fully integrated into the same philosophical framework depends on how features are conceptualized. If features are treated as meaning-laden characteristics requiring interpretation through actors’ perspectives, they could be fully integrated into interpretivist philosophy, but it would also mean that different interpreters could arrive at different conclusions regarding what features are present in the outcome. If features function as relatively objective observables for discriminating between causes, they conflict with core interpretivist commitments. Many interpretivists insist that meanings cannot be reduced to observables that are independent of actors’ interpretations (Kaas et al. 2025, 624). Simply observing that legislation has certain content does not establish causation unless we understand what those features meant to actors in that context. Content features like “policies designed for racial control” already involve interpretation: What counts as “strategic design” requires understanding actors’ intentions and the social meanings of their choices.

This paper does not settle whether the feature-based approach defended here is fundamentally positivist or interpretivist in nature. Many interpretivist approaches to process tracing already incorporate hybrid elements from both interpretivist and positivist traditions,¹⁰ and I see no reason the feature-based approach cannot similarly exist as a hybrid. Following [Pouliot \(2015, 259\)](#), I suggest that positivist, interpretivist, and hybrid approaches can coexist in the methodological toolkit even if their philosophical foundations are somewhat incompatible. If each has valuable use cases for assessing singular causal relationships, including them all strengthens the toolkit while philosophical exploration of their compatibility continues.

Conclusion

This paper has identified and addressed a methodological challenge for process tracing: How to use within-case evidence to assess whether intermediate singular causal relationships in a hypothesized mechanism actually hold in a case. While process tracing claims to trace causal mechanisms through within-case evidence, the mechanisms it traces are themselves composed of chains of singular causal links that must be individually assessed. I have argued that singular causal relationships leave observable traces in the features of outcomes at each stage of a mechanism. If X causes Y, then Y should possess features that are characteristic of having been caused by X rather than by alternative potential causes. This feature-based approach provides a method for establishing intermediate causal links that is grounded in observable within-case evidence while being compatible with philosophical accounts of singular causation that emphasize counterfactual dependence.

The proposed solution represents a partial answer to the problem, contributing to a more complete toolkit for assessing singular causal relationships. A complete toolkit would include multiple approaches: feature analysis for identifying causal signatures in outcomes, causal principles ([Cartwright 2021](#)) for identifying necessary support factors and potential derailers, counterfactual reasoning based on mechanistic hypotheses ([Shan and Williamson 2023](#)), and interpretivist methods for understanding how social meanings and contexts make actions causally effective. These approaches may complement each other because each may be well-suited to certain research contexts but not others. Used together, they offer broader methodological coverage.

The examination of interpretivist approaches also addresses philosophical questions about process tracing's foundations: what assumptions process tracers make about the nature of social phenomena (particularly how they differ from natural phenomena) and what role social meanings and interpretation play in causal explanation. A philosophical question concerns

whether social causation can be established through observation of patterns that are relatively independent of actors' interpretations, or whether it inherently requires interpretive reconstruction of local meanings and contexts. While this paper does not settle this question, process tracers interested in the social dimension of process tracing should continue exploring both positivist and interpretivist approaches and the degree to which they can be integrated or combined.

Acknowledgments

I am grateful to the participants of ANPOSS 2025 and ENPOSS 2025, as well as two anonymous reviewers, for very helpful feedback on earlier versions of this paper.

ORCID iD

Yafeng Wang  <https://orcid.org/0000-0003-2126-4278>

Funding

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

Declaration of Conflicting Interests

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

Notes

1. While [Bennett and Checkel \(2015\)](#) and [Gerring \(2007 \[2017\]\)](#) define cases as instances of types, I also allow in principle for singleton cases representing unique phenomena that do not instantiate repeatable types.
2. This complementary use of process tracing alongside “large-N” quantitative research—where researchers first present statistical analyses of a large number of cases and then use process tracing to conduct detailed analyses of selected individual cases—has been increasingly common in recent scholarship. [Tang et al. \(2017\)](#), for instance, first used quantitative analysis to argue that oil discovery and production within the core territory of a subordinated minority group is correlated with the onset of ethnic war between the central government and the minority group, and then used process tracing to test out several hypothesized mechanisms that connect these two factors.
3. By “standard process tracing methodology,” I refer to process tracing recommendations by qualitative social science methodologists such as Andrew Bennett, Jeffrey Checkel, James Mahoney, and Derek Beach.

4. Even though this argument assumes the interventionist theory of causation, it is easily generalizable to any counterfactual analysis of (singular) causation.
5. In a later work, [Runhardt \(2021\)](#) accepts evidential pluralism in process tracing and argues that the methodological recommendations given by those who accept the “systems view” of process tracing could be reliable in at least some evidential contexts, and that the systems view could coexist with and complement the “interventionist” view of process tracing she defended in her 2015 paper. However, her interpretation of the systems view is that this view seeks “singularist” observations of causation, which are observations where one not only observes two contiguous events but where one also either directly observes that there is a causal relation between them or can infer the existence of such a causal relation a priori. In the case of process tracing, a singularist observation means one observes exactly “how each activity transmits causal force from one entity to another” ([Runhardt 2021](#), 432). In my view, this interpretation of the systems view gives up the attempt to answer the question of how within-case evidence can bear on evaluating actual causal relationships, and instead simply stipulates that certain actual causal relationships can be directly observed. I do not take it to be a satisfactory solution to the problem of evaluating actual causal relationships within a case, because I do not believe that singular causal relationships can be observed directly.
6. I remain neutral regarding whether alternative metaphysical categories—such as conditions, states of affairs, or values of variables—might serve as the relata of singular causal relationships. My claim is that regardless of what constitutes the causal relata in singular causal relationships, these relata possess detailed characteristics or features that can be identified and analyzed.
7. [Shan and Williamson \(2023\)](#) do not elaborate on what they mean by “single-case correlation.” One interpretation is that such a “correlation” amounts to counterfactual dependence between events.
8. According to [Shan and Williamson \(2023, 127\)](#), a single-case correlation may also be supported by a generic correlation in the population to which the case belongs. However, they do not explain in detail when it is appropriate to use such a population-level correlation between event types X and Y to justify a single-case correlation between particular events x and y that instantiate those types.
9. I focus on SPT here because [Kaas et al. \(2025\)](#) provide arguably the most detailed and explicit account of how to assess singular causal claims among the interpretivist approaches discussed.
10. I have used the term “interpretivist approaches to process tracing,” though “approaches inspired by interpretivism” would be more accurate, since all of these approaches are hybrid: [Kaas et al. \(2025, 638\)](#) describe Social Process Tracing as building on both interpretivist and positivist traditions, while [Pouliot \(2015, 259\)](#) and [Norman \(2021, 955\)](#) acknowledge that their versions of process tracing might

satisfy neither pure interpretivists nor positivists. All recognize the value of exploring intersections between interpretivism and more positivist approaches to social science.

References

- Bakke, Kristin M. 2013. "Copying and Learning From Outsiders? Assessing Diffusion From Transnational Insurgents in the Chechen Wars." In *Transnational Dynamics of Civil War*, edited by Jeffrey T. Checkel, 31–62. Cambridge University Press.
- Beach, Derek. 2021. "Evidential Pluralism and Evidence of Mechanisms in the Social Sciences." *Synthese* 199 (3): 8899–919.
- Beach, Derek, and Rasmus Brun Pedersen. 2013 (2019). *Process-Tracing Methods: Foundations and Guidelines*. 2nd ed. University of Michigan Press.
- Bennett, Andrew. 2008. "Process Tracing: A Bayesian Perspective." In *The Oxford Handbook of Political Methodology*, edited by Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, 702–21. Oxford University Press.
- Bennett, Andrew, and Jeffrey T. Checkel. 2015. *Process Tracing: From Metaphor to Analytic Tool*. Cambridge University Press.
- Cartwright, Nancy. 2021. "Rigour Versus the Need for Evidential Diversity." *Synthese* 199 (5-6): 13095–119.
- Clarke, Christopher. 2023. "Process Tracing: Defining the Undefinable?" In *The Oxford Handbook of Philosophy of Political Science*, edited by Harold Kincaid and Jeroen Van Bouwel, 305–27. Oxford University Press.
- Fairfield, Tasha, and Andrew E. Charman. 2017. "Explicit Bayesian Analysis for Process Tracing: Guidelines, Opportunities, and Caveats." *Political Analysis* 25 (3): 363–80.
- Fischer, Frank. 2025. "Interpretivism vs. Positivism in Policy Studies: Is There Common Ground?" *Critical Policy Studies* 19 (3): 371–76.
- Gerring, John. 2007 (2017). *Case Study Research: Principles and Practices*. 2nd ed. Cambridge University Press.
- Goertz, Gary, and James Mahoney. 2012. *A Tale of Two Cultures: Quantitative and Qualitative Research in the Social Sciences*. Princeton University Press.
- Grzymala-Busse, Anna. 2011. "Time Will Tell? Temporality and the Analysis of Causal Mechanisms and Processes." *Comparative Political Studies* 44 (9): 1267–97.
- Hay, Colin. 2016. "Process Tracing: A Laudable Aim or a High-Tariff Methodology?" *New Political Economy* 21 (5): 500–504.
- Kaas, Jonas Gejl, Mathilde Cecchini, and Derek Beach. 2025. "Elucidating the Social: Developing Social Process Tracing as an Integrative Framework." *Sociological Methods & Research* 54 (2): 605–45.

- Lewis, David. 2000. "Causation as Influence." *The Journal of Philosophy* 97 (4): 182–97.
- Mahoney, James. 2010. "After KKV: The New Methodology of Qualitative Research." *World Politics* 62 (1): 120–47.
- Mahoney, James. 2015. "Process Tracing and Historical Explanation." *Security Studies* 24 (2): 200–218.
- Norman, Ludvig. 2021. "Rethinking Causal Explanation in Interpretive International Studies." *European Journal of International Relations* 27 (3): 936–59.
- Pouliot, Vincent. 2015. "Practice Tracing." In *Process Tracing: From Metaphor to Analytic Tool*, edited by Andrew Bennett and Jeffrey T. Checkel, 237–59. Cambridge University Press.
- Punton, Melanie, and Katharina Welle. 2015. "Applying Process Tracing in Five Steps." Centre for Development Impact Practice Paper Annex. Institute of Development Studies. https://assets.publishing.service.gov.uk/media/57a08993ed915d3cfd000304/CDIPPracticePaper_10_Annex.pdf
- Reiss, Julian. 2009. "Counterfactuals, Thought Experiments, and Singular Causal Analysis in History." *Philosophy of Science* 76 (5): 712–23.
- Rellihan, Matthew. 2025. "Causes, Counterfactuals, and Counterfactual History." *Journal of the Philosophy of History*; published ahead of print. doi:10.1163/18722636-12341561.
- Runhardt, Rosa W. 2015. "Evidence for Causal Mechanisms in Social Science: Recommendations From Woodward's Manipulability Theory of Causation." *Philosophy of Science* 82 (5): 1296–307.
- Runhardt, Rosa W. 2021. "Evidential Pluralism and Epistemic Reliability in Political Science: Deciphering Contradictions Between Process Tracing Methodologies." *Philosophy of the Social Sciences* 51 (4): 425–42.
- Shan, Yafeng, and Jon Williamson. 2023. *Evidential Pluralism in the Social Sciences*. Routledge.
- Tang, Shiping, Yihan Xiong, and Hui Li. 2017. "Does Oil Cause Ethnic War? Comparing Evidence From Process-Tracing With Quantitative Results." *Security Studies* 26 (3): 359–90.
- Tannenwald, Nina. 1999. "The Nuclear Taboo: The United States and the Normative Basis of Nuclear Non-Use." *International Organization* 53 (3): 433–68.
- Waldner, David. 2015. "What Makes Process Tracing Good? Causal Mechanisms, Causal Inference, and the Completeness Standard in Comparative Politics." In *Process Tracing: From Metaphor to Analytic Tool*, edited by Andrew Bennett and Jeffrey T. Checkel, 126–52. Cambridge University Press.
- Wang, Samuel S.-H. 2016. "Three Tests for Practical Evaluation of Partisan Gerrymandering." *Stanford Law Review* 68 (6): 1263–321.

- Wang, Yafeng. 2022. "Feature Dependence: A Method for Reconstructing Actual Causes in Engineering Failure Investigations." *Studies in History and Philosophy of Science* 96: 100–111.
- Weaver, Vesla M. 2007. "Frontlash: Race and the Development of Punitive Crime Policy." *Studies in American Political Development* 21 (2): 230–65.
- Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.

Author Biography

Yafeng Wang is an assistant teaching professor at the Department of Philosophy and Religion, Northeastern University. His research interests include philosophy of technology, philosophy of the social sciences, and philosophy of psychiatry. He has published articles in several journals, including *Studies in History and Philosophy of Science* and *European Journal for Philosophy of Science*.