Original Article

Sociological Methods & Research
1–30

DOI: 10.1177/00491241251314039
journals.sagepub.com/home/smr


When to Use Counterfactuals in Causal Historiography: Methods for Semantics and Inference

Tay Jeong[1]

Abstract

According to the interventionist framework of actual causality, causal claims in history are ultimately claims about special types of functional dependencies between variables, which consist not only of actual events but also of corresponding counterfactual states of affairs. Instead of advocating the methodological use of counterfactuals *tout court*, we propose specific circumstances in historical writing where counterfactual reasoning comes in most handy. At the level of semantics, that is, the specification of the variables and their possible values, an explicit specification of the latent contrast classes becomes particularly useful in situations where one may be prompted to take an event that is pre-empted by the antecedent of interest as its proper causal contrast. At the level of inference, we argue that cases in which two or more antecedents appear to be playing a similar role tend to fumble our pretheoretical intuition about cause and propose a sequence of counterfactual tests based on actual examples from causal historiography.

[1]Department of Sociology, McGill University, Montréal, Canada

Corresponding Author:
Tay Jeong, Department of Sociology, McGill University, Canada.
Email: tay.jeong@mail.mcgill.ca

Keywords

historiography, actual causality, process-tracing, causal pre-emption, historical causation, counterfactuals

Introduction

Counterfactuals in history are a contested issue. Some scholars view historical counterfactuals as helpful, if not essential, for the writing of history, which typically *explains* how a certain event or phenomenon came about. Considering that explanation, including historical explanation, involves uncovering the causes of the explanandum (Carr [1961] 1990: 87; Halpern 2016), and causality has often been understood in relation to counterfactuals, counterfactuals may appear to be an indispensable component of historiography.[1] The idea that we must use historical counterfactuals to bring clarity to historiographical arguments is easy to come by in the social sciences and economic history (Fearon 1991; Scheidel 2019), which tend to have a relatively narrow and explicit focus on uncovering causal connections between past events. Yet, even in disciplinary history, the core relevance of counterfactuals has been repeatedly proposed, in large part for its role in uncovering causal relationships that give coherence to sequences of historical events (Bloch 1953, chap.5). For example, Citing Ferguson (1999), Gaddis (2002) states that "counterfactuals…[are] the historian's virtual equivalent of laboratory experimentation (p. 100)," and even argues that "[h]istorians use counterfactual reasoning all the time in establishing causation" (p. 102).

At the same time, counterfactuals in historiography have also received no shortage of rejection among its practitioners. As reflected in Leopold von Ranke's famous dictum that the historian "merely wants to show how it actually was" (Ranke [1824] 1874: VII), a popular idea among historians is that historiography is an enterprise that aims to tell what happened, not what would have happened.[2] This position has had continued support in theoretical studies of historiography as expressed in E. H. Carr's dismissal of historical counterfactuals as "parlor games" and Michael Oakeshott's calling them a "complete rejection of history" (Carr [1961] 1990: 97; Oakeshott 1985: 99; see also Tucker 2004, chap.5). Explicit antipathy to counterfactuals is less pronounced in historiography practiced in sociology and other fields of social science, but even here, the writing of history remains to be predominantly, and in many cases entirely, based on knowledge and representation of the actual world alone.

Historiography can be understood as a genre of academic writing that explains specific social phenomena by referring to their origins or causes

(Bloch 1953: 45). Intuitively, this involves tracing the diachronic development of a certain outcome of interest, narrating how a sequence of events—commonly described as a "process"—led up to it or brought it about (Mandelbaum 1977). A widely pointed out problem in understanding historiography in this way concerns giving a scientific definition to the concept of "process" when applied to history (Rohlfing and Zuber 2021: 1645–6). Following recent methodological developments in the sociological literature, we choose to formalize historical processes using the "interventionist" framework of singular causality (Runhardt 2022). According to this framework, historical cause-effect relationships or "processes" logically imply claims about the unrealized effect of interventions in past events under certain contingencies. The notion of "intervention" here is already familiar to historical scholars, referring to the imaginary experiment of "revisit[ing] the past, varying conditions … to try to see which would produce different results" (Gaddis 2002, 100).

If we understand historiography as arguments about the functional consequences of counterfactual interventions, nonmodal historical narratives that merely weave together sequences of actual events are readily seen as semantically inferior and blurry representations of a latent causal argument defined in the language of structural interventions. Indeed, prior theoretical attempts to account for historiographical explanation in the language of counterfactuals tend to stress the methodological need for a more explicit use of historical counterfactuals (Reiss 2012; Runhardt 2022). We are sympathetic to the idea that an explicit recourse to counterfactuals may help bring clarity to causal historiography and process-tracing, and this article does seek to propose a methodological integration of counterfactuals to the writing of history. Yet, we are also mindful of the "stickiness" of traditional historiographical practice not only in disciplinary history but also in sociology and adjacent fields of social science. As we argue below, the logic of historical causation—even within the interventionist framework—contains features that often obviate the practical need for explicit counterfactual reasoning, both in the domain of semantics and inference. Rather than propose the methodological integration of counterfactuals in historiography *tout court*, we seek to identify specific circumstances that make counterfactual reasoning especially beneficial and propose tools for its implementation.

The rest of the article proceeds as follows. The second section sets the theoretical foundations for subsequent methodological discussions by defining historical causal processes within the so-called interventionist framework of singular causation. The third and fourth sections will provide the main arguments of the article for each of the two logical components of causal

reasoning, namely, inference and semantics. In identifying the types of historiographical indeterminacies that call for systemic counterfactual reasoning, we pay particular attention to two issues that complicate how counterfactuals are implicated in actual causality and interfere with our pre-theoretical intuition about historical causality: Causal pre-emption and competing causes. We make frequent use of examples throughout the article, starting with simple artificial scenarios and subsequently expanding into actual examples from causal historiography. The fifth section concludes with a summary of the arguments of the previous sections.

Historical Processes as Contingent Counterfactuals

Historiography, at least in its modern form that White (1980) calls "history proper," is a narrative, which presents how an intimately linked sequence of events produced or led up to an outcome of interest. Some theoretical approaches to historiography propose to understand such linkages in idealistic or metaphysical terms, by, for example, grounding them in the historian's linguistic emplotment or the supposed perdurance of a macrohistorical entity (Ankersmit 2012; Froeyman 2009). However, in causal historiography popularly used in the social sciences, the transition from one event in a sequence to another is grounded in cause-effect relationships. We may thus understand causal historiography as a genre of writing that explains social events by identifying past events that caused them (Bloch 1953: 45–6).

It is common in the historical sciences to refer to such sequences of cause-effect relationships as a "process." An early work by Mandelbaum (1977: 57) stated that "the connection between the cause and the effect lies in the fact that both are seen as aspects of a single ongoing process, of which the effect is viewed as its end point or result." Goldstone (2003: 47) expressed that process tracing "consists of analyzing a case into a sequence of events and showing how those events are plausibly linked." Mahoney (2000) gives an example in which "King's [Martin Luther King, Jr.'s] death caused the failure of the Poor People's Campaign (B), which in turn led to massive summer riots (C), which heightened welfare militancy (D), which brought about an increase in AFDC acceptance criteria (E), and which fostered an explosion in the AFDC rolls in the late 1960s (F)" (Mahoney 2000: 526, originally cited from Isaac, Street, and Knapp [1994]). "These sequences have the familiar logic of A leads to B, which leads to C, which leads to D, and so on…" (Mahoney 2000: 526).

Despite the ubiquity of causal concepts in our natural language, our causal intuitions do not always converge, which raises the need for a formal

theoretical framework. Multiple theoretical frameworks allow for an analysis of actual causality, but here, consistent with recent works in sociological theory (Runhardt 2022), we will focus on what can be broadly referred to as the "interventionist" framework, which understands causality through a contrast between an actual outcome and a counterfactual outcome that is estimated to have resulted from a surgical "intervention" in the past.

For our purposes, Halpern's (2015) "modified HP account" provides the best theoretical articulation of interventionism. It is the latest version of the "Halpern-Pearl" or "HP" account of actual causality (Halpern and Pearl 2005), which is the most influential formalization of actual causality based on the Structural Causal Modeling (SCM) framework originally developed by Pearl now widely employed in quantitative social science. In addition, as explained in the following discussion, it straightforwardly accommodates the observation that historians often trace outcomes to their historical causes without a conscious appraisal of historical counterfactuals, even though, at least according to any account of actual causality based on the interventionist framework, causal reasoning is ultimately about contrasting the actual and the counterfactual.

A brief note about the term "actual" causality is in order. Sometimes used interchangeably with "singular" causality, this term is widely used in theoretical studies of causation to refer to singular cause-effect relationships that already (i.e., "actually") happened, which is just what we have in mind when we trace singular social phenomena to their historical origins or causes. In the words of Blanchard and Schaffer (2017: 179), "the relation of actual (or token, or singular) causation … is supposed to relate one token [i.e., nomadic or singular—author] event c to another token event e just in case c was in fact causally responsible for bringing about e." In the social-scientific theoretical literature, this corresponds to what Mahoney and Goertz (2006) and Holland (1986) called the "causes-of-effects" approach, as opposed to the "effects-of-causes" approach characteristically used in experimental or statistical studies.

HP theories of actual causality define causality within a given Structural Causal Model (SCM), which is a mathematical model that maps out how a set of variables are functionally related to one another.[3] Formally, an SCM M is a pair $<S, F>$, where S is a signature that "lists the exogenous and endogenous variables and defines their possible values," and F is a set of functions that defines the "modifiable structural equations," which are descriptions of how the value of each endogenous variable in the model is determined as a function of other variables (Halpern and Hitchcock 2015: 418). Specifically, S is defined as a tuple $<U, V, R>$, where U lists all the exogenous variables, V all the endogenous variables, and R the possible

values that can be assigned to U and V. F associates with each endogenous variable $X \in V$ a function F_X that determines the value of X given the values of all other exogenous and endogenous variables. Intuitively speaking, V includes the list of variables that one is willing to consider in causal reasoning, and U can be understood as taken-for-granted "background circumstances" that define each distinct "case" (whether it be individuals, countries, or other historical settings) and determine the values of all the endogenous variables through the functional relationships F.

According to Halpern's (2015) "modified HP account," actual causality is defined as follows (Readers not familiar with the formal syntax may refer to the immediately ensuing text for a nontechnical elaboration of the concepts. Readers already familiar with HP-theories of actual causality may jump to the final two paragraphs of the second section 2).

Let $\vec{X}$ (the "antecedent") stand for a conjunction of primitive events, which can be alternatively written as $X_1 = x_1 \wedge X_2 = x_2 \wedge \ldots \wedge X_k = x_k$.

Likewise, let φ (the "outcome") stand for a Boolean conjunction of primitive events (Halpern 2015: 3023).[4] While φ can be a conjunction of multiple primitive events in Halpern's definition, historiography is almost always concerned with explaining a single historical event or circumstance. We will therefore somewhat roughly, but inculpably, refer to φ as the "outcome" in the subsequent text.

$(M, \vec{u})$ is a notation for the *causal setting*, composed of the causal model M and exogenous assignments $U = \vec{u}$. Given that $\vec{u}$ refers to the "background circumstances" that hold in the actual world, $(M, \vec{u})$ denotes the actual state of affairs. The symbol $\models$ expresses a relationship of entailment. Square brackets are used to specify interventions in antecedent variables.

$\vec{X} = \vec{x}$ is an *actual cause of* φ *in the causal setting* $(M, \vec{u})$ if the following three conditions hold:[5]

AC1. $(M, \vec{u}) \models (\vec{X} = \vec{x})$ and $(M, \vec{u}) \models \varphi$

AC2. There is a set $\vec{W}$ of variables in V and a setting $\vec{x}'$ of the variables in $\vec{X}$ such that if $(M, \vec{u}) \models \vec{W} = \vec{w}^*$, then

$$(M, \vec{u}) \models \left[\vec{X} \leftarrow \vec{x}', \vec{W} \leftarrow \vec{w}*\right] \neg\varphi$$

AC3. $\vec{X}$ is minimal; there is no strict subset $\vec{X}'$ of $\vec{X}$ such that $\vec{X}' = \vec{x}'$ satisfies conditions AC1 and AC2, where $\vec{x}'$ is the restriction of $\vec{x}'$ to the variables in $\vec{X}$.

AC1 merely states the antecedent $\vec{x}$ and outcome φ are both actual events —an arguably trivial first condition for actual causality.

AC2 forms the core of the definition. AC2 is a clause that allows us to admit some events that are not in a relationship of simple counterfactual dependence with the outcome as actual causes. If we merely define a cause as an event without which the outcome would not have occurred (Hume 1748: Section VII; Lewis 1973), we face the problem that we are pretheoretically strongly inclined to call certain events causes even though their nonoccurrence would still have resulted in the same outcome. Such situations mainly occur when the nonoccurrence of the antecedent would have prompted an alternative series of events that would have resulted in the same outcome. In the interventionist framework of causation, this problem is typically tackled by defining causality as counterfactual dependence under certain contingencies that peg some endogenous variables V in the model at specific values. Different attempts to define actual causality vary in the nature of the contingencies under which the outcome must counterfactually depend on the antecedent. In the AC2 of the modified HP account, $\vec{W}$ denotes the subset of the endogenous variables to be pegged, and the expression $\vec{W} \leftarrow \vec{w}^*$ requires that they be intervened on and pegged to their *actual values*. In effect, AC2 states that, for an event $\vec{x}$ to count as a cause of an outcome φ, φ must counterfactually depend on $\vec{x}$ while holding some subset (including a null set) $\vec{W}$ of the endogenous variables V pegged to their actual values.

AC3 is just a minimality condition that filters inessential elements from being included in a cause. To cite Halpern's own example, if having dropped a match qualifies as an actual cause of the forest fire according to AC1 and AC2, AC3 blocks the conjunction "dropping a match *and* doing something irrelevant (such as sneezing)" from also qualifying as a cause.

Having introduced the main theoretical framework for historical causality grounded in counterfactual dependence, we attempt to do two things in the following sections. First, we try to make sense of the strong empiricist tradition in historiography and process-tracing by pointing to features in the logic of actual causality that reduce the need for explicit counterfactual reasoning. Apart from pure theoretical interest, this step is intended to forestall a seemingly logical jump from a *theory* of causality based on counterfactuals to a *methodology* that renders historiography into a thoroughgoing exercise in historical counterfactuals. Second, we make schematic recommendations on when and how focused counterfactual reasoning may come in handy for tracing historical causes or "processes."

Causal reasoning as hitherto theorized comes with two logical components: One must obviously infer the outcomes of interventions, but for that to happen, one must have already described the intervention. In the formal language of the SCM, the two steps correspond to the two components of a causal model M, that is, S which defines all the variables together with

their possible values, and F which relates those variables through functional dependencies. We call these two logical steps "semantics" and "inference", respectively. The following sections discuss the methodological implications of counterfactuals for process-tracing regarding each logical component, starting with inference immediately below.

Counterfactuals for *Inferring* the Outcomes of Interventions

For an illustration of the SCM, we'll start with the simplest possible imaginary scenario. Imagine that there is a bottle, a person (say, Suzy) walks up to it, throws a rock, the rock hits the bottle, and the bottle shatters. The situation can be represented by a simple model M with a set V containing two endogenous variables (ST: Suzy throws, BS: Bottle shatters) and an exogenous variable U that represents all the psychophysiological factors that determine whether Suzy throws or not. R would assign binary values 0 or 1 to all variables, such that Suzy either throws or does not throw, and the bottle either shatters or does not shatter. The functional relationships F between the variables consist of the structural equation BS ← ST, which means that BS is determined by ST through an identity function. BS counterfactually depends on ST, so this trivially makes ST an actual cause of BS: It is merely a special case where a null set satisfies the set $\vec{W}$ in AC2.

This simplest model can be made more complex by adding causal preemption as depicted in Figure 1. We are often concerned about the possible existence of someone or something other than Suzy that may have shattered the bottle had Suzy not thrown, that is, a potential pre-empted cause. Assume

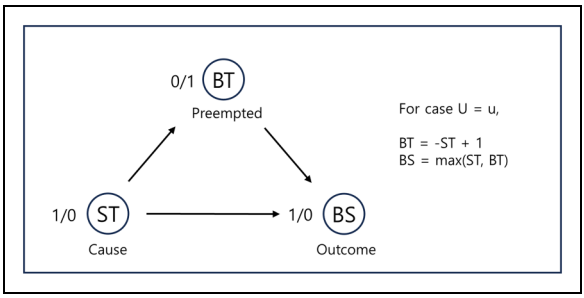

Figure 1. Causal pre-emption in the rock-throwing example.
Note: Values next to the variables indicate possible values, with the one on the left being the actual.

that another person, Billy, was standing behind Suzy ready to throw in case Suzy did not throw. Regardless of our degree of certainty about whether Billy would have managed to hit and shatter the bottle had he gotten his turn, it would be clear from observation that Suzy's throw was an actual cause of the bottle's shattering.

Practically, insofar as causal pre-emption is concerned, our judgments of actual causality are informed by the case where $\widehat{W}$ in AC2 is a set of possible pre-empted causes. The basic intuition behind AC2 is that an antecedent counts as a cause of the outcome if the occurrence of the outcome depends on the occurrence of the antecedent under the contingency that the alternative way(s) that could have brought about the same outcome indeed never took effect (Halpern 2015). Holding fixed the fact that Billy did not throw, if the bottle would not have shattered had Suzy not thrown, we judge that Suzy's throw actually caused the bottle's shattering.

When historians or social scientists talk about historical counterfactuals or "what-ifs," they typically have in mind interventions that do not artificially hold fixed the values of pre-empted events, allowing one to assume such events to have occurred as anticipated in response to the intervention (Bunzl 2004; Fearon 1991; Levy 2015; Reiss 2012; Rosenfeld 2002; Scheidel 2019; Tucker 2004:187–8, 232–3). In cases where the outcome counterfactually depends on the antecedent in such a free-floating way, the antecedent is said to be a "but-for" cause of the outcome, as the outcome would not have occurred but for the occurrence of the antecedent (i.e., it would not have occurred through any alternative path forestalled or pre-empted by the antecedent) (Halpern 2016).

But-for causes play a crucial role in special types of historical studies that do not merely aim to explain a singular historical event by identifying past events that led up to them but explain *divergence*, that is, why one, but not another, type of outcome occurred. Relatedly, testing a hypothesis that the occurrence of a certain past event made the difference between the actual outcome and an unrealized alternative outcome consists of examining whether that past event is a but-for cause. For example, checking whether Suzy's throw made the difference between the bottle's shattering and not shattering would amount to checking whether Suzy's throw was a but-for cause of the bottle's shattering. Fogel's (1964) study of whether the extensive development of railways in the nineteenth century in America made the difference between successful actual industrial development and a potential unsuccessful development at the beginning of the twentieth century comes down to testing whether the expansion of railways was a but-for cause of American economic growth.

Despite their *conceptual* simplicity, identifying or testing the special relationship of but-for causation generally poses more formidable methodological challenges than identifying or testing the broader category of actual causes since the former is more fully exposed to the complications of knowing nonactual possible worlds, which, to quote a well-known expression, is "the fundamental problem of causal inference" (Holland 1986). In the bottle-shattering scenario with pre-emption shown in Figure 1, as long as one is merely interested in tracing the outcome BS to its actual cause, one may ignore the more difficult question of whether Billy would have managed to hit and break the bottle had he been given a chance to throw. Similarly, in Fogel's (1964) study of American railways and economic growth, one may safely bracket off the question of what alternative means of inland transportation would have expanded if it weren't for railways and what consequences they would have had on industrial development—a speculative task that Fogel (1964) could only tackle using complex econometric projections—when considering whether rail transportation actually brought about American industrial growth. In many cases, whether the outcome counterfactually depends on the antecedent of interest while holding the possible pre-empted causes to their actual values would come across as almost self-evident, much so that the historian gets the impression of seeing a connection between two actual events. In contrast, inferring the outcomes of pre-empted events requires seeing a connection between two nonactual events and seldom offers a clear intuition, raising the need for focused counterfactual inference using specialized methods such as theory-guided imagination, mathematical predictions, or systematic comparison with similar cases.

The logic just described lies at the core of the observation that causal historiography and process-tracing usually do not require counterfactual thought through and through even though, according to any interventionist account of causality, cause–effect relationships are logically defined as relationships of contingent counterfactuals. The methodological complexities of inferring the results of pre-empted events and the contrasting ease of counterfactual inference when insulated from the issue of unrealized alternatives also align with the heavy empiricist inclination among historiographical practitioners and the scarcity of focused discussions of counterfactuals in the theoretical literature on process-tracing and causal historiography (Bennett 2010; Clarke 2023; Dowding 2023; Ermakoff 2019; Tucker 2004; Waldner 2012).

Still, merely bracketing off the complexities of causal pre-emption will not always ensure a clear intuition about historical causation. The most common solution for disagreements about historical causality is simply to collect more and better historical evidence (Tucker 2004), adding fine-grained "mediators"

situated between the putative cause and effect capable of bridging gaps in our historical intuition (Clarke 2023:307–8).[6] For example, whether Japanese colonialism actually caused South Korean economic growth spurt from the mid-1960s emerged as a popular question in economic history starting from the late 1980s following South Korea's exceptionally successful developmental outcomes among "late-late" industrializers. An influential body of works argues that Japanese colonial rule laid enduring foundations in human and physical capital, a legacy of a penetrating modern state, and various sociopolitical institutions and knowhows such as cohesive business-state relations that ultimately brought about the growth spurt of the 1960s (Amsden 1989; Kohli 1994; McNamara 1990; Pilat 1994; Yang 2004). However, this thesis has received much controversy, in part due to the highly tumultuous nature of the three-decade interregnum between the end of colonial rule and the beginning of the growth spurt. To express in the language of counterfactuals, it was never intersubjectively clear if the "miraculous" post-mid-1960s growth spurt of South Korea would have still occurred without the experience of the industrialization under Japanese colonial rule even when holding fixed the fact that the alternative developmental possibilities thwarted by Japanese colonialism (such as an independent Korean indigenous modernization during the same period) indeed never took place. Scholars who wish to claim a causal connection are usually prompted to look for evidence of stepping stones or institutional continuities that bridge the logical gap between the two causal relata (Kim 2012), such as a continuity of entrepreneurship (Eckert 1991; McNamara 1990) or the transmission of knowhows of state-led capitalist industrial development that hibernated in the form of institutional memory (Kohli 2004).

There is no doubt that a thorough knowledge of the actual events and circumstances that occurred between the putative cause and effect greatly contributes to resolving historiographical indeterminacies (see Ermakoff, 2019 for a comprehensive account of the types of empirical evidence that can be used for tracing "genetic" causation in historical social science, which we, in the present article, are defining in terms of counterfactuals). However, whether it is always sufficient or optimal is a different matter. Although process historiography does not by default require counterfactual thought, we still argue that focused counterfactual reasoning has an important place in its methodology, especially in situations of historiographical disagreement or uncertainty. Here, historians with a strong empiricist bent may contend: If one is unable to convincingly demonstrate how one event led up to another in plain narrative based on actual events, would explicitly asking and inferring the counterfactuals be any more feasible or convincing, particularly in situations that operate outside of formal theory or quantifiable mechanisms? Yet,

in some cases, the nature of inquiry will be such that it severely clouds our pretheoretical intuition of cause, and thinking about causation in terms of counterfactuals offers a principled guideline for conducting historical causal inference.

Such situations commonly arise when we are confronted with two or more antecedents that can be expected to play a broadly similar role. Causal judgment was easy in the pre-empted scenario of Figure 1 partly because, once the alternative path (Billy's rock) is bracketed off, there was clearly nothing else that could have played the same role as Suzy's rock, that is, come into hard contact with the bottle. The same holds for Fogel's (1964) railway study, since, when alternative means of inland transportation such as roads and canals that could have expanded in lieu of railways in the nineteenth century are bracketed off, there is practically nothing other than railways that could possibly have moved industrial goods over long distances inland in a comparable scale. However, things are not always so simple in history, and there are sometimes other actual events not pre-empted by the antecedent of interest that also appear to have played a similar and potentially important role.

Consider a scenario where both Suzy and Billy throw, their rocks hit the bottle at the same time, and the bottle shatters. Who shattered the bottle (who caused the bottle to shatter)? This question, visualized in Figure 2, can only be answered by knowing the functional relationships that the two antecedents of interest have with the outcome BS. Let SH stand for "Suzy's rock hits" and BH "Billy's rock hits." As we are dealing with actual causality, we will bracket off any other events pre-empted by SH or BH (grayed out in Figure 2). In the conjunctive case where both SH and BH are required for the bottle to shatter, that is, if $BS \leftarrow \min(SH, BH)$, then each is a but-for cause and thereby an actual cause. If SH is required but BH is irrelevant, that is, if $BS \leftarrow SH$, then only SH is a cause. The most confusing case is the disjunctive scenario commonly known as "overdetermination," where either SH or BH is sufficient for BS, that is, $BS \leftarrow \max(SH, BH)$. In this case, the modified HP definition holds that the conjunction $SH \wedge BH$ is a cause of BS, and SH or BH alone is only a *part of a cause* (Halpern 2016: 29). In natural language, what are parts of causes according to the modified HP definition are sometimes just referred to as "causes" (Halpern 2016:25–26), although our pretheoretical intuition of "cause" is less firm in cases of overdetermination (McDermott 1995:527).

In the special rock-throwing scenario just described, the observer would most likely be unable to "see" the causal status of SH and BH and need to rely on counterfactual reasoning if she were to undertake the inquiry (which may or may not be feasible). As there are two antecedents that each come with two possible states, one needs to know the outcomes of four

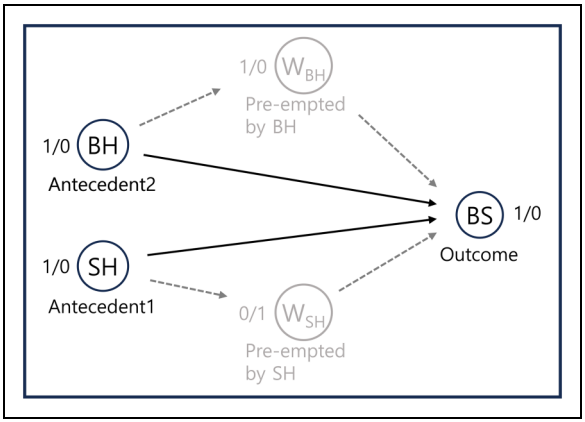

Figure 2. Causal reasoning with two competing antecedents.
Note: Values next to the variables indicate possible values, with the one on the left being the actual. Events pre-empted by each of the two antecedents are grayed out, indicating that they be bracketed off for causal inference.

different combinations to fully identify the function that maps SH and BH to the outcome BS. One of the four consists of events in the actual world, which should already be known to the historian through empirical methods. Among the three counterfactual pairs, it would in most cases also be almost obvious that, if neither antecedent occurred, the outcome would not have occurred: The scholar would not have singled out those two events as potential causes if she had thought otherwise. One may explicitly ask and infer either one or both of the remaining counterfactuals for a principled tracing of historical causes. For extra methodological robustness, researchers may also explicitly consider the case where neither antecedent occurred.

The historiography of Japan's decision to surrender at the end of World War II provides a useful real-world example. Japan's surrender on August 15, 1945 came much earlier than what many had anticipated until just several weeks before, as the Japanese leadership was preparing for a defense of their home islands from an allied landing invasion. Two shocking events that happened within the narrow interval from August 6 to August 9 are widely considered to have made the situation completely unbearable for the Japanese leadership: The dropping of two nuclear bombs by the United State (for convenience, we will group the two A-bombing as one event) and Soviet entry into the War (Akagi and Takita 2016). However, how causally pertinent was each of these two antecedents? In particular, America's

decision to use atom bombs has received continued controversy, with critics characterizing the event basically as a meaningless sacrifice that only served to ensure America's political supremacy after the War (Alperovitz 1965; Nishijima 1968). An implicit argument that underlies such accusations is to depict the scenario as that of "overdetermination," where the atom bombs inflicted massive superfluous injury on an enemy that would have been unable or unwilling to fight much longer anyway.

Situated in such an argumentative context, Asada's (1998) historiography takes a specific interest in the relationship between Japan's early decision to surrender and the atom bombs. Calling historical counterfactuals as "risky" but needed for a full account of Japan's surrender (Asada 1998: 510), the author, as a preliminary check, considers a possible world in which *neither* Soviet entry nor the atom bombings occurred, which means that only conventional bombings continued to put pressure on Japan. Based on the high willingness to keep fighting recorded for most factions in the Japanese military leadership before the "twin shock," Asada convincingly establishes the counterfactual that Japan would not have surrendered at such an early date if neither of the two events occurred. This justifies the exclusive focus on these two events while allowing other events (such as conventional bombings) to be simply embedded in the exogenous variable U representing all the miscellaneous factors that form the background circumstance.

Asada then goes on to imagine the counterfactual *with* Soviet entry but *without* the atom bombs (Asada 1998: 510). The decision to first toggle the atom bomb (rather than the Soviet entry) from actual history is natural considering that the atom bomb is the main interest of his historical-causal inquiry. Note that the backward-looking nature of causal historiography (the so-called "causes-of-effects" approach) requires the imagination of counterfactual absences (i.e., the atom bombs were not dropped) from actual presences (i.e., atom bombs *were* dropped), rather than the manipulation of presences on top of a condition of initial absence typically used in forward-looking "effects-of-causes" approaches. In the absence of causal preemption, the atom bombs will count as a (full) cause of Japan's early acceptance of the Potsdam terms if and only if its absence would have altered the outcome. According to Asada, although the long-dreaded Soviet entry precluded all outcomes other than an utter defeat for Japan, it did not produce the psychological shock effect that could move the most hardline members of the Japanese War Cabinet, who were calling for "the last sacrificial homeland battle" already in partial anticipation of an eventual Soviet entry (Asada 1998: 494). If true, this inference establishes the atom bombs as an individual

cause of Japan's early decision to surrender and logically precludes the possibility of overdetermination.

Having achieved his main goal, Asada does not push the explicit counterfactual queries further. The remaining counterfactual, that is, a scenario without Soviet entry but with the atom bombs, may be additionally conducted if the author was also interested in the causal status of Soviet entry. At least according to Asada's reconstruction of the final weeks of the Japanese War Cabinet, the main preoccupation for the conservative leaders in the Imperial Army by that time of the War was "saving face": Even if they wanted to surrender, their code of conduct prevented them from doing so. And the extraordinariness of the atom bombs perfectly fit into the popular face-saving narrative that Japan lost only because of America's technological prowess. Atom bombings alone without Soviet entry would most probably still have resulted in the same outcome, that is, Japan's early surrender. Such a conclusion would deny Soviet entry as a full actual cause of Japan's timely surrender, leaving only the atom bombs as the cause that brought about its timely acceptance of the Potsdam terms.[7]

Table 1 is an ordered sequence of counterfactual tests that one may use in the type of setting depicted in Figure 2, that is, where there are two actual historical antecedents that appear to have caused the outcome.

It may seem that such cases would be relatively rare in historiography, but they are a very common setting that gives rise to historiographical indeterminacies. The two rival antecedents do not need to "break out" simultaneously or within a narrow time slot as in the example of the atom bombs and Soviet entry. They may also come in the form of two chronologically enduring conditions or traits. Moreover, even historiographical disagreements apparently focused on the causal status of a single controversial antecedent can be fruitfully situated within an expanded framework of the kind just analyzed, that is, a competition between two rival antecedents that appear to have fulfilled a similar role logically required for the outcome's occurrence.

For an illustration of the latter point, we may revisit the historiographical controversy on the colonial origins of the South Korean economic growth spurt mentioned shortly before. The idea that colonialism—still a strong historical trauma in Korea—was the origin of perhaps its biggest modern achievement drew heated historiographical controversies to the dyadic relationship between colonialism and postwar economic development in Korea. The initial rise of this historical hypothesis was not backed by a clear understanding of the mechanisms of an intertemporal transmission of knowhows, capital, and institutions, and subsequent historical debates frequently focused on digging deeper into the specifics of the 30-year

Table 1. Sequential Counterfactual Tests in Situations With Two Competing Antecedents, Based on Asada's (1998) Historiography of Japan's Decision to Surrender.

Actual history Atom bomb (AB)= 1, Soviet entry (SE) = 1, Japan's surrender (JS) = 1			
Counterfactual test	Outcome	Judgment	Note
Test 1 (toggle both) AB = 0, SE = 0	JS = 0	AB or SE causally relevant. Proceed to test 2	Preliminary check; may be omitted if deemed superfluous within the argumentative context
	JS = 1	Rethink causal model; likely missing out a crucial factor from the set of endogenous variables	
Test 2 (toggle AB) AB = 0, SE = 1	JS = 0	AB is an actual cause of JS	Like Asada (1998), one may stop here if the main goal is to show that AB is a full cause
	JS = 1	AB is not a (full) cause of JS	AB may still be causally relevant as part of a cause in overdetermination scenarios
Test 3 (toggle SE) AB = 1, SE = 0	JS = 0 JS = 1	SE is an actual cause of JS SE is not a (full) cause of JS	If JS = 1 in test 2, then each of AB and SE is part of a cause; Overdetermination

interregnum and arguing for the continuities or discontinuities (Eckert 1991; Heo 2012; Kim 2012; Kohli 1994; McNamara 1990).

Perhaps the most intuitive way to use counterfactual reasoning to tackle a disputed causal link between two events is simply to imagine the negation or absence of the hypothesized cause. Such single-variable counterfactual queries have also been posed—mostly unfruitfully and with much confusion—in the academic literature on colonialism and development in Korea, as we discuss in more detail in the fourth section. Yet, researchers may

also consider expanding the scope of the counterfactual query to set it up as a competition between two historical antecedents following the scheme of Figure 2. This has the effect of temporarily zooming out of the focus on the historical details between the posited cause-effect dyad and situating the historical query within the theoretical tension that gives rise to the historiographical disagreements. Methodologically, the introduction of a competing antecedent provides concrete guides for counterfactual inference, specifying what other factors would still have existed even as we imagine the absence of the factor of immediate interest. This dual setup also allows us to talk about historical causality with semantic precision, capable of not only identifying individual causes but also parts of causes in overdetermination scenarios.

In the example of Korean development, the counterfactual query forces us to think what other factors other than colonial development that could possibly have enabled such an unexpected and "miraculous" outcome in one of the poorest countries in the world. Another widely recognized possible origin of the human capital, market-supporting institutions, and state-managerial capacity required for successful capitalist development has roots not in pre-war Japan but in the Cold War regional architecture led by the United States (Krieckhaus 2017), which not only provided direct material and policy assistance to the nascent state of South Korea to curb the tide of socialism (Heo 2012; Stubbs 1999) but also constructed a hierarchically integrated international value chain involving the United States, Japan (with which South Korea, despite popular antipathy, reestablished diplomatic ties in 1965), and South Korea, giving the latter close access to large and advanced foreign economies that offered market access and technological assistance (Chibber 2003). Incorporating this alternative postcolonial developmental origin into the causal query serves to highlight the historical tension in the counterfactual imagination of what would have happened had the capital and institutional legacies of Japanese colonialism been absent. The scholar may then proceed according to the steps outlined in Table 1 depending on the immediate goal of the inquiry. For example, scholars who wish to claim that Japanese colonialism had enduring benefits for South Korean development would need to establish it as a cause or at least a part of a cause in an overdetermination scenario, which would require the toggling of the main antecedent of interest (test 2 in Table 1) and possibly also the alternative antecedent (test 3 in Table 1) in case the previous test failed to establish Japanese colonial development as a full cause.

Counterfactuals for Specifying the *Meaning* of Causal Claims

Any interventionist account of causation defines the notion of "event x causing outcome φ" in terms of a contrast between the actual outcome φ and a counterfactual outcome φ' that would have resulted from x'. Here, φ' denotes a certain state of affairs unidentical to φ, and x' denotes a particular negation of x that is classified together with x as distinct possible values of the *same* endogenous variable X. As such, the meaning of statements such as "x (actually) caused φ" depends on the specification and classification of counterfactual events that undergird the two actual causal relata. In the language of SCM, such steps correspond to the presentation of the endogenous variables V and their possible values R.

The discussion in the second and third sections assumed V and R as given and mainly paid attention to F, that is, the step of inferring or justifying the functional entailment from x' to φ'. However, for the causal claim, "x (actually) caused φ," is the specification of x' and its assignment to X always a straightforward procedure?

To start with more obvious grounds that it isn't: In some cases, the "mere" absence of a certain historical event is undefined both logically and historically, thus calling for an explicit specification. To continue from the preceding example, South Korea's integration into the US-led Cold War regional architecture occurred simultaneously with that system's very emergence, and there was no circumstance in actual history where this regional political-economic system existed before and without South Korea's integration into it. For some elements of the system, such as economic aid, one may plausibly imagine zero value as the logical negation. But other elements such as security relationships or bloc membership do not come with an absolute "null" value, and the absence of one concrete relation necessarily presupposes another concrete relation. An explicit specification of the contrast classes can generally contribute to semantic clarity in such cases.

But ambiguities about contrast classes are not limited to historical antecedents with undefined absences. So long as the target antecedent x—properly specified in its theoretical and spatiotemporal dimensions (c.f. Runhardt 2022)—is an event or action that modularly occurs on top of a preexisting condition, we can schematically define x' either as its mere absence, representing whatever situation there was before x, or as events, institutions, or organizations that could have arisen instead of x to fulfill a comparable social position, settlement, or function. We will slightly abuse notation and denote the former as $\neg\, x$, with the $\neg$ sign now indicating absence. We will

refer to the latter as x^a, with the superscript "a" standing for "alternative." Given such a classification, two major modeling choices emerge. First, x may be grouped with $\neg\, x$ as different possible values of the same variable X while x^a is separated off to another variable W. Second, one may group the pre-empted event x^a with x and forgo variable W. Figure 3 schematically compares these two modeling choices.[8]

Normally, our causal intuition favors the former modeling option (Figure 3(a)), since our pretheoretical conception of historical causation or "process" is based on a *chronological transformation* from one state to another, that is, from an initial state to the putative cause and then to the observed outcome. If Suzy threw a rock, then the situation that existed before she threw would properly represent its contrast class. When it comes to expanding transportation networks, then the infrastructure that existed before the expansion would most intuitively represent its contrasting negation. If one political system was superseded by another one in a certain country, then the original system that existed just prior to the replacement would in many cases provide an intuitive contrast class. This intuition about historical causation crucially facilitates historiographical communication in practice and allows us to intersubjectively understand many historical causal arguments that only speak of actual events while leaving their underlying contrast classes implicit.

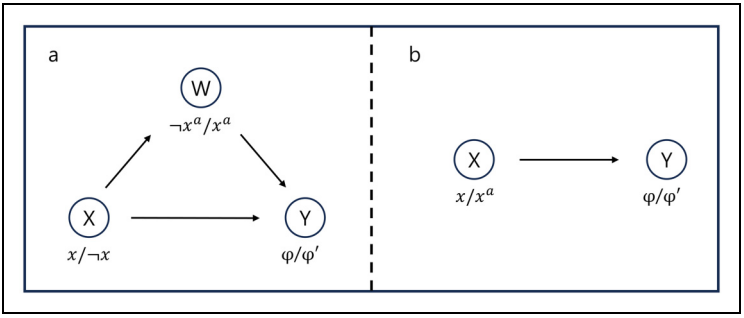

Figure 3. Two modeling choices with different implications for judgments of actual causality.
Note: The ¬ sign is used liberally here to denote the mere absence of a certain historical event, representing whatever institution or state of affairs that existed just prior to its occurrence. x^a denotes alternative developments that could have occurred in lieu of x. Values below the variables indicate possible values, with the one on the left being the actual.

Despite the strong intuitive appeal of Figure 3(a), we may sometimes be inclined to group x together with the pre-empted event x^a for judgments of actual causality following Figure 3(b). Such a modeling choice is prone to emerge when x is seen as an event that thwarted an object that was already in motion, or more generally, when one has a strong reason to take an unrealized course of history as natural, expected, or good. In such cases, historical processes may be derailed from their commonsensical understanding as a chronological transformation from one actual state to another and perceived in terms of a contrast between a "proper" or "natural" counterfactual state and the corresponding actual one.

Historical debates on colonialism and development in Korea again contain a relevant example. For historians sympathetic to the Korean standpoint, there are both historical and moral reasons to see colonialism as an abnormal turn of history that thwarted a projected or proper course of development. Their opponents also know this, and the possibility of causal semantics based on Figure 3(b) emerges seriously in historiographical communication.

Consider, for example, the debate between Kohli (1994, 1997) and Haggard, Kang, and Moon (1997), the former of which identifies Japanese colonialism as the origin of South Korean development and the latter of which opposes it. For Haggard et al. (1997), there was no "process" from colonialism to postcolonial development. While acknowledging that some forms of state-led capitalist development did happen under colonial rule, the authors describe how colonial capitalist development came with severe distortions such as ethnic discrimination and also stress the substantial institutional ruptures after decolonization that arguably undid much of colonialism's positive legacies.

However, parallel to these arguments, Haggard et al. (1997) also criticize Kohli's historical narrative by proposing to think in terms of counterfactuals (Haggard et al. 1997: 875). After acknowledging that "there are both firms and entrepreneurs who had their start in the interwar period," the authors state that "such firms would have emerged had Japan *not* occupied the country [emphasis original]" (p. 868). Unfortunately, the authors do not specify what they mean by the phrase "had Japan not occupied the country" and why they are suddenly proposing to think in terms of such counterfactuals. Are they insinuating an unrealized independent development of Korea that would have occurred had the Japanese not taken over? And if so, could it be that they proposing to define actual causality following the general scheme of Figure 3(b) instead of Figure 3(a) presumably implied by Kohli? Alternatively, are Haggard and colleagues merely proposing to think about the counterfactual while setting aside the possibility of pre-empted causes within Figure 3(a), merely following the logic of AC2?

We personally find the last option to be most consistent with the overall argument context of the debate between Haggard and colleagues and Kohli: the authors may not be hinting at any alternative path of Korean development but are merely proposing to think what would have been the result if Japan had not developed Korea and Korea's level of development remained near precolonial levels until 1945. This would be consistent with their repeated argument that Japan's contributions were unimpressive and mostly did not outlast colonial rule. However, such an interpretation is not self-evident, not only because the authors do not accurately describe the counterfactual intervention that undergirds the causal antecedent, but also because the other interpretive option just mentioned (i.e., Figure 3(b)) is already strongly insinuated by a sector of Korean "nationalist" scholarship. This body of work criticizes the so-called Colonial Modernization Thesis by arguing that Japanese colonialism *thwarted*, rather than promoted (i.e., caused), Korean development, in the sense that Korea was arguably developing rapidly and successfully under the late reformist Korean court unlike the distorted form of development that accompanied colonial rule (Korea Education Development Institute 1994, 1997; Lee 2000). Indeed, Kohli (1997), in his response article to Haggard et al. (1997), understood their proposal about historical counterfactuals as pointing to a hypothetical state of Korean indigenous development, which he quickly dismissed as being overly speculative and historically unrealistic (Kohli 1997: 886).

Presenting a simple structural causal model or even merely describing what counterfactual state of affairs constitutes the intervention setting in the main causal variable and what events should be seen as pre-empted ones could greatly reduce the risk of miscommunication in such argumentative contexts. One could, for example, describe a causal model akin to Figure 3(a) applied to this specific historical context, thereby ruling out the semantic possibility of Figure 3(b). Given Figure 3(a), the claim that Japan heavily restricted the development of a Korean managerial class, for example, would directly serve the argument that the causal effect of X (the developmental experience under colonial rule did or did not exist) on Y (post-colonial growth spurt did or did not occur) while holding the pre-empted event W pegged to its actual value $\neg\, x^a$ may be null or at least smaller than what has been argued. The critic would then be prompted to point to another antecedent $Z = z$ that must really have caused Y, which directly leads to the dual-antecedent setup of Figure 2 and the methodological steps suggested in Table 1.

There may or may not be the need to describe the pre-empted event $W = w$ in more detail depending on what Haggard and colleagues intend by their

proposal to think about counterfactuals. If they merely aim to reject the actual-causal relationship between Japanese colonial development and post-colonial South Korean growth spurt, then W (such as indigenous Korean development) may be harmlessly left underspecified in the causal model as it will need to be bracketed off. On the other hand, if they also intend to take a more adventurous path and seek to argue that an alternative path of Korean development was possible had it not been for colonialism, they will need to describe that $W = w$ in sufficient theoretical, historical, and spatiotemporal specificity and infer what consequences it would have had for Korean development in the second half of the twentieth century.

Conclusion

Causal analysis grounded on counterfactual "interventions" has seen widespread adaptation across the social and behavioral sciences in the past couple of decades. While the bulk of their application has taken the form of forward-looking, type-based causal analysis used in statistical methods, counterfactual-based theories can be effectively applied to analyses of backward-looking historical causation (Halpern 2016), and recent methodological studies in sociology have also advocated their use for this purpose (Runhardt 2022). The language of Structural Causal Modeling—familiar to many sociologists from causal analysis in statistics—provides a powerful tool for historical causal analysis in the interventionist framework.

Accepting a counterfactual-based *theory* of historical causality makes it prima facie natural to integrate counterfactual reasoning into the *practice* of historiography. In response to the thoroughgoing methodological empiricism of disciplinary historiography, we argue that the systemic application of counterfactual methods can be contained within specific types of situations that give rise to historiographical ambiguities or disagreements. We have sought to identify such situations for each of the two logical components of historical causality, which we roughly labeled as semantics and inference. The former concerns describing and then grouping various nomadic events or circumstances—both actual and counterfactual—into paired bundles or "variables" that populate a causal model, thereby designating, for instance, what counterfactual event should form the latent causal contrast of the actual antecedent. The latter involves inferring or justifying the outcomes that would have resulted from the designated causal contrast under whatever logical contingencies required by the causal query.

For semantics, we pointed to the standard intuition to take the "mere" absence of the historical antecedent as its negation or contrast class. Such

intuitions can be unproblematically implemented when the event under consideration exhibits logical modularity, which typically also means that one may simply refer to whatever historical state of affairs that existed prior to its occurrence. As such, our mundane idea about historical causation can be intuitively represented as a chronological transformation from one historical state to another, which straightforwardly converges into the common metaphor of "process."

An explicit specification of the variables and their possible actual and counterfactual values are practically useful in cases where the modular absence of a certain event is undefinable or because the modular absence of the event may, for special reasons, not intersubjectively represent the proper negation of the event. Paying particular attention to the latter case, we argued that we may sometimes be prompted to contrast an actual event to a pre-empted event that would have occurred in its stead. Such impulses arise when we have a strong reason to see the pre-empted event as good or natural, either because the situation was already dynamically changing toward it or because one prefers it for nonepistemic reasons. In such cases, actual-causal claims in history, generally expressed in the form that event x "caused" or "brought about" or "is the origin" of event y, are prone to misunderstanding at a semantic level between scholars of opposing political-historical orientation.

For example, if it is true that South Korea was developing or was on the verge of developing very successfully on its own prior to colonization as some nationalist historians claim, then should we say that Japanese colonialism developed (or caused the development of) Korea for what it has actually done? Should we instead not say that it thwarted it? The assumptions of a dynamic pre-colonial developmental trajectory of Korea, coupled with political commitments about colonialism, fumbles at least some people's pre-theoretical understanding of what it means for one historical event to cause another, and explicitly translating the actual-causal claim in terms of underlying counterfactuals can greatly reduce the risk of miscommunication.

Regarding the many situations where we may intuitively identify the logical negation of an actual event as its mere absence akin to Figure 3(a), inferring actual causality amounts to inferring that the outcome would not have occurred if the event had not occurred, *and* none of the events of a similar class that could have occurred in its stead occurred. While inferring the progression of counterfactual events, in principle, constitutes a fundamental methodological challenge for causal inference, bracketing off pre-empted events this way greatly simplifies inference, which, in line with the common understanding that historical causation occurs as a transformation or 'process'

from one event to another, often makes it seem as we are seeing a connection between two actual events. The widespread anxiety about counterfactuals among historiographical practitioners can be inculpably toned down once we acknowledge this crucial distinction between the contingent counterfactuals that ground judgments of actual causality and the "free-floating" counterfactuals that must account for alternative paths of historical development in the absence of any guardrails provided by actual history.

Notwithstanding the general ease of judgments of actual causality compared to unconditioned counterfactuals, historiography still has no shortage of disagreements about actual causality, not only because of miscommunication about meaning but more fundamentally because of disagreements about inference. We noted that difficulties in inference are prone to occur in situations where two or more antecedents expected to play a similar role are competing for causal explanation. Likewise, difficulties in inferring the causal status of a single particularly controversial historical antecedent are often prompted by the existence of another antecedent that could also have served a similar role. We argued that sequential counterfactual reasoning with two antecedents—as schematized in Table 1—may offer methodological guidelines in these situations that fumble our pretheoretical intuition about historical causation. It may do so by designating what transformative forces would still remain even in the absence of one antecedent and explicitly directing the scholar to arbitrate this tension in the counterfactual query. In addition, it allows the scholar not only to check for (full) causes but also parts of causes in overdetermination scenarios—a situation that greatly challenges our mundane ability to understand actual–causal relationships.

Finally, a comment about the use of formal models is in order. In case the scholar decides to explicate counterfactuals for describing the variables and their possible values or for making inferences from those variables, should she use the technical language of Structural Causal Models, possibly using the directed acyclical graphs like the ones used in Figures 1 and 3 displaying the list of possible values for each variable, or is it sufficient to merely express them in plain words without such graphs? It may seem that the suggestion to introduce such "formal" methods into historiography may represent an overly large deviation from any actual or potential historiographical practice. However, we argue this distinction is not too important, for causal graphs are merely visually aided representations of the list of antecedents, their latent contrast classes, and relationships of pre-emption that can be unproblematically, albeit more verbosely, described in plain words. In potentially using such graphical representations, the scholar may selectively focus on a local part of the entire causal chain (i.e., only a part of the entire historical

process that she wishes to present) that invokes special concern about meaning or inference.

Author's Note

Department of Sociology, Sungkyunkwan University, South Korea.

Declaration of Conflicting Interests

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

Funding

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

ORCID iD

Tay Jeong https://orcid.org/0000-0002-1018-2923

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Notes

1. Nothing in this article depends on whether there are noncausal explanations of token events (Skow 2014). We are concerned about historiography as a mode of writing that offers a causal explanation of particular social events. If a certain work of historiography explains a social phenomenon without identifying its causes, then it lies outside the scope of the present article.
2. See also the statement by Walther Rathenau (1918: 82), "history does not conjugate in conditionals, it talks about what is and was, not what would and would have been."
3. Halpern (2016) contains a detailed introduction to the SCM in the context of actual causality. Blanchard and Schaffer (2017) offers a gentle and abridged introduction.
4. We would like to thank an anonymous reviewer for pointing out that φ in Halpern's definition is not necessary an "event" but a sentence.
5. This sentence and AC1–AC3 are directly quoted from Halpern (2016: 23–25).
6. It should be noted that this strategy logically presumes transitivity in historical causation, such that if A causes B and B causes C, then A causes C. There are

good reasons to think that transitivity is not a necessary feature of causation (McDermott 1995); however, most proposed examples of causal nontransitivity presented in the philosophical literature feature highly artificial scenarios that are scarcely relevant to historiography.

7. While we find Asada's use of historical counterfactuals a methodological strength, we do not intend to suggest that his arguments are necessarily correct. See Akagi and Takita (2016) for a review of rebuttals that have been published in response to Asada's influential contribution to the "orthodox" view.
8. Note that the choice between Figure 3(a) and (b) places different formal constraints on jointly possible interventions. When x and x^a are modeled as alternative states of the same variable as in Figure 3(b), then a joint intervention that realizes both events is precluded by the rules of the SCM. This is formally possible in Figure 3(a) where x and x^a are separated into distinct variables. In practice, historical (rather than formal) logic precludes such a joint intervention in Figure 3(a) as well.

References

Akagi, Kanji and Ryosuke Takita. 2016. "Japan's Decision to Surrender: A Historiographical Review." *Journal of Law, Politics, and Sociology* [法學研究: 法律・政治・社会] 89(9):1-43. In Japanese.

Alperovitz, Gar. 1965. *Atomic Diplomacy: Hiroshima and Potsdam; the Use of the Atomic Bomb and the American Confrontation with Soviet Power*. New York, NY: Vintage Books.

Amsden, Alice. 1989. *Asia's Next Giant: South Korea and Late Industrialization*. New York: Oxford University Press.

Ankersmit, Frank. 2012. *Meaning, Truth, and Reference in Historical Representation*. Ithaca, NY: Cornell University Press.

Asada, Sadao. 1998. "The Shock of the Atomic Bomb and Japan's Decision to Surrender: A Reconsideration." *Pacific History Review* 67(4):477-512.

Bennett, Andrew. 2010. "Process Tracing and Causal Inference." " Pp. 207-220 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards* 2nd ed., edited by H. E. Brady and D. Collier. UK: Rowman & Littlefeld.

Blanchard, Thomas and Jonathan Schaffer. 2017 "Cause Without Default." Pp. 175-214 in *In Making a Difference: Essays on the Philosophy of Causation*, edited by Helen Beebee, Christopher Hitchcock, and Huw Price. Oxford: Oxford University Press.

Bloch, Marc. 1953. *The Historian's Craft*. New York, NY: Vintage Books.

Bunzl, Martin. 2004. "Counterfactual History: A User's Guide." *The American Historical Review* 109(3):845-58. doi: 10.1086/530560.

Carr, E. H. 1990. *What Is History?* 2nd ed. London: Penguin Books.

Chibber, Vivek. 2003. *Locked in Place: State-Building and Late Industrialization in India*. Princeton, NJ: Princeton University Press.

Clarke, Christopher. 2023 "Process Tracing: Defining the Undefinable?" Pp. 305-27 in *The Oxford Handbook of Philosophy of Political Science*, edited by H. Kincaid and J. Van Bouwel. Oxford, UK: Oxford University Press.

Dowding, Keith. 2023 "Process Tracing: Causation and Level of Analysis." Pp. 328-42 in *The Oxford Handbook of Philosophy of Political Science*, edited by H. Kincaid and J. Van Bouwel. Oxford, UK: Oxford University Press.

Eckert, Carter J. 1991. *Offspring of Empire: The Kochang Kims and the Colonial Origins of Korean Capitalism, 1876–1945*. Seattle and London: University of Washington Press.

Ermakoff, Ivan. 2019. "Causality and History: Modes of Causal Investigation in Historical Social Sciences." *Annual Review of Sociology* 45(1):581-606.

Fearon, James D. 1991. "Counterfactuals and Hypothesis Testing in Political Science." *World Politics* 43(2):169-95. doi: 10.2307/2010470.

Ferguson, Niall. 1999 "Virtual History: Towards a 'Chaotic' Theory of the Past." Pp. 1-90 in *Virtual History: Alternatives and Counterfactuals*, edited by N. Ferguson. New York, NY: Basic Books.

Fogel, Robert. 1964. *Railroads and American Economic Growth: Essays in Economic History*. Baltimore, MD: The Johns Hopkins University Press.

Froeyman, Anton. 2009. "Concepts of Causation in Historiography." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 42(3):116-28. doi: 10.3200/HMTS.42.3.116-128.

Gaddis, John. 2002. *The Landscape of History: How Historians Map the Past*. New York, NY: Oxford University Press.

Goldstone, Jack. 2003 "Comparative Historical Analysis and Knowledge Accumulation in the Study of Revolutions." Pp. 41-90 in *Comparative Historical Analysis in the Social Sciences*, edited by J. Mahoney and D. Rueschemeyer. Cambridge, UK: Cambridge University Press.

Haggard, Stephan, David Kang, and Chung-In Moon. 1997. "Japanese Colonialism and Korean Development: A Critique." *World Development* 25(6):867-81. doi: 10.1016/S0305-750X(97)00012-0.

Halpern, Joseph. 2015 "A Modification of the Halpern-Pearl Definition of Causality." Pp. 3022-33 in *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*. AAAI Press

Halpern, Joseph. 2016. *Actual Causality*. Cambridge, MA: The MIT Press.

Halpern, Joseph Y. and Christopher Hitchcock. 2015. "Graded Causation and Defaults." *The British Journal for the Philosophy of Science* 66(2):413-57. doi: 10.1093/bjps/axt050

Halpern, Joseph Y. and Judea Pearl. 2005. "Causes and Explanations: A Structural-Model Approach. Part I: Causes." *The British Journal for the Philosophy of Science* 56(4):843-87. doi: 10.1093/bjps/axi147

Heo, Soo Youl. 2012. "Korea's Liberation in 1945 and its Economic Development." *Journal of Korean Independence Movement Studies* [한국독립운동사연구] 43:463-509. In Korean.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81(396):945-60.

Hume, David. 1748. *An Enquiry Concerning Human Understanding*.

Isaac, Larry W., Debra A. Street, and Stan J. Knapp. 1994. "Analyzing Historical Contingency with Formal Methods." *Sociological Methods & Research* 23(1):114-41.

Kim, Naknyeon. 2012 "Janggitonggye: Haebang Huwaui Yeongyeol." Pp. 295-332 in *Hanguk Ui Janggitonggoe: Gukmin Gyejeong [National Accounts of Korea 1911-2010]*, edited by Naknyeon Kim. Seoul, South Korea: Seoul National University Press. In Korean.

Kohli, Atul. 1994. "Where Do High Growth Economies Come From? The Japanese Lineage of Korea's 'Developmental State.'." *World Development* 22(9):1269-93.

Kohli, Atul. 1997. "Japanese Colonialism and Korean Development: A Reply." *World Development* 25(6):883-88. doi: 10.1016/S0305-750X(97)00009-0.

Kohli, Atul. 2004. *State-Directed Development: Political Power and Industrialization in the Global Periphery*. Cambridge: Cambridge University Press.

Korean Educational Development Institute. 1994. *100 Years of Modern Korean School Education(I)—School Education in the Enlightenment Period*. RR 94-7. Jincheon, Republic of Korea: Korean Educational Development Institute. In Korean.

Korean Educational Development Institute. 1997. *100 Years of Modern Korean School Education(II)—The Japanese Occupation Period*. RR 97-10. Jincheon, Republic of Korea: Korean Educational Development Institute. In Korean.

Krieckhaus, Jonathan. 2017. "Geopolitics and South Korea's Economic Success." *Asian Perspective* 41(1):43-69. doi: 10.1353/apr.2017.0003.

Lee, Tae-jin. 2000. *Gojongsidaeui Jaejomyeong*. Paju, Republic of Korea: Taehaksa. In Korean.

Levy, Jack S. 2015. "Counterfactuals, Causal Inference, and Historical Analysis." *Security Studies* 24(3):378-402. https://doi.org/10.1080/09636412.2015.1070602.

Lewis, David. 1973. "Causation." *The Journal of Philosophy* 70(17):556-67.

Mahoney, James. 2000. "Path Dependence in Historical Sociology." *Theory and Society* 29(4):507-48.

Mahoney, James and Gary Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14(3):227-49. doi: 10.1093/pan/mpj017.

Mandelbaum, Maurice. 1977. *The Anatomy of Historical Knowledge*. Baltimore, MD: The Johns Hopkins University Press.

McDermott, Michael. 1995. "Redundant Causation." *British Journal for the Philosophy of Science* 46:523-44.

McNamara, Dennis L. 1990. *The Colonial Origins of Korean Enterprise 1910-1945*. New York, NY: Cambridge University Press.

Nishijima, Ariatsu. 1968. *Genbatsu Wa Naze Touka Saretanoka: Nihon Koufuku Wo Meguru Senryaku to Gaikou*. Tokyo, Japan: Aoki Shoten Publishing Co. In Japanese.

Oakeshott, Michael. 1985. *Experience and Its Modes*. New York, NY: Cambridge University Press.

Pilat, Dirk. 1994. *The Economics of Rapid Growth: The Experience of Japan and Korea*. Aldershot, UK: Edgar Elgar.

Ranke, Leopold von. 1874. *Geschichten Der Romanischen Und Germanischen Völker von 1494 Bis 1514*. Leipzig: Duncker und Humblot.

Rathenau, Walther. 1918. *Die Neue Wirtschaft*. Berlin: S. Fischer Verlag.

Reiss, Julian. 2012 "Counterfactuals." Pp. 154-183 in *The Oxford Handbook of Philosophy of Social Science*, edited by H. Kincaid. Oxford: Oxford University Press.

Rohlfing, Ingo and Christina Isabel Zuber. 2021. "Check Your Truth Conditions! Clarifying the Relationship Between Theories of Causation and Social Science Methods for Causal Inference." *Sociological Methods & Research* 50(4):1623-59. doi: 10.1177/0049124119826156.

Rosenfeld, Gavriel. 2002. "Why Do We Ask 'What If?' Reflections on the Function of Alternate History." *History and Theory* 41(4):90-103.

Runhardt, Rosa W. 2022. "Concrete Counterfactual Tests for Process Tracing: Defending an Interventionist Potential Outcomes Framework." *Sociological Methods & Research* 53 (4): 1591-628. doi: 10.1177/00491241221134523.

Scheidel, Walter. 2019. *Escape from Rome: The Failure of Empire and the Road to Prosperity*. Princeton, NJ: Princeton University Press.

Skow, Bradford. 2014. "Are There Non-Causal Explanations (of Particular Events)?" *The British Journal for the Philosophy of Science* 65(3):445-67.

Stubbs, Richard. 1999. "War and Economic Development: Export-Oriented Industrialization in East and Southeast Asia." *Comparative Politics* 31(3):337. doi: 10.2307/422343.

Tucker, Aviezer. 2004. *Our Knowledge of the Past: A Philosophy of Historiography*. Cambridge: Cambridge University Press.

Tucker, Aviezer. 2008 "Causation in Historiography." Pp. 98-108 in *A Companion to the Philosophy of History and Historiography*, edited by H. Kincaid. Chichester, United Kingdom:Wiley-Blackwell.

Waldner, David. 2012. "Process Tracing and Causal Mechanisms." Pp. 65-84 in *The Oxford Handbook of Philosophy of Social Science*, edited by H. Kincaid.

White, Hayden. 1980. "The Value of Narrativity in the Representation of Reality." *Critical Inquiry* 7(1):5-27.

Yang, Jonghoe. 2004. "Colonial Legacy and Modern Economic Growth in Korea: A Critical Examination of Their Relationships." *Development and Society* 33(1):1-24.

Author Biography

Tay Jeong is a sociologist with broad-based research interests in international conflict and social philosophy. He obtained a Ph.D. degree at McGill University and is currently an SSHRC postdoctoral fellow at Kangwon National University, where he conducts research on Korean politics and society. He is particularly interested in thinking about contemporary political issues using historical or philosophical methods.