

NOTICE

This material may be protected by
Copyright Law (Title 17 U.S. code)

In Janet Box-Steffensmeier, Henry Brady, and David Collier, eds.,
Oxford Handbook of Political Methodology. New York:
Oxford University Press, 2008. Pp. 627-44.

CHAPTER 27

COUNTERFACTUALS
AND CASE STUDIES

JACK S. LEVY

In Henry James's "The Jolly Corner," Spencer Bryden pursues the ghost of the man he might have been had he not left New York City three decades earlier for a more leisurely life abroad. In Dickens's "A Christmas Carol," Ebenezer Scrooge transforms his life after encountering the Ghost of Christmas Yet to Come and learning how the future would play out if he were to continue his old ways. Robert Frost is less explicit about what he would have encountered on "the road not taken," but he knows that the road less traveled "made all the difference."

Scholars frequently speculate about what might have been in history. Pascal famously wrote that "Cleopatra's nose: had it been shorter, the whole face of the world would have changed." It is often said that the First World War would not have occurred without the assassination of Archduke Ferdinand (Lebow 2007), that the Second World War would not have occurred without Hitler (Mueller 1991), and that the end of the cold war would have been significantly delayed without Gorbachev (English 2007). In a more general theoretical claim, Skocpol (1979) argues that without either peasant revolts or state breakdown, social revolutions will not occur.

Opposition politicians frequently invoke counterfactuals in arguing that if the current administration had acted differently, the country would have been better off. Hillary Clinton later defended her 2002 vote to authorize the president to use force in Iraq by saying that "if we knew then what we know now, there wouldn't have been a vote ... and I certainly wouldn't have voted that way."¹

¹ NBC, *Today show*, December 18, 2006.

Some historians are skeptical about the analytic utility of counterfactual analysis, and regard it as entertaining “after dinner history” (Ferguson 1999a, 15) or a “parlour game” (Carr 1964, 97), but not analytically sound scholarship. Croce (cited in Ferguson 1999a, 6) states that it is necessary “to exclude from history the ‘conditional’ which has no rightful place there. . . .” What is forbidden is . . . the “anti-historical and illogical ‘if.’” Fischer (1971) includes an index item for “counterfactual questions” in his book *Historians’ Fallacies*, but it says “see fictional questions.” Fischer describes Fogel’s (1964) pathbreaking counterfactual analysis of the impact of railroads on American economic development as a step “down the methodological rathole” and a return to “ancient metaphysical conundrums” (p. 18). Oakeshott (1966, 128–9) argues that if the historian were to consider “what *might* have happened” and treat “great events” or “turning points” as causally “decisive,” “the result is not merely bad or doubtful history, but the complete rejection of history . . . a monstrous incursion of science into the world of history.”

Other historians, and most social scientists, recognize that counterfactuals are unavoidable. They understand that “the study of history is a study of causes” (Carr 1964, 87), and they recognize that any causal statement involves assumptions about what did not happen but could have happened. Bueno de Mesquita (1996, 229) argues that in applied game theory “we cannot understand what happened in reality without understanding what did not happen but might have happened under other circumstances.” The historian Ferguson (1999a, 87) argues that “To understand . . . [history] as it, actually was, we therefore need to understand *how it actually wasn’t*—but how, to contemporaries, it might have been.”

The question is how to validate counterfactual claims about what would have happened in a hypothetical or alternative world in which the hypothesized cause took on a different value, whether in a particular case or in a more general theoretical relationship. Whereas causal statements are in principle amenable to a direct empirical test, the same is not true for counterfactual statements, since the conditional upon which the counterfactual rests does not exist and cannot be fully realized in order to examine the effects that flow from it (Goodman 1983). In the absence of direct empirical confirmation, by what criteria can we say that some counterfactuals are more legitimate or more valid than others, and for what theoretical purposes? Since we cannot avoid counterfactuals, the question, in response to Oakeshott, is how to introduce science into the world of history in a way that enhances our understanding of history.

I focus on the methodologically normative use of counterfactual arguments to advance causal understanding of the social and political world. This differs from psychologists’ more descriptive focus on such cognitive science questions as how people actually use counterfactuals, how they judge the validity of counterfactual arguments, and how their cognitive and motivational biases affect those judgments and influence what kinds of counterfactuals they find most persuasive (Roese and Olson 1995; Tetlock and Belkin 1996; Tetlock and Lebow 2001).

Counterfactuals are relevant in any kind of causal analysis, but I focus primarily on the role of counterfactuals in case study analysis. This subject is particularly

important because qualitative-comparative researchers are more likely than quantitative scholars to posit necessary conditions, which automatically generate explicit counterfactuals (Goertz and Starr 2003).

After explaining why counterfactuals are important for history and social science, I identify the criteria by which we can evaluate the utility of counterfactuals for supporting idiographic and nomothetic causal claims. Throughout I assume that counterfactuals are best conceived as a method, to be used in conjunction with other methods, to generate and validate explanations about social and political behavior. The primary goal of studying what did not happen is to better help us to understand what did happen.

1 THE IMPORTANCE OF COUNTERFACTUALS

I follow Tetlock and Belkin (1996, 4) and define counterfactual as a “subjective conditional in which the antecedent is known or supposed for purposes of argument to be false.” It is a “contrary-to-fact” conditional that identifies a “possible” or “alternative” world in which the antecedent did not actually occur.

All causal statements imply some kind of counterfactual. A historical argument that a particular set of conditions, processes, and events caused, influenced, or contributed to a subsequent set of conditions, processes, and events implies that if the antecedent conditions had been different, the outcome would have been different. Similarly, a theoretical statement that x is a cause of y implies that if the value of x were different, the outcome y would be different.² The interpretative statement that British and French appeasement of Hitler contributed to the Second World War implies that if Britain and France had stood firm against Hitler the Second World War would not have occurred, or perhaps that it would have been considerably shorter and less destructive. The theoretical proposition that appeasement only encourages further aggression implies the counterfactual that a more hard-line strategy would reduce the likelihood of further aggression.

While all causal statements imply a counterfactual, some counterfactuals are more explicit than others. Historical interpretations and theoretical propositions that posit necessary conditions, which are fairly common in political science and in social science more generally (Goertz and Starr 2003), are particularly explicit about their counterfactual implications. The logical expression of necessary conditions—“if not x then not y ”—directly specifies the consequent of a counterfactual world. Necessary condition counterfactuals are central in “window of opportunity” models (Kingdon 1984), “powder keg” models (Goertz and Levy 2007),

² This is particularly clear if one says that c is a cause of e if the conditional probability that e occurs, given that c occurs, is greater than the unconditional probability that e occurs. For a discussion of statistical approaches to counterfactual analysis see Morgan and Winship (2007).

causal chains involving necessary conditions, and explanations based on path dependency and critical junctures (Mahoney 2000), all of which are common in political science.

Necessary conditions are also important because they are at the core of one of the two leading conceptualizations of causation: *x* is a cause of *y* if *y* would not occur in the absence of *x* (Lewis 1973).³ The central role of necessary conditions in causal explanation, particularly in case studies, and the fact that the primary way of testing hypotheses involving necessary conditions is to analyze the validity of the counterfactual associated with it, add significantly to the importance of developing a valid methodology of counterfactual analysis.⁴

Counterfactuals play an essential role in common historiographical debates about whether a particular outcome was inevitable or contingent. The most effective way of supporting an argument that an outcome was contingent is to demonstrate that a slight change could easily have led to a different outcome. An effective way to demonstrate that an outcome was inevitable at a certain point is to demonstrate that there was no change in existing conditions that was both conceivable and capable of leading to a different outcome.

Some theoretical approaches are more explicit than others about their counterfactual implications. A good example is game theory. A game tree specifies exactly what happens if actors make different choices, how other actors react, and the set of possible outcomes. Actors' choices are causally dependent on their expectations of what would happen if they made other choices. "Off the equilibrium path" behavior is a counterfactual prediction.

Since causal statements—whether about particular historical events or about theoretical relationships between variables—imply counterfactuals, the validity of the counterfactual bears on the validity of the original causal proposition. Compelling evidence that Hitler would have initiated a war regardless of whether Britain or France had pursued a strategy of appeasement would falsify the claim that appeasement causally contributed to the Second World War.⁵ Evidence that social revolutions sometimes occur in the absence of a peasant revolt and a state crisis would disconfirm Skocpol's (1979) theory of social revolution.

Thus the empirical validation of counterfactual statements is an important step in hypothesis testing. Evidence from several case studies is generally more persuasive than evidence from a single case study in testing a theoretical proposition

³ The alternative conception of causation is the "regularity" model, which includes constant conjunction, temporal precedence, and nonspuriousness. This view is often traced to Hume, but in fact Hume emphasized both conceptions of causation (Goertz and Levy 2007).

⁴ While all necessary conditions imply a specific counterfactual, not all counterfactuals imply the presence of necessary conditions. The statement "if not *x*, then *y* will still occur" posits a counterfactual, but without a necessary condition. Also, unlike statements of necessary conditions, statements of sufficient conditions do not imply a counterfactual, since "if *x* then *y*" says nothing about the consequences of "not *x*." The widely accepted empirical generalization that a democratic dyad is sufficient for peace implies no counterfactual about what would happen if at least one of two states is not a democracy.

⁵ Khong (1996) explores the complexities of validating this counterfactual. Ripsman and Levy (2007) argue that neither British nor French leaders expected that appeasement would avoid war.

(and quantitative researchers argue that statistical evidence is far superior to multiple case studies), but single case studies can also serve this purpose if the hypothesis posits necessary or sufficient conditions, if the hypothesis generates precise predictions, or if the proposition permits a most likely or least likely research design (George and Bennett 2005; Levy 2008).

This point relates to a larger principle about theory construction and research design. The widely accepted injunction to derive as many observable implications as possible from an historical interpretation or theoretical proposition, and to test them against the evidence (King, Keohane, and Verba 1994), applies to a theory's counterfactual implications as well as to its more direct implications.⁶ *Ceteris paribus*, the more explicit the counterfactual implications of a theory, the better the theory. A theory that specifies the consequences of both x and not x tells us more about the empirical world than a theory that specifies only the consequences of x .

The preceding discussion emphasizes the utility of counterfactuals in testing both interpretations of individual cases and more general theoretical propositions. The first is more idiographic in orientation, and the second more nomothetic. Interpretations of individual cases can be either inductive or guided by theory.⁷ Although theory-guided case studies are presumably more explicit about their counterfactual predictions, at least at the beginning of a research study,⁸ inductively driven interpretive case studies can also end up suggesting what would have happened if certain things had been different. Ferguson's (1999b) argument about what would have happened had Britain "stood aside" in the First World War (probably no Bolshevik Revolution, no Second World War, no Holocaust) is based on an inductive case study.

In idiographic case studies, the role of counterfactual analysis is often to explore the question of whether history could have turned out differently, and how alternative worlds might have developed. Such "what if" scenarios include both relative short-term forecasts, such as the consequences of a failed assassination attempt in 1914, and long-term macrohistorical forecasts, such as the plausibility of alternative scenarios through which the rise of the West to a position of world dominance might have been blocked (Tetlock, Lebow, and Parker 2006).

Idiographic counterfactuals can also play a normative role—of passing moral judgment on individual leaders by asking whether or not they could have acted differently under the circumstances (Tetlock and Belkin 1996, 8). Presumably we could not blame a leader for a costly war if a counterfactual analysis were to provide persuasive evidence that the same war would have occurred if another leader had been in power. Neville Chamberlain is widely criticized for his policy of appeasing Hitler, but that

⁶ Qualitative researchers qualify this argument by de-emphasizing the "as many" phrase and arguing that some observable implications are more important than others for theory testing.

⁷ Scholars often mischaracterize the idiographic/nomothetic distinction. Idiographic refers to descriptions or explanations of the particular, whereas nomothetic refers to the construction or testing of general theoretical propositions (Levy 2001).

⁸ In the analytic narrative research program (Bates et al. 1998), for example, game-theoretic models are used to guide individual case studies.

judgment rests on the counterfactual argument that another British leader would have acted differently and that the outcome would have been better for Britain, which some scholars have questioned.

Exploring a theory's counterfactual implications can also be useful for theory development. This is a deductive analysis that does not involve empirical case studies.⁹ As Tetlock and Belkin (1996) note, "The goal is not historical understanding... [but] to pursue the logical implications of a theoretical framework." For this purpose, the counterfactual conditional itself need not necessarily be plausible, in the sense that we could imagine a path through which it might arise. Such nonplausible counterfactuals are often called "miracle counterfactuals" (Fearon 1996).

Any complete theory specifies the consequences if key causal variables were to take on other values. Economic models specify the consequences for the economy if the Federal Reserve Board were to raise interest rates by a full percent, even if such actions are quite implausible under the circumstances. Computer simulations are used to explore the consequences of possible worlds (including socially complex worlds) that are not really accessible through empirical methods (Cederman 1996). Similar models can be applied historically to trace the consequences of a particular set of initial conditions, however implausible they might be. To assess the contribution of the railroads to American economic growth, Fogel (1964) constructed a model of the American economy and explored how it might have developed in the absence of railroads.

Counterfactual thought experiments can also be used deductively for other purposes, including the pedagogical purposes of encouraging someone to think through the implications of his or her arguments or beliefs, confront uncomfortable arguments, and generally open his or her mind to new ideas. This may help reveal double standards in moral judgment, contradictions in causal beliefs, and the influence of cognitive or motivational biases. This mode of counterfactual analysis combines the descriptive and the normative, in that descriptive/causal knowledge about how people use counterfactuals in making judgments and decisions is used for the normative purpose of inducing them to think in less biased ways about causality and counterfactuals (Tetlock and Belkin 1996, 12–16; Weber 1996; Lebow 2000).

2 CRITERIA FOR EVALUATION

We now turn to the criteria by which we might evaluate the validity of counterfactual arguments in explaining cases or testing theoretical propositions. What kinds of counterfactuals are more legitimate than others for these purposes, recognizing

⁹ In mathematics, an "indirect proof" or "proof by contradiction" assumes the theorem is false, traces the logical consequences, reveals a contradiction, and concludes that the theorem is true.

that different standards might be appropriate for different theoretical and descriptive purposes?¹⁰

Counterfactual propositions are similar to other theoretical propositions in that they have premises or initial conditions, hypothesized consequences, and a “covering law” or causal mechanisms explaining how the former leads to the latter. A theoretical proposition must be logically consistent and falsifiable in principle. We generally prefer theories that apply to broader empirical domains, make more (and more varied) predictions about the empirical world, generate substantial support from the empirical evidence, and are consistent with other well-accepted theories (Hempel 1966, ch. 4). Criteria for evaluating counterfactual arguments should be grounded in these widely accepted standards of theory evaluation. These are best organized around the categories of the clarity of the antecedents and consequents, the plausibility of the antecedent, and the conditional probability of the consequent given the antecedent.¹¹

2.1 Clarity

Counterfactual arguments must include well-specified antecedents and consequents. The consequent should be more specific than “the outcome would have been different,” which is only moderately helpful. Unless we specify *how* it would have been different, the statement is nearly nonfalsifiable and hence not particularly useful. For a counterfactual to be scientifically useful, the consequent must be clearly specified by the analyst, not left to the imagination of the reader.

Necessary condition counterfactuals are quite explicit, since they posit an antecedent of $\sim x$ and a consequent of $\sim y$. To say that war would not have occurred in the absence of an assassination is a powerful statement. It would be more discriminating, and in many ways more useful, to specify whether the absence of war meant peace for several years or peace punctuated by ongoing crises and a high risk of war in the future. With respect to the antecedent, it might have made a difference, in terms of the likely response, if the failed assassination attempt had been discovered or not. Non-necessary condition counterfactuals can also be clear. The statement “if Hitler had been ousted in a coup, the Second World War would still have occurred” has well-specified antecedents and consequents.

While clarity is good, there is a trade-off between the specificity of the consequent and the probability of its occurrence. The more detailed a consequent, the more likely it is to be false. The probability of multiple outcomes is the product of their individual probabilities, and the probability of a single outcome is the product of the individual probabilities of each of the steps leading to it (Lebow 2000, 583). This suggests that

¹⁰ The main difference between empirically oriented and deductively oriented counterfactuals is that the latter are not constrained by need for the antecedent to be plausible.

¹¹ This differs somewhat from the Tetlock and Belkin (1996, 16–31) and Lebow (2000, 577–85) conceptualizations.

counterfactuals that are either too detailed or the result of a long causal chain are likely to be false. Thus at the two extremes counterfactuals are either nonfalsifiable or false. Neither is particularly useful.¹² The best counterfactuals have specific but not too specific antecedents.

Relatedly, arguments suggesting that a particular counterfactual antecedent will lead to a particular consequent with a high degree of certainty are quite suspect. It would be problematic to invoke a counterfactual to demonstrate that history did not have to happen the way it did, only to predict a counterfactual consequent that deterministically flowed from the antecedent, since that would be supporting a statement of contingency with an argument based on determinism.

2.2 Plausibility of the Antecedent

It is widely agreed that for counterfactual arguments to be useful in exploring whether history might have turned out differently, the antecedent must be plausible or realistic as well as well specified. One must be able to imagine how that antecedent might arise. It is not useful to say that 1960s America would have been different if Abraham Lincoln had been president.

For the purposes of assessing causality, counterfactual analysis has the same general task as experimental, statistical, and comparative methods: to organize evidence to show that a change in the value of an outcome variable can be traced to the effects of a change in a causal variable, and not to changes in other variables. Just as experimental research manipulates one variable at a time in a controlled setting, and comparative research tries to select matched cases in which covariation with the outcome variable is limited to a single causal variable, counterfactual analysis ideally posits an alternative world that is identical to the real world in all theoretically relevant respects but one, in order to explore the consequences of that difference.

That is easier said than done, and the same problems that plague comparative research constrain counterfactual analysis as well: In a system of interconnected behavior, a change in one variable reverberates through the system and induces changes in many other variables, so that "we can never do merely one thing" (Hardin, cited in Jervis 1997, 10). In counterfactual thought experiments, as in a matching cases strategy, it is often quite difficult to hold everything else equal. Hence Lebow (2000) questions the utility of "surgical" counterfactuals.

An attempt to assess the impact of US nuclear superiority on the outcome of the Cuban Missile Crisis by imagining the crisis under conditions of Soviet strategic superiority would have to change too much history to be useful. Soviet superiority would be conceivable only if the US economy and technological capacity had been much weaker, if American society did not support a competitive military establishment, and so on. The presence of these other conditions would almost certainly

¹² Admittedly, an empirical confirmation of any detailed counterfactual prediction would be particularly compelling precisely because the *ex ante* probability was so low (Popper 1965).

have changed the world in other ways as well (the status of Berlin, for example), further complicating any effort to say that it was Soviet superiority, rather than these other changes, that caused the outcome. Moreover, a Soviet Union with strategic superiority would have had no incentive to put offensive missiles in Cuba in the first place.

As this example suggests, a counterfactual conditional is not complete in itself. It relies on other changes to sustain it. Goodman (1983) called these *connecting principles*, which themselves involve counterfactual propositions about the consequences of those other changes. Any counterfactual analysis needs to specify the secondary counterfactuals, or "enabling counterfactuals" (Lebow 2000), that must be introduced to sustain the primary counterfactual.

A good counterfactual requires that these connecting principles and enabling counterfactuals be specified with reasonable precision (like the primary counterfactual) and that they be consistent with each other and with the antecedent of the primary counterfactual.¹³ Goodman (1983, 15) refers to this requirement of logical consistency as *cotenability*.

The counterfactual of Soviet strategic superiority in 1962 fails to satisfy the cotenability criterion. So does the argument that if Nixon had been president during the Cuban Missile Crisis, he would have ordered an air strike rather than a naval blockade. Lebow and Stein (1996) argue persuasively that if Nixon had been president he, unlike Kennedy, would have probably authorized the use of US forces in the Bay of Pigs operation, Castro would have been overthrown, and the Soviets would not have put offensive missiles in Cuba.

In his attempt to assess the likely development of the American economy in the absence of railroads, Fogel (1964) identified other developments that were likely to occur in the absence of railroads, including the introduction of the internal combustion engine and the automobile (along with its demands for iron and other materials) fifty years before its actual appearance. Elster (1978, 204–8) basically argues that this assumption of the early emergence of the automobile is not cotenable with the assumption of the delay of the railroad, since the technology upon which the automobile was based surely would have led to the railroad.¹⁴

For these reasons, most analysts accept Max Weber's (1949) argument that for the purposes of causal analysis the best counterfactual worlds to examine are those that require as few changes as possible in the real world. This is the "minimal rewrite of history" rule (Tetlock and Belkin 1996), which is defined in terms of the magnitude of the changes as well as their number.¹⁵

¹³ Tetlock and Belkin (1996, 21) include consistency with the consequent, which I discuss in the next section on the conditional probability of the consequent given the antecedent.

¹⁴ See also Tetlock and Belkin (1996, 22) and Lebow (2000, 582).

¹⁵ Similarly, King and Zeng (2005) use statistical modeling to demonstrate that whereas the validity of counterfactual propositions positing relatively modest changes in the real world can be evaluated based on the data, "extreme counterfactuals" are quite sensitive to assumptions built into the model that have little to do with the data.

One example of a minimal rewrite counterfactual is the proposition that if George W. Bush had not won the 2000 election, the United States would not have gone to war in Iraq. This counterfactual involves a minimal rewrite of history because one does not have to change much—just a modest number of Florida voters reading their ballots correctly. So the counterfactual's antecedent is quite plausible.¹⁶ The hypothesized consequent (President Al Gore not invading Iraq) is quite plausible but not certain, and the argument would have to include enabling counterfactuals such as how Gore would have reacted to 9/11.

Lebow's (2007) argument that if the assassination attempt against the Archduke had failed the First World War probably would not have occurred is another highly plausible, minimal rewrite counterfactual. The antecedent is quite easy to imagine, as Lebow explains in great detail. Indeed the *ex ante* probability of the alternative world was undoubtedly higher than that of the real world, which is another possible criterion for the evaluation of counterfactuals.¹⁷ Lebow's analysis is particularly useful as an illustration of the value of specifying the *conditions and processes* through which one could get from the real world to the counterfactual world. Another good example of this is Mueller's (1991) detailed assessment of the identity and policy preferences of other possible German leaders who might have held the position of chancellor, to support his argument that in the absence of Hitler the Second World War would not have occurred.

Scholars will disagree, of course, on how minimal a rewrite of history has to be for the antecedent to be considered plausible. There is no single answer to this question, as there may be a trade-off between maximizing the plausibility of a counterfactual by minimizing the number of additional conditions necessary to sustain it, and selecting historically or theoretically meaningful counterfactuals. As the historian Schroeder (2007, 149–50) argues, “a major counterfactual . . . will change too much, and a minor one too little, to help us explain what really did happen and why, and why alternative scenarios failed to emerge.”¹⁸ Still, counterfactuals relating to the 1914 assassination and the 2000 American election demonstrate that minimal rewrite counterfactuals can be consequential.

Ferguson (1999a, 86) offers his own answer of how we distinguish plausible and implausible counterfactuals: “We should consider as plausible or probable only those alternatives which we can show on the basis of contemporary evidence that contemporaries actually considered.” Ferguson further narrows the range of acceptable counterfactuals by restricting them to “hypothetical scenarios which contemporaries not only considered, but also committed to paper.”

The second criterion is unnecessarily restrictive. While paper records of actors' alternative options might provide particularly compelling evidence of choices not made, we should not exclude evidence based on oral interviews with first-hand

¹⁶ By contrast, counterfactual claims about the consequences of George W. Bush not winning the 2004 election posit a more problematic antecedent.

¹⁷ The validity of the hypothesized consequent is another, analytically distinct question which I return to later.

¹⁸ See also Weber (1996) and Lebow (2000).

observers. Nor should we be precluded from exploring alternative histories in repressive political systems in which actors are afraid to commit their thoughts to paper or among nonelite groups for which there are no written records.

Even Ferguson's (1999a) first criterion might be too restrictive. It would exclude alternative choices that actors failed to consider, or failed to consider seriously, because of psychological biases, political constraints, or failure of imagination—alternative possibilities that other decision-makers, had they been in office, might have considered (assuming a change in personnel was a plausible antecedent). It is not clear, for example, whether Ferguson's criterion would allow us to consider the consequences of the US responding to the Soviet missiles in Cuba by doing nothing, since they gave that option little or no serious consideration. In addition, some options are formally considered only for political reasons but given no serious thought; as Janis (1982) suggests for many "devil's advocate" arguments.

Schroeder (2007, 151–2) agrees that the counterfactuals actually considered by the actors themselves are important, but recognizes that they are too restrictive. He argues that the historian needs to pose counterfactual questions of his own: "What other decisions and actions could the historical actors have made under the existing circumstances? To what extent did they recognize and consider these? What circumstances made these choices or alternative courses genuinely possible or merely specious and actually unreal? What might the alternative results of these choices have been?" These are precisely the right questions to ask, but even they might be too restrictive. They fail to give enough emphasis to counterfactual possibilities introduced by external events such as assassinations, battles or elections won or lost, fatal illness, personal tragedies, and other highly contingent events (Lebow 2000, 560).¹⁹

Game theory provides a still more restrictive set of criteria for the selection of acceptable counterfactuals from the enormous number of possible counterfactuals. First of all, game-theoretic models in extensive form specify exactly what will happen if actors were to make different decisions at various choice nodes. All sequences of choices "off the equilibrium path" are possible counterfactuals. But actors did not go down those paths for a reason. In a large class of games, however, there are multiple equilibria, multiple ways all actors could have behaved that were fully consistent with their interests given the constraints of the game. Each of these paths constitutes a fully acceptable counterfactual (Bueno de Mesquita 1996; Weingast 1996).²⁰

The consideration of the alternatives that actors did consider or might have considered within a choice-theoretic framework raises an interesting asymmetry in counterfactual analysis. Both events and nonevents generate counterfactuals, but it is often easier to examine event-generated counterfactuals than those generated

¹⁹ The outcomes of a small percentage of battles hinge on accidents, luck, insubordination, unexpected weather, and other contingencies, and the long-term political consequences of such reversals of outcome can be profound (Cowley 1999).

²⁰ The existence of alternative subgame perfect equilibria provides an excellent criterion for good counterfactuals where the specification of the game provides a reasonable fit to the situations faced by actors and choices they consider. We need to remember, however, that the simplifications necessary to create tractable games, including limits on the number of choices available to actors, may be too restrictive in some situations, including those involving enormous social complexity.

by nonevents. We can explore the consequences of a counterfactual failure of the assassination plot in 1914, since there is ample evidence from pre-assassination records of how Austrian and German leaders defined their decision problems and options, and thus how they might have acted. In the absence of an assassination, however, it would be far more difficult to explore the counterfactual possibility of an assassination of the Archduke, since presumably no one at the time gave serious thought to that possibility and what they would do if it happened, and we would have to introduce more secondary counterfactuals that would be somewhat speculative. Similarly, it is easier (though not easy) to explore counterfactuals created by the hypothetical failure of the assassination attempt against Kennedy than those that would have been created by the hypothetical success of the assassination attempt against Reagan. The *ex ante* probability, and hence counterfactual plausibility, of each of these antecedents is quite different.

If the 9/11 plot had failed and gone undetected, and if the US had not invaded Iraq (which is quite plausible but not certain), it would be quite difficult to undertake a counterfactual analysis, set in the absence of 9/11, of the possibility that a major terror attack against the United States originating from Afghanistan might trigger an American war against Iraq in 2003. In a world in which 9/11 (or something like it) had not occurred in 2001 (or soon thereafter), would any of our criteria for evaluating counterfactuals permit the characterization of a 9/11 attack as a plausible antecedent and an American invasion of Iraq as a likely consequent?²¹

2.3 The Conditional Plausibility of the Consequent

Thus far we have emphasized that the characteristics of a useful counterfactual include a well-specified antecedent and consequent and an antecedent or conditional that is plausible, involves a minimal rewrite of history, and is sustainable through conditions that are cotenable with each other and with the antecedent. The next question is whether the antecedent, along with the conditions that are necessary to support it, is likely to lead to the hypothesized consequent. The basic requirement, like those for any theoretical proposition, is that the causal linkages be clearly specified, logically complete, and consistent with the empirical evidence. It is admittedly difficult to separate theoretical and empirical criteria, since social science research is ideally characterized by an ongoing dialogue between theory and evidence (Levy 2007), but listing them separately is useful for our purposes here.

In many ways, the most important requirement for a good counterfactual, besides a plausible antecedent, is a good theory. As Fogel (1964, cited in Tetlock and Belkin 1996, 26) writes, "Counterfactual propositions... are merely inferences from hypothetico-deductive models." The greater the extent to which the hypothesized

²¹ For an intriguing counterfactual analysis set in a counterfactual world, but one that involves a much more extended timeframe than the one posited here, see Lebow (2006). Imagining a world in which Mozart lived to sixty-five and in which neither of the world wars of the twentieth century occurred, Lebow considers the possibility that an early death of Mozart triggered a chain of historical happenings that led to these events, and considers the critiques and defenses of such a counterfactual.

causal mechanisms leading from a specific antecedent to a particular consequent are consistent with well-established and empirical confirmed theories, the greater the plausibility of the counterfactual. The better the theory—defined in terms of its logical coherence, precision, deductive support from well-established theories, and empirical support—and the more it makes explicit predictions about what will happen under a variety of counterfactual conditions, the better the counterfactual.

For the purposes of evaluating a particular counterfactual proposition, we are more interested in the predictive power of a relevant theory under the specific conditions defined by the antecedent than in its predictive power under a wide range of conditions, although the latter increases our confidence in the former. For this reason, a well-established empirical law, even in the absence of consensus regarding how to explain it theoretically, would also provide useful support for a counterfactual proposition. Admittedly, we have relatively few empirical laws in political science, but some propositions have far more empirical support than others.

Thus *consistency with the empirical evidence* is another criterion for a good counterfactual proposition. Tetlock and Belkin (1996, 27–30) include “consistency with well-established statistical generalizations” as one of their six criteria of a good counterfactual. I would construe evidence more broadly. While statistical (and experimental) evidence is highly desirable—assuming a sufficient number of comparable cases, variables that are operationalizable and measurable over those cases, accessible data, etc.—such evidence is not always available. Comparative studies and, less frequently, single case studies (if based on a compelling most or least likely design) can sometimes provide an adequate evidentiary basis to support a counterfactual proposition, and such evidence, when combined with statistical evidence, can often provide substantially better support than statistical evidence alone (George and Bennett 2005).

Although “direct” support for a specific counterfactual proposition is important, our confidence in its validity is enhanced if the counterfactual generates other theoretical predictions that are also supported. A counterfactual proposition has implications not only for the final outcome, but also for the intervening paths between antecedent and consequent, and those intervening predictions should be specified in advance and tested if at all possible. This is the criterion of *projectibility* (Tetlock and Belkin 1996, 30–1).

Let me return to the point I made above about the likelihood of a counterfactual being a function of the length of the causal chain leading to it. This is just the commonsense notion that short-term predictions are more plausible than long-term predictions, which applies to counterfactual propositions as well as to those with factual antecedents. Even where one finds short-term regularities, the compounding of small uncertainties generates enormous irregularity and unpredictability over the long term. Even deterministic processes can generate highly unpredictable outcomes if they are sensitive to initial conditions, as chaos theory demonstrates.

Fearon (1996) demonstrates this in a compelling way in his discussion of an extremely simple process involving cellular automata that follow rules of behavior that are well understood, precisely specified, but stochastic, with well-defined probabilities of moving from one state to another. He demonstrates how such processes generate local regularities but global unpredictability. Fearon’s (1996) analysis leads

him to suggest a *proximity* criterion, so that we can assess plausibility “only where the counterfactuals involve causal mechanisms and regularities that are well understood and that are considered at a spatial and temporal range small enough that multiple mechanisms do not interact, yielding chaos” (p. 66).

A related reason for insisting on a proximity criterion is that even if the antecedent is plausible and there are good theoretical and empirical reasons to believe that the presence of the antecedent would lead to predicted consequents with fairly high probability, it is always possible that subsequent developments might return history to its original course, before it was diverted by the hypothesized antecedent. Lebow (2000, 584) labels these “second-order counterfactuals,” but I prefer the term *redirecting* counterfactuals, since primary counterfactuals generate numerous secondary counterfactuals and only some of them reroute history back to its original course.

Consider the argument that the Vietcong attack on the US military base at Pleiku in 1965 was a significant cause of US escalation in the Vietnam War. In the absence of that Vietcong attack, however, it is quite plausible to argue that another incident would have occurred, whether randomly or deliberately instigated for strategic advantage, and led to an American escalation. As Bundy argued, “Pleikus are like streetcars.” Wait long enough and one will come along.²² This remark is often used in support of structural arguments and against arguments about the importance of contingency in historical processes.

Another example of a redirecting counterfactual arises with respect to Lebow’s (2007) own counterfactual argument that in the absence of the assassination of the Archduke the First World War probably would not have occurred. Lebow argues that in the absence of the assassination, existing trends in the balance of military power that favored Russia would have forced Moltke (or his successor as Chief of the German General Staff) to abandon the Schlieffen Plan and adopt instead a more defensive strategy, which would have eliminated the incentives for preventive or preemptive military action in any crisis that might arise.

This counterfactual is clear and the antecedent constitutes a minimal rewrite. The causal linkages to the consequent are not plausible, however, because they would have probably generated self-refuting strategic behavior. Lebow’s assessment of military trends is right on the mark. German military and political leaders would almost certainly have accepted his vision of a 1917 world, and precisely for that reason they never would have allowed that world to come about. The same fears for the future that generated Germany’s preventive motivation for war in 1914 (which Lebow implicitly acknowledges) would have led German leaders to initiate or provoke the preventive war they thought they needed before Russia grew too strong (Fischer 1967). This critique, of course, generates its own counterfactual, which would need to be evaluated against Lebow’s counterfactual. Fortunately, the expansive literature on the First World War provides enough evidence to resolve this debate, even if not with complete certainty.

²² National Archives and Records Administration, Lyndon Baines Johnson Library, oral interviews of Frederick W. Flott, <http://webstorage1.mcpa.virginia.edu/library/mc/poh/transcripts/flott_frederick_1984_0927.pdf>.

3 CONCLUSION

All causal statements generate counterfactuals about what would happen if certain variables were to take on different values, and all nonexperimental methodologies must deal with this in one way or the other. I focus here on the role of counterfactuals in case studies. I argue that if counterfactuals are made explicit and used according to scientifically acceptable rules of inference, a study of history as it really wasn't can help us understand history as it really was (to borrow from Ferguson 1999a).

Counterfactuals serve different theoretical and descriptive purposes. Different theoretical goals and normative values call for different trade-offs among various research objectives, and consequently there is no single set of methodological criteria applicable to all counterfactuals. Leaving aside the use of counterfactual thought experiments to stimulate the imagination, which is useful but which follows a different set of rules, I suggest criteria for the evaluation of counterfactuals for the purposes of explaining historical cases or using cases to assess more general theoretical propositions.

The analyst should clearly specify a counterfactual's antecedent, consequent, and the causal linkages between them. Counterfactuals should change as few aspects of the real world as possible in order to isolate their causal effects (the "minimal-rewrite-of-history" rule). The analyst should specify both the "supporting conditions" or "connecting principles" that are needed to sustain the primary counterfactual and the secondary counterfactuals that lead from the antecedent (and its supporting conditions) to the consequent. The consequent should be more specific than "the outcome would have been different," but not so specific that it becomes implausible, given the fact that the probability of a highly specific outcome is far less than the probability of a range of outcomes.

Counterfactual analysis, perhaps even more so than other kinds of analysis, is a theory-driven process. We cannot directly trace the consequences of an unobservable antecedent, so we must rely on theoretical knowledge. The stronger the theory, and the more the analyst can resort to theory or empirical laws to justify his or her hypothesized causal mechanisms linking the antecedent to the consequent, the better the counterfactual. Counterfactuals that generate additional observable implications and that can themselves be "tested" against the evidence provide additional confidence in the validity of the counterfactual. This is the "projectibility" criterion.

Logical consistency is another important criterion. The supporting conditions for a primary factual must be consistent with each other and with the antecedent. Their linkages to the consequent must also be theoretically sensible. It does little good to posit an antecedent that could only lead to the consequent if supported by connecting principles and secondary counterfactuals that were themselves inconsistent with the antecedent. Researchers should be particularly alert to strategic behavior that might return history to its original path ("redirecting counterfactuals")—where actors recognize that the hypothesized consequences of the antecedent are both accurate and undesirable, and act to head off those outcomes.

It would also be inconsistent to try to support an argument that a particular outcome was contingent by demonstrating that a small change in the situation would have invariably led to a different outcome—one cannot support an argument for contingency by invoking a deterministic process. To the extent possible, based on existing theory and empirical evidence, the author should give a sense of the likelihood that the consequent would have followed from the antecedent.

Counterfactual analysis, like any causal analysis, must recognize that uncertainty about initial conditions and the pervasiveness of stochastic behavior limits even the best of our theories to relatively short-term predictions. After that, too many things interact in too many unpredictable ways. The longer the causal chain, the lower the probability of occurrence of the outcome at the end of the chain, particularly given the relative absence of sufficient conditions in social theory. This suggests temporal and spatial “proximity” as an additional criterion for evaluating counterfactuals.

While the application of these criteria might make “after dinner history” somewhat less entertaining, it disciplines the use of counterfactuals and provides an additional methodological tool for evaluating causality in a nonexperimental world in which many confounding variables interact in unpredictable ways.

REFERENCES

- BATES, R., GREIF, A., LEVI, M., ROSENTHAL, J.-L., and WEINGAST, B. 1998. *Analytic Narratives*. Princeton, NJ: Princeton University Press.
- BUENO DE MESQUITA, B. 1996. Counterfactuals and international affairs: some insights from game theory. Pp. 211–29 in *Counterfactual Thought Experiments in World Politics*, ed. P. E. Tetlock and A. Belkin. Princeton, NJ: Princeton University Press.
- CARR, E. H. 1964. *What is History?* Harmondsworth: Penguin.
- CEDERMAN, L.-E. 1996. Rerunning history: counterfactual simulation in world politics. Pp. 247–67 in *Counterfactual Thought Experiments in World Politics*, ed. P. E. Tetlock and A. Belkin. Princeton, NJ: Princeton University Press.
- COWLEY, R. 1999. *What If?* New York: G. P. Putnam's Sons.
- ELSTER, J. 1978. *Logic and Society*. New York: John Wiley and Sons.
- ENGLISH, R. 2007. Perestroika without politics: how realism misunderstands the Cold War's end. Pp. 237–60 in *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals*, ed. G. Goertz and J. S. Levy. New York: Routledge.
- FEARON, J. D. 1996. Causes and counterfactuals in social science: exploring an analogy between cellular automata and historical processes. Pp. 39–67 in *Counterfactual Thought Experiments in World Politics*, ed. P. E. Tetlock and A. Belkin. Princeton, NJ: Princeton University Press.
- FERGUSON, N. 1999a. Virtual history: toward a “chaotic” theory of the past. Pp. 1–90 in *Virtual History: Alternatives and Counterfactuals*, ed. N. Ferguson. New York: Basic Books.
- 1999b. *The Pity of War: Explaining World War I*. New York: Basic Books.
- FISCHER, D. 1971. *Historians' Fallacies: Toward a Logic of Historical Thought*. London: Routledge and Kegan Paul.
- FISCHER, F. 1967. *Germany's Aims in the First World War*. New York: Norton.

- FOGEL, R. 1964. *Railroads and American Economic Growth: Essays in Econometric History*. Baltimore: Johns Hopkins University Press.
- GEORGE, A. L., and BENNETT, A. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, Mass.: MIT Press.
- GOERTZ, G., and LEVY, J. S. 2007. Causal explanation, necessary conditions, and case studies. Pp. 9–45 in *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals*, ed. G. Goertz and J. S. Levy. New York: Routledge.
- and STARR, H. (eds.) 2003. *Necessary Conditions: Theory, Methodology, and Applications*. Lanham, Md.: Rowman and Littlefield.
- GOODMAN, N. 1983. *Fact, Fiction, and Forecast*, 4th edn. Cambridge, Mass.: Harvard University Press.
- HEMPEL, C. G. 1966. *Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice Hall.
- JANIS, I. L. 1982. *Groupthink*, 2nd rev. edn. Boston: Houghton Mifflin.
- JERVIS, R. 1997. *System Effects*. Princeton, NJ: Princeton University Press.
- KHONG, Y. F. 1996. Confronting Hitler and its consequences. Pp. 95–118 in *Counterfactual Thought Experiments in World Politics*, ed. P. E. Tetlock and A. Belkin. Princeton, NJ: Princeton University Press.
- KING, G., KEOHANE, R., and VERBA, S. 1994. *Designing Social Inquiry*. Princeton, NJ: Princeton University Press.
- and ZENG, L. 2006. The dangers of extreme counterfactuals. *Political Analysis*, 14: 131–59.
- KINGDON, J. 1984. *Agendas, Alternatives, and Public Policies*. Boston: Little, Brown.
- LEBOW, R. N. 2000. What's so different about a counterfactual? *World Politics*, 52: 550–85.
- 2006. If Mozart had died at your age: psychologic versus statistical inference. *Political Psychology*, 27: 157–72.
- 2007. Contingency, catalysts, and nonlinear change: the origins of World War I. Pp. 85–111 in *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals*, ed. G. Goertz and J. S. Levy. New York: Routledge.
- and STEIN, J. G. 1996. Back to the past: counterfactuals and the Cuban missile crisis. Pp. 119–48 in *Counterfactual Thought Experiments in World Politics*, ed. P. E. Tetlock and A. Belkin. Princeton, NJ: Princeton University Press.
- LEVY, J. S. 2001. Explaining events and testing theories: history, political science, and the analysis of international relations. Pp. 39–83 in *Bridges and Boundaries: Historians, Political Scientists, and the Study of International Relations*, ed. C. Elman and M. F. Elman. Cambridge, Mass.: MIT Press.
- 2007. Theory, evidence, and politics in the evolution of research programs. Pp. 177–97 in *Theory and Evidence in Comparative Politics and International Relations*, ed. R. N. Lebow and M. Lichbach. New York: Palgrave Macmillan.
- 2008. Case studies: types, designs, and logics of inference. *Conflict Management and Peace Science*, 25: 1–18.
- LEWIS, D. 1973. *Counterfactuals*. Cambridge, Mass.: Harvard University Press.
- MAHONEY, J. 2000. Path dependence in historical sociology. *Theory and Society*, 29: 507–48.
- MORGAN, S. L., and WINSHIP, C. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. New York: Cambridge University Press.
- MUELLER, J. 1991. Changing attitudes towards war: the impact of the First World War. *British Journal of Political Science*, 21: 1–28.
- OAKESHOTT, M. 1966. *Experience and its Modes*. London: Cambridge University Press.
- POPPER, K. 1965. *The Logic of Scientific Discovery*. New York: Harper Torchbacks.
- RIPSMAN, N. M., and LEVY, J. S. 2007. The preventive war that never happened: Britain, France, and the rise of Germany in the 1930s. *Security Studies*, 16: 32–67.

- ROESE, N. J., and OLSON, J. M. 1995. *What Might Have Been: The Social Psychology of Counterfactual Thinking*. Mahwah, NJ: Lawrence Erlbaum.
- SCHROEDER, P. W. 2007. Necessary conditions and World War I as an unavoidable war. Pp. 147–93 in *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals*, ed. G. Goertz and J. S. Levy. New York: Routledge.
- SKOCPOL, T. 1979. *States and Social Revolutions*. Cambridge: Cambridge University Press.
- TETLOCK, P. E., and BELKIN, A. 1996. Counterfactual thought experiments in world politics: logical, methodological, and psychological perspectives. Pp. 1–38 in *Counterfactual Thought Experiments in World Politics*, ed. P. E. Tetlock and A. Belkin. Princeton, NJ: Princeton University Press.
- and LEBOW, R. N. 2001. Poking counterfactual holes in covering laws: cognitive styles and historical reasoning. *American Political Science Review*, 95: 829–43.
- and PARKER, G. (eds.) 2006. *Unmaking the West: "What If?" Scenarios that Rewrite World History*. Ann Arbor: University of Michigan Press.
- WEBER, M. 1949. *The Methodology of the Social Sciences*. Glencoe, Ill.: Free Press.
- WEBER, S. 1996. Counterfactuals, past and future. Pp. 268–88 in *Counterfactual Thought Experiments in World Politics*, ed. P. E. Tetlock and A. Belkin. Princeton, NJ: Princeton University Press.
- WEINGAST, B. R. 1996. Off-the-path behavior: a game-theoretic approach to counterfactuals and its implications for political and historical analysis. Pp. 230–43 in *Counterfactual Thought Experiments in World Politics*, ed. P. E. Tetlock and A. Belkin. Princeton, NJ: Princeton University Press.