Counterfactuals, Causal Inference, and Historical Analysis

JACK S. LEVY

I focus primarily on the utility of counterfactual analysis for help-ing to validate causal inferences in historical analysis. How can we use what did not happen but which easily could have hap-pened to understand what did happen? With an infinite number of things that might have happened, and with temptations to con-struct "counterfactuals of convenience" to bolster one's preferred historical interpretations or political preferences, we need a set of rules or best practices for evaluating the scientific legitimacy of counterfactuals. Building on earlier work in several disciplines, I develop a set of criteria for the conduct of counterfactual analysis in historical case studies. The best counterfactuals begin with clearly specified plausible worlds involving small and easily imaginable changes from the real world. They make relatively short-term pre-dictions based on empirically validated theoretical generalizations and on secondary counterfactuals that are mutually consistent. These counterfactuals are also sensitive to strategic behavior that might return history to its original course, and they are explicitly tested against competing counterfactuals.

It is often said that the US Civil War would not have occurred in the ab-sence of slavery, that the First World War would not have happened without the assassination of Austrian archduke Franz Ferdinand, and that the United States would not have initiated the 2003 Iraq War if Al Gore rather than George W. Bush had been the US president. Political leaders defend their policies by emphasizing the disastrous consequences that would have fol-lowed from alternative choices. They also invoke counterfactuals to distance themselves from failed policies. Four years after Hillary Clinton voted in the

Jack S. Levy is Board of Governors' Professor of Political Science at Rutgers University and an Affiliate of the Saltzman Institute of War and Peace Studies at Columbia University.

Senate to authorize the president to use force in Iraq, she stated that "if we knew then what we know now... I certainly wouldn't have voted that way."[1] Counterfactuals are part of everyday life. Most of us occasionally think about the girl or boy who got away, about "the road not taken," or about the one play that lost our favorite team the championship.

Each of these statements involves a counterfactual—a "possible," "parallel," or "alternate" world in which key features of the real world were not present or took on different values. A counterfactual involves "the mental construction of a course of events which is altered through modification in one or more 'conditions.'"[2] More technically, it is a "subjunctive conditional in which the antecedent is known or supposed for purposes of argument to be false."[3]

Historians have invoked counterfactuals ever since Herodotus argued that if the Athenians had not defeated the much stronger Persian forces in the Battle of Salamis (480 BCE), "Hellas would have been conquered by the Persians."[4] Four centuries later, Livy argued that if Alexander the Great had invaded Rome he would have been defeated.[5] Historians have been quite divided, however, as to the analytic utility of counterfactuals. Michael Oakeshott argued that counterfactual history "is not merely bad or doubtful history, but the complete rejection of history... a monstrous incursion of science into the world of history."[6] Criticisms of counterfactual history (or "what if" history, or "allohistory") are nicely summarized by E. H. Carr's argument that although counterfactuals might be entertaining "parlour games," they are too arbitrary, speculative, and self-serving to contribute to historical understanding. Counterfactuals are too arbitrary in the selection of some key pivotal moments rather than others as starting points in the alternative world, too speculative in the "imaginative history" about what follows from those pivotal moments, and too self-serving in the construction of counterfactuals to support historians' own preferred interpretations or ideological agendas.[7] This tendency to adopt "counterfactuals of convenience"

---

[1] Interview with Hillary Clinton, *Today*, NBC, 18 December 2006.

[2] Max Weber, *The Methodology of the Social Sciences*, trans. Edward A. Shils and Henry A. Finch (New York: Free Press, [1905] 1949), 173.

[3] Philip E. Tetlock and Aaron Belkin, "Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives," in *Counterfactual Thought Experiments in World Politics*, ed. Tetlock and Belkin (Princeton, N.J.: Princeton University Press, 1996), 4.

[4] Herodotus, *The Histories*, in *The Landmark Herodotus*, ed. by Robert B. Strassler, trans. Andrea L. Pervis (New York: Pantheon Books, 2007), 7.139.

[5] Livy, *History of Rome from Its Foundation*, Book IX, trans. Betty Radice (Harmondsworth, UK: Penguin, 1982), 9.17–19.

[6] Michael Oakeshott, *Experience and Its Modes* (Cambridge: Cambridge University Press, 1933), 129.

[7] E. H. Carr, *What is History?* (Harmondsworth, UK: Penguin Books, 1964), 97. See also Philip E. Tetlock and Geoffrey Parker, "Counterfactual Thought Experiments: Why We Can't Live without Them &

is hardly new. Some argue that Livy used the counterfactual of Alexander's failure to conquer Rome as a veiled critique of the great man theory of history, of the anti-Republican tradition in Rome, and possibly of Augustus himself.[8]

Other historians, while often acknowledging potential problems inherent in the use of counterfactuals, emphasize the potential utility, or at least inescapability, of counterfactual history. In response to arguments like Walter Rathenau's that history should focus on "what is and what was, not what would be and what would have been," others respond that to explain what is or what was, one must identify causes.[9] Statements about causes carry implications about what could have been if actors had made different choices or if conditions had been slightly different. Knowledge about what might have been can help explain what actually was. Historical actors themselves think in terms of the alternative futures associated with the different choices they might make, and to understand the choices actors made the historian must consider those alternative futures, which are now alternative pasts. Thus Niall Ferguson 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."[10] Similarly, Paul Schroeder argues that the value of counterfactual reasoning for historians is "to shed light on what actually did happen, why it did, and what it means."[11] This view is shared by most social scientists. Max Weber writes that "in order to penetrate to the real causal interrelationships, *we construct unreal* ones." Bruce Bueno de Mesquita claims that "we cannot understand what happened in reality without understanding what did not happen but might have happened under other circumstances."[12]

How We Must Learn to Live with Them," in *Unmaking the West: "What If?" Scenarios That Rewrite World History*, ed. Tetlock, Richard Ned Lebow, and Parker (Ann Arbor: University of Michigan Press, 2006), 28–33. Richard J. Evans develops the self-serving theme and argues that "wishful thinking is everywhere in the world of historical counterfactual," in Evans, *Altered Pasts: Counterfactuals in History* (Waltham, MA: Brandeis University Press, 2013), chap. 1. Gavriel Rosenfeld argues that "[a]lternative history is inherently presentist," designed to advance a policy agenda. Rosenfeld, *The World Hitler Never Made: Alternate History and the Memory of Nazism* (New York: Cambridge University Press, 2005), 10.

[8] Ruth Morello, "Livy's Alexander Digression (9.17–19): Counterfactuals and Apologetics," *Journal of Roman Studies* 92 (2002): 62–85.

[9] Quoted in Evans, *Altered Pasts*, 125. As Carr himself noted, "the study of history is a study of causes." Carr, *What Is History?* 87.

[10] Niall Ferguson, "Virtual History: Toward a 'Chaotic' Theory of the Past," in *Virtual History: Alternatives and Counterfactuals*, ed. Ferguson (New York: Basic Books, 1999), 87.

[11] Paul W. Schroeder, "Embedded Counterfactuals and World War I as an Unavoidable War," in Schroeder, *Systems, Stability, and Statecraft: Essays on the International History of Modern Europe*, ed. David Wetzel, Robert Jervis, and Jack S. Levy (New York: Palgrave, 2004), 157.

[12] Weber, *Methodology of the Social Sciences*, 185–86; Bruce Bueno de Mesquita, "Counterfactuals and International Affairs: Some Insights from Game Theory," in *Counterfactual Thought Experiments*, ed. Tetlock and Belkin, 229.

At the same time, scholars recognize that counterfactuals raise difficult methodological questions. Causal statements about the real world are, at least in principle, more amenable to empirical tests than those about the counterfactual world because the key causal and contextual variables of the former are observable or have observable implications.[13] A counterfactual, however, rests on a nonexistent conditional whose consequences cannot be empirically traced. As a result, a counterfactual proposition "can never be subjected to any direct empirical test."[14] There are an endless number of paths that history might have taken from any given starting point. In the absence of observable data, how do we select among them? How can we say that some counterfactuals are more scientifically legitimate, valid, or useful than others?[15]

There are many different uses of counterfactuals, and scholars in numerous disciplines have taken an interest in counterfactuals.[16] In this article I focus primarily on the utility of counterfactual analysis for helping to validate causal inferences in case study and historical research. How can counterfactuals be used to support arguments about the causal impact of particular variables in a particular historical episode? How can we use what did not happen but that might have happened to help understand what did happen? This engages the closely related question of whether history might have turned out differently if key conditions or leaders had been different or if particular events had not occurred. To put this discussion in context, however, I begin by reviewing the different types of counterfactuals and uses to which they have been put. I then develop criteria for evaluating the scientific utility or legitimacy of counterfactuals in case study analysis.

---

[13] As Hume reminds us, however, there are limits to the empirical validation of a hypothesized causal relationship.

[14] Nelson Goodman, *Fact, Fiction, & Forecast*, 4th ed. (Cambridge, MA: Harvard University Press, 1983), 4.

[15] A similar problem plagues quantitative analysis. Although one can observe the effects of a treatment on one individual/unit and the effects of a nontreatment on another individual/unit, it is impossible to observe the effects of both a treatment and nontreatment on a single individual/unit under identical conditions. The growing recognition of this central problem of causal inference among large-N researchers has led to attempts to develop statistical procedures to deal with it, such as the "potential outcomes framework." Stephen L. Morgan and Christopher Winship, *Counterfactuals and Causal Inference Methods and Principles for Social Research* (New York: Cambridge University Press, 2007).

[16] In philosophy, see David Lewis, *Counterfactuals* (Cambridge, MA: Harvard University Press, 1973); John Collins, Ned Hall, and L. A. Paul, eds., *Causation and Counterfactuals* (Cambridge, MA: MIT Press, 2004). In economics, see Robert Fogel, *Railroads and American Economic Growth: Essays in Econometric History* (Baltimore, MD: Johns Hopkins University Press, 1964); Peter D. McClelland, *Causal Explanation and Model Building in History, Economics, and the New Economic History* (Ithaca, NY: Cornell University Press, 1975). In literature, film, and popular culture, see Ferguson, "Virtual History," 1–90; Rosenfeld, *The World Hitler Never Made*, chap. 1; Evans, *Altered Pasts*.

TYPES AND USES OF COUNTERFACTUALS

Scholars differentiate among different types and uses of counterfactuals.[17] One distinction is between the methodological, epistemological, or evaluative use of counterfactuals and the descriptive analysis of counterfactuals.[18] The first, which is the primary focus of this article, addresses the methodological question of how we should use counterfactuals in social science for the purposes of exploring causality or of analyzing whether history had to turn out as it did. What are the criteria by which we should evaluate the utility of counterfactual arguments in social science or history? The descriptive analysis of counterfactuals, common among social psychologists and cognitive scientists, focuses on how people actually use counterfactuals in thinking about the world. Do people think about factual and counterfactual questions differently? How do their cognitive and motivational biases affect counterfactual judgments and inferences?[19]

Another important distinction is between "plausible world" counterfactuals and "miracle counterfactuals."[20] Plausible world counterfactuals, or "minimal rewrite" counterfactuals, involve antecedents that one could easily imagine having occurred (a failed assassination attempt, for example).[21] They include historical "near misses" or missed opportunities. We could ask, for example, how close the Cuban missile crisis came to escalating to war, including nuclear war.[22] Miracle counterfactuals impose no constraints on the values of key variables in the alternative world and no requirement for an explanation of how the alternative world might have emerged from the real world.

---

[17] For more discussion see Tetlock and Belkin, "Counterfactual Thought Experiments," 3–16. I leave aside the "entertainment" value of counterfactual history.

[18] Ibid., 4–5.

[19] Neal J. Roese and James M. Olson, *What Might Have Been: The Social Psychology of Counterfactual Thinking* (Mahwah, NJ: Lawrence Erlbaum, 1995); Philip E. Tetlock, *Expert Political Judgment* (Princeton, NJ: Princeton University Press, 2005); Richard Ned Lebow, *Forbidden Fruit: Counterfactuals and International Relations* (Princeton, NJ: Princetown University Press, 2010), chaps. 5–6.

[20] For "plausible world" counterfactuals, see Gregory Hawthorn, *Plausible Worlds: Possibility and Understanding in History and the Social Sciences* (New York: Cambridge University Press, 1991). For "miracle counterfactuals," see Lewis, *Counterfactuals*, 75–77; James D. Fearon, "Causes and Counterfactuals in Social Science: Exploring an Analogy between Cellular Automata and Historical Processes," in *Counterfactual Thought Experiments in World Politics*, ed. Tetlock and Belkin, 39–67.

[21] For the term "minimal rewrite counterfactuals," see Tetlock and Belkin, "Counterfactual Thought Experiments," 23.

[22] For an analysis of US and Soviet missed opportunities to de-escalate the Cold War, see Deborah Welch Larson, *Anatomy of Distrust: U.S.-Soviet Relations during the Cold War* (Ithaca, NY: Cornell University Press, 1997).

The plausible/miracle world distinction is closely linked to the distinction between idiographic (or historiographical) counterfactuals and nomothetic (or generalizing) counterfactuals.[23] The former aims to explain or understand a particular historical episode or development, while the latter aims to explore more general theoretical arguments. Scholars using nomothetic counterfactuals begin with a theoretical model and then trace the logical implications of the model by analyzing its predictions under a wide range of different conditions, regardless of whether those conditions are historically realistic. The goal of nomothetic counterfactuals "is not historical understanding; rather, it is to pursue the logical implications of a theoretical framework."[24]

Counterfactuals are also used pedagogically—to encourage people to think through the logical implications of their arguments or theories, to confront inconsistencies, to recognize moral double standards, to empathize with those in particular situations or cultures, to recognize biases in their thought processes, to assess moral responsibility, or more generally to think about the world in different ways.[25] If I am convinced my historical interpretation is correct, but reject a counterfactual implication that follows logically from it, I need to modify my argument. Similarly, counterfactual thought experiments also can help people identify double standards in moral judgments, by forcing people to think about whether their emotional response to an event in the real world would be any different in a hypothetical world in which identities were changed. These kinds of consistency probes are one useful role for miracle counterfactuals, where there are no constraints on the feasibility of alternative worlds.

Scholars often use idiographic counterfactuals to facilitate judgments of moral responsibility by asking whether it is reasonable to expect that an individual political leader could have acted differently under the circumstances.[26] That requires the analyst to identify the range of choices facing political leaders and the likely consequences of alternative choices given existing constraints. Counterfactuals also provide a normative standard in the law; we often ask "what would a reasonable person have done."

Counterfactual thought experiments can be used to think through the implications of a theory when it is impossible to perform an experimental or empirical analysis. A famous example from physics is "Schrödinger's Cat."

---

[23] Tetlock and Belkin, "Counterfactual Thought Experiments," 7–10. The idiographic/nomothetic distinction refers to what is being explained, rather than how one explains it. A theory-driven explanation of an individual historical case is idiographic. Jack S. Levy, "Explaining Events and Testing Theories: History, Political Science, and the Analysis of International Relations," in *Bridges and Boundaries: Historians, Political Scientists, and the Study of International Relations*, ed. Colin Elman and Miriam Fendius Elman (Cambridge, MA: MIT Press, 2001), 45–47.

[24] Tetlock and Belkin, "Counterfactual Thought Experiments," 9.

[25] Ibid., 12–16; Lebow, *Forbidden Fruit*; Daniel Nolan, "Why Historians (and Everyone Else) Should Care about Counterfactuals," *Philosophical Studies* 163, no. 2 (March 2013): 317–35.

[26] Tetlock and Belkin, "Counterfactual Thought Experiments," 8.

Erwin Schrödinger proposed this experiment to help think through the implications of the "Copenhagen interpretation" of quantum mechanics, and the puzzle that a cat (if unobserved) could be simultaneously dead and alive.[27]

Scholars have used counterfactual thought experiments as a "debiasing tool" to help overcome certain inherent psychological biases, particularly the "hindsight bias."[28] There is substantial evidence that peoples' knowledge of a historical outcome increases the a priori probability they subsequently attach to that outcome. As a consequence, people lean more toward determinism than toward contingency in their interpretations of history and see more contingency in future events than in past events. Events that were once considered to be highly improbable are sometimes seen as overdetermined once they happen (the end of the Cold War, for example). People who are exposed to counterfactual scenarios about the past can, to a certain extent, overcome the hindsight bias. They are less likely to hold deterministic views and more likely to recognize the role of contingency.[29]

COUNTERFACTUALS AND CAUSATION

Much of the importance of counterfactuals derives from their relationship to the analysis of causation. Nearly all causal statements imply some kind of counterfactual. This is particularly clear for statements of necessary conditions, which are defined by a counterfactual. *Necessary condition counterfactuals* take the form "if $\sim x$, then $\sim y$" (if not *x* then not y). They are particularly important because many scholars define causation in terms of necessary conditions and because theoretical propositions and especially historical interpretations positing necessary conditions are fairly common in international relations and diplomatic history.[30] In fact, some qualitative

---

[27] John Gribbin, *In Search of Schrödinger's Cat: Quantum Physics and Reality* (New York: Bantam Books, 1984).

[28] Tetlock and Belkin, "Counterfactual Thought Experiments," 15. Baruch Fischoff and Ruth Beyth, "I Knew It Would Happen: Remembered Probabilities of Once-Future Things," *Organizational Behavior and Human Performance* 13, no. 1 (February 1975): 1–16.

[29] Philip E. Tetlock, *Expert Political Judgment*; (the full citation already appears in footnote 19); Lebow, *Forbidden Fruit*, chaps. 5–6. On the tension between promoting more disciplined thinking about counterfactuals and the psychological barriers that must be overcome in the process, see Ian S. Lustick, "Tetlock and Counterfactuals: Saving Methodological Ambition from Empirical Findings," *Critical Review* 22, no. 4 (2010): 427–47.

[30] For the necessary condition conception of causation, see Lewis, *Counterfactuals*. For alternative conceptions of causation see Henry E. Brady, "Causation and Explanation in Social Science," in *Oxford Handbook of Political Methodology*, ed. Janet Box-Steffensmeier, Brady, and David Collier (New York: Oxford University Press, 2008), 217–70. For examples of theoretical propositions or historical interpretations involving necessary conditions, see Gary Goertz and Jack S. Levy, eds., *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals* (New York: Routledge, 2007). H. Stuart Hughes wrote that "the most satisfactory type of causal explanation in history simply tries to locate the factor which, when removed, would make the decisive difference in a given sequence of events... [and]

methodologists argue that qualitative researchers generally conceive of causation in terms of identifying necessary and sufficient conditions for particular outcomes.[31]

Although the most dramatic form of counterfactuals involves necessary conditions and a reversal of outcomes, counterfactuals are not limited to necessary conditions. Any causal statement of the form "x is a cause of y" implies a counterfactual, because it implies that if the value of x were different, the outcome y probably would be different.[32] It is conceivable, however, that under some conditions the counterfactual world can have the same outcome as the real world. Some argue, for example, that the First World War would still have occurred if the assassination attempt against Franz Ferdinand had failed.

One kind of causal statement that has no specific counterfactual implications is one positing sufficient conditions. The statement if "x then y" is uninformative about what would happen in the absence of x. There may be another causal factor z that is also sufficient for y, and depending on the presence or absence of z, y may or may not occur in the absence of x. The statement "if x then y" does imply the contrapositive "if $\sim y$ then $\sim x$." This is a counterfactual, but one that refers to a change in the *outcome* variable, not the causal variable. It has no implications for what outcomes might look like in an alternate world in which a key causal variable takes on a different value. But the issue of counterfactuals involving sufficient conditions is complicated. If x is sufficient for y, and if x is absent, one possible causal path to y is eliminated. This presumably reduces the probability of y.

A persuasive confirmation or disconfirmation of a counterfactual hypothesis would have an important bearing on causal relationships in the real world. For hypothesized necessary conditions, counterfactual analysis is the only way to bring argument and evidence to bear on the causal statement. For other causal statements, evaluating their counterfactual implications provides an additional methodological tool for assessing their validity and an additional testable implication of the primary hypothesis. As James Fearon argues, "arguments about the relative importance of possible causes become arguments about the relative plausibility of different counterfactual scenarios."[33]

As noted above, counterfactual analysis is an additional tool for analyzing causal relationships. It is not a stand-alone method. For a particular

---

render the events in question inconceivable." Hughes, "The Historian and the Social Scientist," *American Historical Review* 66, no. 1 (October 1960): 29. See, also, Evans, *Altered Pasts*, 119.

[31] Goertz and Mahoney, *Tale of Two Cultures Qualitative and Quantitative Research in the Social Sciences* (Princeton, NJ: Princeton University Press, 2012), chaps. 3–6.

[32] This is particularly clear if one adopts a probabilistic conception of causality. Ellery Eells, *Probabilistic Causality* (New York: Cambridge University Press, 1991).

[33] James D. Fearon, "Counterfactuals and Hypothesis Testing in Political Science," *World Politics* 43, no. 2 (January 1991): 178.

historical case, counterfactual analysis needs to be combined with process tracing. It should also be combined, wherever possible, with the comparative method. If another historical case closely resembles the parallel world generated by a minimal-rewrite counterfactual in this case, it should be employed. That is a big "if," however, and one argument for the use of counterfactuals to assess causation is the absence of comparable cases.

One should not necessarily infer that if a small counterfactual change in a causal variable would lead to a substantial change in the outcome, then the causal variable in question is necessarily the "most important" cause of the outcome. This caution applies even in the case of necessary condition counterfactuals. Evaluating the relative weights of causal variables when they include necessary conditions raises difficult analytic issues. There may be other necessary conditions earlier on the causal chain leading to the outcome; sufficient conditions may play a critical role; and there may be multiple causal paths that could lead to the outcome.[34] A potentially game-winning pass might be dropped in the end zone near the end of the game, or the coach may make a controversial call, but there are usually countless actions earlier in the game that help explain the outcome. The last link in a causal chain is not always the most important.

Counterfactual analysis is often seen as a tool to support a contingent and anti-determinist world view. This is a central theme in Ferguson's essay, subtitled "Toward a 'Chaotic Theory' of the Past."[35] Jeremy Black states that his focus is on "the role of counterfactualism in demonstrating the place of contingency" and the limitations of determinist accounts. He subsequently argues that "Counterfactualism... is centrally linked to contingencies and conjunctures" and to human agency.[36] Similarly, Gavriel Rosenfeld argues that "alternate history is inherently anti-deterministic."[37] This helps to explain why many structuralists like E. H. Carr are so hostile to counterfactuals.

True, arguments about contingency draw strong support from a demonstration that an easily imaginable change in political leadership or key events would likely have led to a different outcome. Richard Ned Lebow, for example, argues that the First World War would almost certainly not have occurred without the assassination of the archduke and concludes that the war was contingent rather than inevitable or structurally determined.[38] It is

---

[34] On multiple causal paths, see Goertz and Mahoney, *A Tale of Two Cultures*; Jack S. Levy, "The 'Paths-to-War' Concept," in *What Do We Know about War?* 2nd ed., ed. John A. Vasquez (Lanham, MD: Rowman & Littlefield, 2012), 281–90.

[35] Ferguson, "Virtual History."

[36] Jeremy Black, *What If? Counterfactualism and the Problem of History* (London: Social Affairs Unit, 2008), vii, 10.

[37] Rosenfeld, *World Hitler Never Made*, 6.

[38] Richard Ned Lebow, "Contingency, Catalysts, and Nonlinear Change: The Origins of World War I," in *Explaining War and Peace*, ed. Goertz and Levy, 85–111; Lebow, *Forbidden Fruit*, chap. 3.

also true, however, that determinist arguments positing that a particular outcome was driven by structural and social forces derive strong support from a counterfactual analysis that persuasively argues that a different decision or a different leader probably would not have led to a different outcome.[39] Although the best candidates for counterfactual antecedents are those involving manipulations of contingent causal variables, this does not necessarily imply that the counterfactual outcome would deviate from the observed real-world outcome. That can only be determined by a rigorous and systematic counterfactual analysis. Contingency in antecedents does not necessarily imply contingency in consequents.

In fact, many serious applications of counterfactual methodology have concluded that the counterfactual world would not have been different. In his widely praised but contested counterfactual analysis of an Al Gore presidency in 2000, Frank Harvey argues that a Gore administration, like the Bush administration, would probably have gone to war in Iraq. Harvey uses this finding to question arguments for contingency and for the causal importance of George W. Bush in Iraq War decision making.[40] Paul Schroeder rejects the argument of Lebow and others that the First World War was contingent by arguing that the changes necessary to avoid the First World War were too many in number and too difficult to implement.[41]

Schroeder's counterfactual analysis departs from the standard method of asking whether a different outcome might have resulted from small changes in the real world. Instead, he analyzes the stability of an outcome by exploring its sensitivity to a range of variations in its antecedents. He basically asks the question of the conditions that would have been necessary for alternative outcomes to occur. This is an equally valid form of counterfactual analysis and a useful contribution to historical understanding.[42]

Those who emphasize contingency and those who emphasize the primacy of structures and social forces often reach diametrically opposed conclusions from their respective counterfactual analyses of the same historical cases. This is precisely what Carr warned about when he argued that counterfactuals are too speculative and self-serving to be analytically useful. To overcome these concerns, discipline counterfactual analysis, and prevent it from deteriorating into a "literature of the imagination," we need to develop

---

[39] For this reason it is quite problematic to define counterfactuals as changes that lead to different outcomes, as Evans does in *Altered Pasts*, xv.

[40] Frank P. Harvey, *Explaining the Iraq War: Counterfactual Theory, Logic, and Evidence* (New York: Cambridge University Press, 2012).

[41] Schroeder, "Embedded Counterfactuals."

[42] See also Charles Maier, "Wargames: 1914–1919," *Journal of Interdisciplinary History* 18, no. 4 (Spring 1988): 821; Martin Bunzl, "Counterfactual History: A User's Guide," *American Historical Review* 109, no. 3 (June 2004): 857.

a rigorous set of criteria, or best practices, for evaluating the scientific legitimacy and utility of idiographic counterfactuals.[43]

CRITERIA FOR EVALUATION

Counterfactual propositions have the same logical structure as conventional theoretical explanations or historical interpretations. They have initial conditions (the counterfactual antecedent), hypothesized consequences (the counterfactual consequent), and a theoretical generalization or causal path that explains the linkage between the antecedent and the consequent. The criteria for evaluating a counterfactual proposition are similar to those for evaluating any theoretical proposition: conceptual clarity, logical consistency, falsifiability, consistency with well-established theories and empirical generalizations, empirical scope, number and diversity of observable implications, and parsimony, among others.[44] In addition, in an idiographic counterfactual, the antecedent and the path from which it emerges from the real world must be plausible.

Philip E. Tetlock and Aaron Belkin were the first to systematically incorporate these criteria into rules for counterfactual analysis.[45] These rules include clarity, logical consistency or cotenability, historical consistency (the minimal-rewrite rule), consistency with well-established theoretical laws and statistical generalizations, and "projectability." Subsequent scholars have generally adopted these rules, with some modifications, and suggested some additional criteria: the need for sensitivity to second-order or redirecting counterfactuals and for comparative counterfactual analysis.[46] Here I present a slightly modified, reorganized, and expanded list of criteria.

Clarity

The analyst must clearly specify the counterfactual antecedent, consequent, the causal path and mechanisms linking the two, and the nature of the relationship (deterministic or probabilistic). If there are several causal paths through which the antecedent might lead to the consequent (counterfactual equifinality), or if the antecedent might lead to several possible consequents (counterfactual multifinality), those should be identified.

Although the general counterfactual statement that "if x were different, y would have been different" is helpful in differentiating the counterfactual

---

[43] On literature of the imagination, see Hawthorn, *Plausible Worlds*, 167.

[44] Carl G. Hempel, *Philosophy of Natural Science* (Englewood Cliffs, NJ: Prentice-Hall, 1966).

[45] Tetlock and Belkin, "Counterfactual Thought Experiments," 16–31.

[46] Lebow, *Forbidden Fruit*, chap. 2; Jack S. Levy, "Counterfactuals and Case Studies," in *The Oxford Handbook of Political Methodology*, ed. Box-Steffensmeier, Brady, and Collier, 640; [The full citation for the edited volume already appears in fn 30]

world from the real world, we want to know more about the nature and magnitude of those differences. In some cases the differences would seem to be obvious. If condition/event/individual *x* led to war, the necessary condition counterfactual that war would not have occurred in the absence of *x* indicates an unambiguous change in outcomes and constitutes a powerful statement of the causal importance of *x*. Even here, however, greater specificity might be useful.

With respect to the counterfactual world defined by the hypothetical non-assassination of Franz Ferdinand in June 1914, it would make a difference, in terms of Austria-Hungary's response and the likelihood of war, whether this outcome was the result of a missed shot by Gavrilo Princip, a shot by Princip that missed Ferdinand but killed his wife Sophie, no shot by Princip after the explosion of the first bomb on the bridge, or no assassination attempt at all. With respect to the outcome, does the hypothesized absence of war refer to the absence of a local Austro-Serbian war or of a world war, and in the summer of 1914 or during the next several years? How one interprets a counterfactual argument might depend on the precise specification of its antecedent and consequent.

The theoretical utility of a counterfactual, however, is not necessarily a monotonically increasing function of its degree of specificity.[47] The greater the detail in the specification of the antecedent, the less likely it is to occur, undermining the plausibility criterion. Moreover, 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, so that the probability of a detailed, compound outcome resulting from a long causal chain is fairly small. As Lebow argues, "it is highly unlikely that hypothesized antecedents will produce *specific* consequences at any temporal remove."[48] This suggests that the most useful counterfactuals involve antecedents and consequents with a moderate degree of specificity.

The Minimal-Rewrite Rule

The assessment of causality—whether through experimental, statistical, comparative, or counterfactual methods—requires the demonstration that any change in the value of an outcome variable can be traced to the effects of a single causal variable or combination of variables and not to confounding variables or extraneous influences. The model research design is an experimental study with random assignment, but that is not feasible for many

---

[47] Specificity would be increased by narrowing the timeframe for the occurrence of an event, identifying multiple consequences or outcomes, or specifying a long causal chain involving many links.

[48] Lebow, *Forbidden Fruit*, 56.

research questions. The method of controlled comparison, following John Stuart Mill's methods of agreement and of difference, tries to select carefully matched cases that are similar in every respect but one.[49] Quantitative matching methods follow a similar logic.[50] Ideally, counterfactual analysis should do the same.[51] As Weber argued, counterfactual analysis should make as few changes as possible from the real world. Gregory Hawthorn emphasizes the "closest possible world" test, Tetlock and Belkin suggest the "minimal-rewrite of history" rule, and Martin Bunzl emphasizes "the horizon of possibility."[52] Note that the minimal-rewrite rule applies to the antecedent, not to the consequent. Assessing the magnitude of the change in the consequent is the end point, not the starting point, of counterfactual analysis.

Just as Mill's methods, wherein controlled comparison and matching face the difficulty of finding real-world cases that are identical in all respects but one, the method of counterfactual analysis has difficulty identifying counterfactual antecedents that are identical to the real world in all respects but one. The social world is a highly interconnected system. A change in one variable induces changes in other variables, which then ripple through the system, generating further changes. As Garrett Hardin states, "we can never do merely one thing."[53]

Although "surgical" counterfactuals that change only one factor are ideal types that are never achieved in practice, they can sometimes be approximated.[54] The argument that the First World War would not have occurred in the absence of the assassination of Archduke Ferdinand is an excellent example of a minimal-rewrite counterfactual, given the number of highly contingent events leading to the assassination and how easily the plot could have failed. In fact, the ex ante probability of a failed assassination attempt was undoubtedly greater than that of a successful assassination. This suggests a corollary to the minimum-rewrite rule: the higher the ex ante probability of the occurrence of the counterfactual antecedent, and the lower the ex ante probability of the real world outcome, the better the minimal rewrite.

Another good example of a minimal-rewrite counterfactual is the proposition that if Al Gore rather than George W. Bush had been US president in 2003, the United States would not have invaded Iraq. One has to change

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

[50] Morgan and Winship, *Counterfactuals and Causal Inference*, chap. 4.

[51] John Lewis Gaddis argues that counterfactuals are "the historian's virtual equivalent of laboratory experimentation." Gaddis, *The Landscape of History: How Historians Map the Past* (New York: Oxford University Press, 2002), 100. Ingo Rohlfing suggests the minimum-rewrite rule basically uses the counterfactual to re-create Mill's method of difference. Rohlfing, *Case Studies and Causal Inference: An Integrative Framework* (New York: Palgrave Macmillan, 2012), 176.

[52] Weber, *Methodology of the Social Sciences*, chap. III; Hawthorn, *Plausible Worlds*, 31–60; Tetlock and Belkin, "Counterfactual Thought Experiments," 23–25; Bunzl, "Counterfactual History," 857.

[53] Cited in Robert Jervis, *System Effects* (Princeton, NJ: Princeton University Press, 1997), 10.

[54] For "surgical" counterfactuals, see Lebow, *Forbidden Fruit*, 50.

very little history to imagine a Gore victory in the 2000 election—just a few hanging chads on the ballots in Florida—to generate this alternative world.[55] Whether a Gore victory would have prevented a war with Iraq is a different and more difficult question. The plausibility of the antecedent and the conditional plausibility of the consequent given the antecedent are two distinct questions.

Many examples of minimal-rewrite counterfactuals involve a change in individual political leaders. Assassinations, deaths in office, and close elections all provide easily imaginable ways to substitute one political leader (and their belief systems, personalities, risk orientations, etc.) for another without changing the structure of the political and economic systems, social attitudes, and system-level structures. Imagining a change in the state's leading decision maker that does not involve changes in other causal variables, and providing a compelling argument that the outcome would or would not have changed, is a useful way of analyzing the causal role of individual leaders.[56] Benajmin Jones and Benjamin Olken, arguing that success or failure of assassination attempts (unlike assassination attempts themselves) is exogenous, demonstrate that successful (but not unsuccessful) assassination attempts statistically affect the likelihood of war and of institutional change, providing strong evidence of the causal role of individual leaders.[57] Similarly, in her study of the impact of threat perception on a leader's intervention strategies and policies, Elizabeth Saunders argues that the shift from the John Kennedy to the Lyndon Johnson presidency was an exogenous event that led to a change in threat perception (from Kennedy to Johnson) but in few other causal variables.[58]

Military battles that historians agree could easily have turned out differently, and especially if they should have turned out differently given the balance of military forces and contemporaries' expectations, are often good candidates for plausible world counterfactuals.[59] Herodotus's counterfactual of a Persian victory over the Greeks at Salamis in 480 BCE is a good example, given the significant numerical superiority of the Persians. For the same reason, a reversal of the "miracle on the Marne" in September 1914, and a German victory, would have been a minimal rewrite of history. Exogenous natural events can also be used as minimal rewrites. Geoffrey Parker argues that in the absence of unexpected shifts in the wind that helped defeat the

---

[55] For other minimal-rewrite paths to a Gore presidency, see Harvey, *Explaining the Iraq War*, 25; Lebow, *Forbidden Fruit*, 44–45.

[56] Robert Jervis, "Do Leaders Matter and How Would We Know?" *Security Studies* 22, no. 2 (April 2013): 153–79. For a useful application see Harvey, *Explaining the Iraq War*.

[57] Benjamin Jones and Benjamin Olken, "Hit or Miss: The Effects of Assassination on Institutions and War," *American Economic Journal: Macroeconomics* 1, no. 2 (July 2009): 55–87.

[58] Elizabeth N. Saunders, *Leaders at War: How Presidents Shape Military Interventions* (Ithaca, NY: Cornell University Press, 2011), chaps. 4–5.

[59] For examples see Robert Cowley, *What If? The World's Foremost Military Historians Imagine What Might Have Been* (New York: G. P. Putnam's Sons, 1999).

Spanish Armada, the war could easily have ended with a Spanish victory and Spanish control over England, at least for a time.[60]

Hypothetical changes in the nature of a state's political or economic systems, cultural attitudes, organizational culture, or in the structure of power or wealth in the international system are much less likely to satisfy the minimal-rewrite rule. These structural and cultural characteristics usually emerged gradually over time. If the processes giving rise to them are changed, those changes would have introduced other changes with consequences that would have rippled through the system in unpredictable ways. The counterfactual proposition that the Cold War would not have occurred if the Soviet Union had been democratic is a good example. Whatever it might have taken to create a democratic Soviet Union by 1945, those processes would have changed so many other things that one would be unable to trace the hypothesized absence of the Cold War to democratic institutions and political cultures.[61]

Another approach to the problem of identifying plausible counterfactual worlds focuses on the choice sets of political leaders. Ferguson argues that "we should consider as plausible or probable only those alternatives which we can show on the basis of contemporary evidence that contemporaries actually considered."[62] As we implied earlier, the closer leaders came to making another choice, the more minimal the rewrite.[63] Paul Schroeder accepts the criterion of alternatives actually considered but insists that historians also examine "both those possibilities and alternatives contemporaries saw and considered, and those they failed to see at all or to consider seriously." They should ask "What other decisions and actions could the historical actors have made under the existing circumstances" and with what likely consequences?[64]

The analyst must also explain why the option was neglected. If an option was ignored because of an accidental interruption in the flow of information in a system that normally functions smoothly, or because an

---

[60] Geoffrey Parker, "Repulse of the English Fireships: The Spanish Armada Triumphs, August 8, 1588," in *What If?* ed. Crowley, 139–54.

[61] This does not mean that the democratic USSR counterfactual could not be used (as a reviewer suggests) to probe the argument that bipolar systems are inherently unstable. Such an analysis, however, would be more of a miracle counterfactual than a minimal-rewrite counterfactual. It would be used to develop the implications or logic of a theoretical argument rather than to assess causality in this specific case.

[62] Ferguson, "Virtual History," 86–87. Ferguson imposes the further restriction that these alternatives must not only be considered but also committed to paper. This is too restrictive, because some things are not committed to paper, and some records are lost. Levy, "Counterfactuals and Case Studies," 636–37; Lebow, *Forbidden Fruit*, 48.

[63] Giovanni Capoccia and R. Daniel Kelemen suggest that legitimate counterfactuals include "only policy options that were available, considered, and narrowly defeated by relevant actors." Capoccia and Kelemen, "The Study of Critical Junctures: Theory, Narrative and Counterfactuals in Historical Institutionalism," *World Politics* 59, no. 3 (April 2007): 356.

[64] Schroeder, "Embedded Counterfactuals," 159–60, 342.

erratic and unpredictable leader chose for unclear reasons not to consider that option, then the use of that option as a counterfactual antecedent would qualify as a minimal rewrite. It would not be a minimal rewrite, however, if leaders neglected an option because they believed that the state did not have resources to implement it, that its costs would exceed its benefits, or that they would suffer politically from an adverse domestic reaction.

This scenario is clear in game-theoretic models of interactive decision making. These models explicitly specify what would happen if actors made different choices, how other actors would strategically respond, and the payoffs that each actor would receive for every possible combination of choices. Choices not made, and the sequence of choices that would have followed from them, are "off the equilibrium path," as defined by the criterion of subgame perfect equilibrium. Those choices would not qualify as minimal-rewrite counterfactuals. However, many classes of games have multiple equilibria. These multiple equilibria satisfy the minimal-rewrite criterion for plausible world counterfactuals because the choices that define them leave all actors no worse off.[65]

Some identify a tension between the need for rigorous criteria for counterfactual analysis—particularly the minimum-rewrite and proximity criteria—and the goal of generating historically or theoretically meaningful counterfactuals. Schroeder, for example, argues that "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."[66] Schroeder raises an important point, but he underestimates the extent to which some contingent events can have potentially profound consequences. The examples of the 1914 assassination of the Austrian Archduke and the 2000 US presidential election, along with numerous military contests that could easily have ended differently, suggest that small changes sometimes lead to large consequences. This is also clear in the theoretical literature on critical junctures, positive feedback, and path dependency.[67] Moreover, counterfactual analyses can also contribute to our understanding of particular historical cases if they demonstrate that a different leader or choice would have had little impact on the outcome.

Lastly, there is a neglected benefit of the minimal-rewrite criterion. Merely asking whether a hypothetical antecedent satisfies this criterion can contribute to historical understanding by forcing the investigator to confront

---

[65] Bueno de Mesquita, "Counterfactuals and International Affairs"; Barry R. Weingast, "Off-the-Path Behavior: A Game-Theoretic Approach to Counterfactuals and Its Implications for Political and Historical Analysis," in *Counterfactual Thought Experiments in World Politics*, ed. Tetlock and Belkin, 230–43.

[66] Schroeder, "Embedded Counterfactuals, 158.

[67] Ruth Berins Collier and David Collier, *Shaping the Political Arena: Critical Junctures, the Labor Movement, and Regime Dynamics in Latin America* (Princeton, NJ: Princeton University Press, 1991); Capoccia and Kelemen, "Study of Critical Junctures."

new questions about causal connections and interdependencies. For example, the question of whether the United States would have entered the First World War were it not for German submarine warfare, and whether the absence of the latter could be treated as a minimal-rewrite antecedent, raises important questions about German preferences, perceptions, risk propensities, and politics.

Cotenability

In introducing a counterfactual antecedent and explaining how history would have developed, the analyst must make additional assumptions about what else might change or stay the same, given the interconnectedness of the social world. The election of Al Gore as US president in 2000 would have changed not only the belief systems and personality of the president but also the identity and views of the president's leading advisors and top-level governmental officials, the relationship between the president and the Congress, and possibly other variables as well. The counterfactual analysis would have to make assumptions about these other changes and their potential consequences for US foreign policy and particularly for the key question of the decision (or not) to invade Iraq. These assumptions—labeled "connecting principles" (Goodman) or "enabling" counterfactuals" (Lebow)—are necessary to sustain the primary counterfactual.[68] These connecting principles should be explicit and consistent with the hypothesized linkages from the real world to the antecedent and from the antecedent to the hypothesized consequent, and with each other. Goodman refers to this requirement of logical consistency as "cotenability."[69]

The failure to satisfy the cotenability criterion is a common problem in the use of counterfactuals in historical analysis. One widely repeated example is Robert Fogel's pathbreaking attempt to assess the contribution of railroads to American economic development through a quantitative counterfactual analysis of how the American economy would have developed in the absence of railroads. Fogel argues that in the absence of the railroads other technologies (as well as an expanded system of canals) would have developed in their place, including the internal combustion engine and the automobile.[70] Jon Elster points out, however, these assumptions are not internally consistent, because the technology upon which the automobile was based would have allowed for the development of the railroad.[71]

---

[68] Good, *Fact, Fiction, & Forecast,* 17; [the full citation already appears in footnote 14]

[69] Goodman, *Fact, Fiction, & Forecast*, 15.

[70] Fogel, *Railroads and American Economic Growth*.

[71] Elster, *Logic and Society*, 204–8. It is conceivable that non-technological factors—laws, financial regulations, or political pressure—might have blocked the development of railroads but not of alternative technologies.

Consistency with Well-Established Theoretical Generalizations

A good counterfactual requires not only a plausible antecedent and an explanation for how it might arise, but also a credible causal path leading from the antecedent to the hypothesized consequent. Given that the counterfactual antecedent never occurred, we have no observations to validate our predictions, so standard empirical methods are not viable. We have information about the initial conditions in the counterfactual world, which are nearly identical to the point of divergence or branching point in the real world. But how do we predict the likely consequences of those initial conditions, given the nearly infinite number of possibilities? The absence of observable empirical data rules out a bottom-up, inductive, process-tracing approach. Theory must play a central role. In the absence of theory there is no constraint on the kinds of counterfactual arguments that might be made, fulfilling Carr's worst fears. As Edgar Kiser and Margaret Levi argue, "If an analyst approaches an historical problem purely inductively... the number of potential counterfactuals is practically infinite."[72] These considerations lead J. D. Gould to argue that "what the historian must look for in using counterfactual arguments... is *relevant lawlike statements*." Similarly, the economic historian Fogel argues that "statements about [counterfactuals] can only be justified by hypothetico-deductive systems." Tetlock and Belkin argue that good counterfactuals should be consistent with "well-established theoretical laws."[73]

Counterfactual analysis is even more of a theory-driven activity than is conventional empirical analysis. As Tetlock and Parker suggest, counterfactuals are "theoretically constructed" rather than "empirically discovered."[74] The problem is that we have few well-established laws of international relations. With precious few exceptions, we have, at best, contested theories and probabilistic relationships of modest strength that are conditional and often contextually dependent rather than universal. However, given the aim of explaining a single case and exploring a counterfactual under a specific set of historical conditions, contingent generalizations can often substitute for more general theoretical laws. We can also use a plurality of different theories to substantiate different links in the causal chain, as long as each theory is well established in its own domain and the propositions derived from those theories are reasonably consistent with each other.[75] This represents

---

[72] Edgar Kiser and Margaret Levi, "Using Counterfactuals in Historical Analysis: Theories of Revolution," in *Counterfactual Thought Experiments in World Politics*, ed. Tetlock and Belkin, 191.

[73] J. D. Gould, "Hypothetical History," *Economic History Review*, 2nd ser., 22, no. 2 (August 1969): 197; Fogel, *Railroads and American Economic Growth*, 246; Tetlock and Belkin, "Counterfactual Thought Experiments," 25–27.

[74] Tetlock and Parker, "Counterfactual Thought Experiments," 41n33; Levy, "Counterfactuals and Case Studies," 641.

[75] Rudra Sil and Peter J. Katzenstein, *Beyond Paradigms: Analytic Eclecticism in the Study of World Politics* (New York: Palgrave Macmillan, 2010).

a step back from the use of genuine theoretical laws, but recall that even Carl G. Hempel argued that most historical explanations were "explanation sketches" that fall short of a deductive-nomological model.[76]

Tetlock and Belkin identify *consistency with well-established statistical generalizations* and *projectability* as two additional criteria for the evaluation of counterfactuals.[77] Each of these criteria, however, is in fact a precondition for the existence of a well-established theoretical generalization. For this reason it is preferable to treat consistency with well-established statistical generalizations and projectability as subsets of the broader category of consistency with well-established theoretical generalizations.

It is necessary, however, to make one modification in this criterion. Although consistency with well-established statistical generalizations is clearly an important criterion by which we can say that a theory is well established, it should not be treated as a necessary criterion for a valid counterfactual argument. It goes too far to say that "counterfactual inferences are justified... if and only if they are embedded in a system of statistical contingency for which we have reasonable evidence."[78] The set of well-established statistical generalizations in security studies, as in many other fields, is relatively limited.[79] What statistical evidence would suffice for validating a counterfactual inference that US president Al Gore would (or would not) have decided to go to war with Iraq in 2003? The conventional wisdom and Harvey's critique are each based on the cumulation of different kinds of evidence, including comparative and single case studies of belief systems, as well as quantitative studies of economic and social trends.[80] Although statistical evidence is often the most compelling, in its absence other forms of evidence can be informative. In some cases, the documentary record leaves traces of what leaders intended to do under various contingencies, though these must be approached critically. Thus I prefer the category of "consistency with the empirical evidence" to the category of consistency with well-established statistical generalizations.

Tetlock and Belkin, adopting Nelson Goodman's concept of projectability, argue that law-like generalizations underlying counterfactual propositions must be able "to predict what will happen in new, hitherto unobserved cases"

---

[76] Carl G. Hempel, "The Function of General Laws in History," *Journal of Philosophy* 39, no. 2 (January 1942): 42.

[77] Tetlock and Belkin, "Counterfactual Thought Experiments," 27–31.

[78] Robyn M. Dawes, "Counterfactual Inferences as Instances of Statistical Inferences," in *Counterfactual Thought Experiments in World Politics*, ed. Tetlock and Belkin, 304.

[79] The existence of statistically based sports analytics that report the probability of the final win/lose outcome at every single point in the game provide a much stronger basis for counterfactual inferences for "what if" events during the game. For football, see http://www.advancedfootballanalytics.com. I thank Paul Poast for alerting me to this source.

[80] On different case study designs, see George and Bennett, *Case Studies and Theory Development*; Jack S. Levy, "Case Studies: Types, Designs, and Logics of Inference," *Conflict Management and Peace Science* 25, no.1 (Spring 2008): 1–18.

and be falsifiable.[81] They go on to say that "counterfactuals that are devoid of testable implications in the actual world leave us marooned in hypothetical worlds of our own subjective making."[82] In other words, we cannot support a counterfactual inference with ad hoc theoretical arguments that we construct solely for that specific instance. The theoretical argument underlying the path from the counterfactual antecedent to consequent must have observable implications that can be tested somewhere in the real world. The greater the number and variety of those implications and successful tests, the more confidence we can have in the hypothesized theoretical linkage.[83]

Historical Accuracy

Although Tetlock and Belkin focus their criterion of "consistency with well-established historical facts" around the minimal-rewrite rule in assessing the plausibility of the antecedent, historical accuracy is also relevant in evaluating the plausibility of the predicted consequences of the antecedent in the alternative world.[84] A counterfactual explanation of expected behavior in the hypothetical world begins with the initial conditions defined by the antecedent, which are identical (other than the minimal rewrite) to the conditions in the real world at the point of divergence. These initial conditions include descriptive historical information about individual values, preferences, beliefs, perceptions, and personalities; social and economic conditions; the nature of the political system, key stakeholders and their interests and relative influence, and public attitudes; relationships with friends and adversaries; and system structures. Any analysis of the path from the antecedent to the consequent in the counterfactual world must be consistent with these initial conditions and with plausible arguments (based on theoretical generalizations invoked to support the counterfactual) about how those initial conditions might (or might not) have changed.

Harvey's examination of the likelihood of an Iraq War under a Gore presidency provides a good example of how to bring a wide array of factual information to bear in a rigorous and systematic counterfactual analysis. Harvey analyzes the belief system and personality of Al Gore, the beliefs of his likely political and security advisors, organizational constraints, societal pressures and public opinion, the failures of the US intelligence system, Gore's likely interactions with Saddam Hussein and allied leaders, and other theoretically relevant factors. Harvey's analysis demonstrates that good

---

[81] For Goodman's concept of projectability, see Goodman, *Fact, Fiction, & Forecast*, chap. 4. Tetlock and Belkin, "Counterfactual Thought Experiments," 30–31 (quotation).

[82] Tetlock and Belkin, "Counterfactual Thought Experiments," 30–31.

[83] Gary King, Robert Keohane, and Sydney Verba, *Designing Social Inquiry* (Princeton, NJ: Princeton University Press, 1994).

[84] Tetlock and Belkin, "Counterfactual Thought Experiments," 23–25.

counterfactual analysis includes a significant empirical as well as theoretical component.[85] This requires that the analyst be well immersed in the history of the case and in the secondary literature. Earlier I noted Fogel's argument that hypothetico-deductive models, or at least approximations of them, are necessary for counterfactual analysis. Fogel concludes by arguing that these models should be applied "by scholars deeply immersed in the materials of history. Only the scholar who knows what is *unique*, *special*, and *particular* about a given historical problem can successfully adopt powerful general methods to the study of that problem. The casual interloper cannot possess this knowledge."[86]

Temporal Proximity

Counterfactuals involve predictions about what would have followed from a change in real-world conditions. Our confidence in the validity of counter-factual predictions is a function of the temporal distance and length of the causal chain from antecedent to consequent. The greater the distance and the longer the causal chain, the greater the number of opportunities for breaks in the causal chain, for other events and processes to intervene and shift history onto a different path, making the final outcome less predictable. As Gould argues, "almost all of the obstacles to accurate prediction grow, some of them exponentially, as the time-horizon is extended."[87] Similarly, Sidney Hook states that "when we draw the line of possible eventuality too far out of the immediate period, the mind staggers under the cumulative weight of the unforeseen."[88] These considerations lead Ian Kershaw to argue that historians should use "short-range counterfactuals" rather than engage in an "intellectual guessing-game of looking into some distant future."[89]

Even if each link in a causal chain involves short-term regularities that are well understood and highly predictable, small disturbances, especially if relationships are nonlinear, can reverberate through the system, generate large irregularities, and make it impossible to make anything other than short-term predictions.[90] Fearon illustrates this pattern in his discussion of the dynamics of a very simple system of cellular automata, which produce short-term regularities but long-term irregularities. These patterns are magnified in more complex social worlds. These considerations lead Fearon to propose a *proximity* criterion for evaluating counterfactuals. He argues that we should

---

[85] Harvey, *Explaining the Iraq War*.

[86] Fogel, *Railroads and American Economic Growth*, 249.

[87] Gould, "Hypothetical History," 199–200.

[88] Sidney Hook, *The Hero in History: A Study in Limitation and Possibility* (New York: John Day, 1943), 134; cited in Fearon, "Causes and Counterfactuals in Social Science," 66.

[89] Ian Kershaw, *Fateful Choices: Ten Decisions that Changed the World, 1940–1941* (New York: Penguin, 2007), 6.

[90] James Gleick, *Chaos: Making a New Science* (New York: Viking, 1987).

evaluate the plausibility of a counterfactual "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."[91] This does not necessarily rule out counterfactuals involving longer causal chains, but it does require that we acknowledge the much greater uncertainty associated with those outcomes.[92]

Redirecting Counterfactuals

Long causal chains provide more opportunity for subsequent developments that change the course of history, sometimes leading history back to its original course. Consider a "powder keg" model, in which certain structural conditions and a proximate cause are individually necessary and jointly sufficient to generate a particular outcome.[93] If a specific spark is not present we must ask whether another spark might arise later, either exogenously or through strategic behavior. The "streetcar" metaphor expresses the same idea. It is often argued that an important cause of US escalation in the Vietnam War was the 1965 Vietcong attack on the US military base at Pleiku.[94] McGeorge Bundy, President Kennedy's national security advisor, provided an implicit critique of this argument when he said that "Pleikus are like streetcars; if you miss one, another will come along shortly."[95] Note that the source of the spark could be another unpredictable Vietcong attack or deliberately provocative behavior by the United States to escalate the war under conditions where it could blame its adversary.

Thus the counterfactual analyst must be sensitive to the possibility that an exogenous or endogenous event might return history to its original course. Lebow labels these "second-order counterfactuals."[96] I prefer the term "reversionary counterfactuals" or "redirecting counterfactuals," to focus on events or actions that redirected history back to the path it was on prior to the point of divergence into the counterfactual world and to distinguish them from other kinds of second-order counterfactuals.[97]

---

[91] Fearon, "Causes and Counterfactuals in Social Science," 66.

[92] Some trends, like the value of the overall stock market, are easier to predict over the long term than the short term.

[93] Goertz and Levy, "Causal Explanation, Necessary Conditions, and Case Studies," 36–39.

[94] In the *Pentagon Papers* the discussion of the American escalation begins with a chapter entitled "Introduction—Pleiku Pulls the Trigger." See Mike Gravel, *The Pentagon Papers: The Defense Department History of United States Decision-making on Vietnam*, vol. III (Boston: Beacon Press, 1971), 286–87.

[95] Quoted in James G. Blight, Janet M. Lang, and David A. Welch, *Vietnam If Kennedy Had Lived: Virtual JFK* (Lanham, MD: Rowman & Littlefield, 2009), 171.

[96] Lebow, *Forbidden Fruit*, 57–58.

[97] For "reversionary counterfactuals," see Tetlock and Parker, "Counterfactual Thought Experiments,"19. For "redirecting counterfactuals," see Levy, "Counterfactuals and Case Studies," 680.

We can apply the concept of a redirecting counterfactual to Lebow's argument that in the absence of the assassination of the archduke in 1914, the continued growth in Russian power would have compelled the German military, by 1917, to abandon the Schlieffen Plan and adopt a more defensive strategy. This would have minimized the incentives for a preventive or preemptive war in any subsequent crisis, generated more risk-averse German behavior, and significantly reduced the probability of war.[98] Lebow's vision of the world in 1917 in the absence of the Great War is quite plausible—sufficiently plausible, in fact, that it would have been shared by German military and political elites, who would have been highly motivated to prevent that world from ever arising. The same preventive logic that many scholars (including Lebow) argue were critical in driving the German decision for war in 1914 would have been even more compelling as Russian power continued to grow and as the point of no return approached.[99] In all likelihood, German strategic behavior in response to shifting power would have led the counterfactual world back to the world war of the real world.

Comparative Counterfactual Analysis

In questioning Lebow's 1914 counterfactual, I am suggesting an opposing counterfactual, one predicting the probability of war rather than peace in the aftermath of a failed assassination attempt in 1914. Who is right? The answer requires a systematic comparative analysis of each of the two counterfactuals, pitting one against the other and based on the criteria outlined above.[100] As Harvey effectively argues, all counterfactual analysis is inherently comparative.[101] Imre Lakatos's conception of a three-cornered test—testing a theory against both the evidence and leading alternative theories—applies to the counterfactual world as well as to the real world.[102]

---

[98] Lebow, "Contingency, Catalysts, and Nonlinear Change"; Lebow, *Forbidden Fruit*, chap. 3. Lebow's counterfactual includes additional causal paths leading away from war after a failed assassination attempt in 1914.

[99] For factors driving Germany to war, see Fritz Fischer, *Germany's Aims in the First World War* (New York: Norton, 1976), chap. 2; Jack S. Levy, "The Sources of Preventive Logic in German Decision-making in 1914," in *The Outbreak of the First World War: Structure, Politics, and Decision-Making*, ed. Levy and John A. Vasquez (Cambridge: Cambridge University Press, 2014), 139–66.

[100] I would begin with the question of whether the factors that combined with shifting power and preventive logic to facilitate the German decision for war in 1914 would have been likely to persist or reoccur in the two or three years after 1914.

[101] Harvey, *Explaining the Iraq War*, chap. 1. Similarly, Goodman argued that "[T]he counterfactual *V* is invalidated not by lack of a law upholding it but by conflict with a more strongly upheld conditional." Goodman, *Fact, Fiction, and Forecast*, 14, 122.

[102] Imre Lakatos, "Falsification and the Methodology of Scientific Research Programs," in *The Methodology of Scientific Research Programs, Philosophical Papers*, vol. 1, by Imre Lakatos, ed. John Worrall and Gregory Currie (New York: Cambridge University Press, 1978), 8–101.

THE ANALYTIC UTILITY OF PLAUSIBLE WORLDS

Counterfactuals serve many useful purposes. I focus on the role of idiographic (historiographical) counterfactuals in case study research, with the primary aim of suggesting how scholars can use counterfactuals, preferably in combination with comparative methods and process tracing, to help validate causal inferences in particular historical episodes. Nearly all causal statements have counterfactual implications. An analysis of the validity of those counterfactual implications is the only within-case way of assessing the validity of causal arguments specifying necessary conditions and a useful additional tool for evaluating other kinds of causal statements. By asking new questions about an historical episode, counterfactual researchers identify additional testable implications of the original causal proposition and bring additional leverage to bear on their historical interpretation. Many debated the causal weight of George W. Bush's beliefs and personality on his decision for war in Iraq, but few thought that an analysis of the belief systems of Gore and his prospective advisors might be relevant for resolving that debate, as Harvey's counterfactual analysis makes clear.[103]

E. H. Carr speaks for many critics in arguing that counterfactual analysis is arbitrary, speculative, self-serving, and consequently non-falsifiable. In the last two decades, however, social scientists and historians have begun to develop standards for evaluating the utility of historiographical counterfactuals in exploring how (and how easily) history could have moved in a different direction and how such knowledge might be used to help explain what actually happened. I build on the existing literature but go beyond it in constructing a revised set of criteria. Those criteria include the clarity of antecedents, consequents, and the causal path(s) linking them; a minimal rewrite of history from the real world to the counterfactual antecedent; cotenability, or logical consistency among secondary counterfactuals necessary to support the primary counterfactual; consistency with well-established theoretical generalizations, including consistency with the empirical evidence and the projectability of the theoretical generalization underlying the counterfactual prediction back to observable real world cases; historical accuracy with respect to the initial conditions at the point of divergence; temporal proximity; sensitivity to redirecting counterfactuals that return history to its original path; and comparative counterfactual analysis.

In specifying these criteria, I am not suggesting that counterfactual analysis in case studies is an easy and straightforward exercise. To the contrary, it is extraordinarily difficult. Social scientists and historians have enough trouble establishing causality in the observable world. In counterfactual analysis, they must do so in a hypothetical world where the full set of initial conditions for the counterfactual argument do not exist and where there is no

---

[103] Harvey, *Explaining the Iraq War*.

sequence of observable events to analyze. This places an even greater reliance on theory than in conventional empirical analysis, in a context where we have few validated empirical laws of international relations. Predicting the consequences of a counterfactual antecedent involves the same analytic and empirical problems as forecasting the future. Today's predictions of the future are tomorrow's historical counterfactuals. The well-known difficulties of forecasting the future should serve as a cautionary tale for those expressing too much confidence in their counterfactual arguments about the past. It is ironic that some scholars use counterfactuals to argue that the world is contingent, but then make essentially deterministic arguments about what would have happened after a contingent change in the real world.

Until recently, substantive debates about the "what ifs" of history have far outpaced the methodology of counterfactual analysis for evaluating those debates. That has begun to change, as analysts have developed rigorous criteria for counterfactual analysis. The aim is to move counterfactual analysis from speculative acts of the imagination to a system for disciplined thinking about alternative paths that history might have taken in order to enhance our understanding of what actually did happen. Contrary to Oakeshott's assertion that counterfactuals are "a monstrous incursion of science into the world of history," we can use science to enhance our understanding of history through the judicious use of counterfactuals, ideally in conjunction with within-case process tracing, the comparative method, and other approaches.[104] If we make counterfactuals explicit, ground them in theory, and evaluate them in terms of logically acceptable rules of inference, counterfactuals can provide an additional methodological tool for evaluating causality in a complex world. Paraphrasing Ferguson, and invoking the Rankean dictum that the historian should study history "as it really was," the rigorous study of history as it really wasn't can help us understand history as it really was.[105]

ACKNOWLEDGMENTS

---

[104] Oakeshott, *Experience and Its Modes*, 128–29 (quotation).

[105] Ferguson, "Virtual History," 87; Leopold von Ranke. *The Theory and Practice of History*, ed. Georg G. Iggers and Konrad von Moltke (Indianapolis, IN: Bobbs-Merrill, 1973).