Project MUSE®
http://muse.jhu.edu

Review Articles

WHAT'S SO DIFFERENT ABOUT A COUNTERFACTUAL?

By RICHARD NED LEBOW*

Niall Ferguson, ed. *Virtual History: Alternatives and Counterfactuals*. New York: Basic Books, 1999, 500 pp.
Niall Ferguson. *The Pity of War: Explaining World War I*. New York: Basic Books, 1999, 563 pp.

BECAUSE the siege of Troy was proving difficult, the frustrated Argives went down to their ships to sail for home. According to Homer, "A homecoming beyond fate might have been accomplished, had not Hera spoken a word to Athene." The grey-eyed Athena sped down the peak of Olympus and instructed Odysseus to prevent the departure of the Argives, which he did.[1] The drama of the *Iliad* rests on a counterfactual, as do many other great works of literature. Counterfactuals are also taken seriously in the physical and biological sciences, where researchers routinely use them to develop and evaluate sophisticated, nonlinear models.[2] They have been used with telling effect in American politics, and other political scientists could greatly benefit from wider use of this useful research tool.[3] But for most members of

* I would like to thank Richard Hamilton, Richard Herrmann, Edward Ingram, Friedrich V. Kratochwill, Eli Bohmer Lebow, Geoffrey Parker, Janice Gross Stein, and Philip E. Tetlock for their helpful comments.

[1] *The Iliad of Homer*, trans. Richard Lattimore (Chicago: University of Chicago Press, 1951), bk. 2, lines 135–210.

[2] There are numerous examples from physics. On using information from nonevents to test nuclear weapons, see R. Penrose, *Shadows of the Mind: A Search for the Missing Science of Consciousness* (Oxford: Oxford University Press, 1994); and A. C. Elitzur and L. Vaidman, "Quantum-Mechanical Interaction-Free Measurement," *Foundations of Physics* 23, no. 7 (1993), 987–97. On using interaction-free measurements with a test particle to determine the presence of an object, see G. Kwiat et al., "Interaction Free Measurements," *Physics Review Letters* 74, no. 12 (June 12, 1995), 4763–66. And on counterfactual computation, see Graeme Mitchison and Richard Josza, "Counterfactual Computation," *Quantum-Ph/9907007* (July 2, 1999).

[3] Electoral studies are a case in point. See Raymond E. Wolfinger and Benjamin Highton, "Can More Efficient Purging Boost Turnout" (Paper presented at the annual meeting of the American

our profession counterfactual arguments appear to have no scientific standing. They are flights of fancy, fun over a beer or two in the faculty club, but not the stuff of serious research.[4]

I begin my essay with the proposition that the difference between so-called factual and counterfactual arguments is greatly exaggerated; it is one of degree, not of kind. I go on to discuss three generic uses of counterfactual arguments and thought experiments. In the process, I distinguish between "miracle" and "plausible" world counterfactuals and identify the uses to which each is suited. I critique two recent historical works that make extensive use of counterfactuals and contend that they are seriously deficient in method and argument. I then review the criteria for counterfactual experimentation proposed by social scientists who have addressed this problem and find many of their criteria unrealistic and overly restrictive. The methods of counterfactual experimentation need to be commensurate with the purposes for which they are used, and I conclude by proposing eight criteria I believe appropriate to plausible-world counterfactuals.

Counterfactuals versus Factuals

Counterfactuals are "what if" statements, usually about the past. Counterfactual experiments vary attributes of context or the presence or value of variables and analyze how these changes would have affected outcomes. In history and political science these outcomes are always uncertain because we can neither predict the future nor rerun the tape of history.

The speculative nature of counterfactuals makes many scholars wary of them. But counterfactual analysis is not always based solely on speculation. In the aftermath of the arrest of Aldrich Ames as a Soviet spy, the Central Intelligence Agency convened a team of counterintelli-

Political Science Association, New York, 1994); Ruy A. Teixeira, *The Disappearing American Voter* (Washington, D.C.: Brookings Institution, 1992); Steven J. Rosenstone and John Mark Hansen, *Mobilization, Participation, and Democracy in America* (New York: Macmillan, 1993); Richard J. Timpone, "Past History Reverses: Counterfactuals and the Impact of Registration Reform on Participation and Representation" (Manuscript, 1999).

[4] According to A. J. P. Taylor, "A historian should never deal in speculation about what did not happen." See Taylor, *The Struggle for the Mastery in Europe, 1848–1918* (London: Oxford University Press, 1954); E. H. Carr, *What Is History*? (London: Macmillan, 1961), 127; E. P. Thompson, *The Poverty of Theory and Other Essays* (New York: Monthly Review Press, 1978), 300. M. M. Postan writes: "The might-have beens of history are not a profitable subject of discussion"; quoted in J. D. Gould, "Hypothetical History, " *Economic History Review*, 2d ser., 22 (August 1969), 195–207. See also David Hackett Fischer, *Historians' Fallacies* (New York: Harper Colophon Books, 1970), 15–21; Peter McClelland, *Causal Explanation and Model-Building in History, Economics, and the New Economic History* (Ithaca, N.Y.: Cornell University Press, 1975).

gence experts to figure out how he might have been unmasked earlier. The team imagined a series of procedures that might have been in place and asked which, if any of them, might have tripped up Ames. Using their extensive knowledge of Ames's personality, motives, and behavior and of the modus operandi of his Soviet spymasters, they were able to conduct their inquiry with some degree of precision.[5] And in New York City authorities only recently prepared for the possibility of widespread Y2K computer failures by running tests of the ability of city agencies to respond to a complex emergency. Officials were confident that such breakdowns would not occur, but they used the opportunity to consider the scenarios as "future counterfactuals" to gather useful data.[6] The controversy surrounding the strategy of deterrence provides an example of the use of counterfactuals in international relations. One of the principal policy lessons of the 1930s was that appeasement whets the appetites of dictators while military capability and resolve restrains them. The failure of Anglo-French efforts to appease Hitler is well established, but the putative efficacy of deterrence rests on the counterfactual that Hitler *could* have been restrained *if* France and Britain had demonstrated willingness to go to war in defense of the European territorial status quo. German documents make this an eminently researchable question, and historians have used these documents to try to determine at what point Hitler could no longer be deterred.[7] Their findings have important implications for the historical assessment of French and British policy and for the strategy of deterrence.

The Cuban missile crisis is an evidence-rich environment in which counterfactuals drove policy and have since been subjected to testing. Nikita Khrushchev's decision to send missiles to Cuba and then to remove them and Kennedy's decision to impose a blockade were contingent upon hypothetical antecedents. Kennedy believed—incorrectly—that Khrushchev sent missiles to Cuba because he doubted the president's resolve and that he would not have done so had Kennedy taken a stronger stand at the Bay of Pigs or in Berlin. Kennedy reasoned that he had to compel Khrushchev to remove the missiles to convince him of his resolve and to deter a subsequent and more serious Soviet challenge to the Western position in Berlin. Evidence from Soviet and American archives and interviews with former officials make

[5] Private communication to the author.

[6] *New York Times*, November 27, 1999, A1.

[7] Yuen Foong Khong, "Confronting Hitler and Its Consequences," in Philip E. Tetlock and Aaron Belkin, eds., *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives* (Princeton: Princeton University Press, 1996).

it possible to explore the validity of most of these counterfactuals and thus to evaluate the choices of Soviet and American leaders and the subsequent scholarly analyses of the crisis.[8]

Quantitative counterfactual analysis is another possibility. Jay Winter exploits counterfactual projections of mortality rates based on prewar data from Prudential Life Insurance policies to determine the age structure of British war losses. He combines data from the life tables for 1913 and 1915 in roughly two to one proportions, as the war did not begin until August 1914, to create a counterfactual table for 1914. On the basis of the prior decade, he then calculates what the life tables would have been for the period 1914–18 in the absence of war. By comparing the actual death rates in each age group with the counterfactual estimates, he is able to determine the death rates of five-year cohorts for each year of the war.[9]

Even when evidence is meager or absent, the difference between counterfactual and "factual" history may still be marginal. Documents are rarely smoking guns that allow researchers to establish motives or causes beyond a reasonable doubt. Actors only occasionally leave evidence about their motives, and historians rarely accept such testimony at face value. More often historians infer motives from what they know about actors' personalities and goals, their past behavior, and the constraints under which they operated. In his highly acclaimed study of the Peloponnesian War, Donald Kagan argues that Pericles wanted to ally with Corcyra in the expectation that it would deter Sparta from coming to the aid of Corinth. If deterrence failed, Athens, protected by its city walls and the long walls to its harbor at Piraeus, would refuse to engage the main body of Spartan forces even if they invaded Attica and laid waste its olive groves and vineyards. After the Spartans experienced a few years of frustration, Pericles expected them to recognize the futility of waging war against Athens. Further, he thought that the peace faction, led by King Archidamus, would regain power and that the two hegemons would reach a more lasting accommodation.[10] The purpose of this purely speculative scenario is to explain away behavior that otherwise appears unenlightened and warlike.

The problem of motive is not unique to ancient history, where sources are notoriously meager. Janice Gross Stein and I spent several

[8] Richard Ned Lebow and Janice Gross Stein, "Back to the Past: Counterfactuals and the Cuban Missile Crisis," in Tetlock and Belkin (fn. 7).

[9] Winter, *The Great War and the British People* (London: Macmillan, 1986), 76–83.

[10] Kagan, *The Outbreak of the Peloponnesian War* (Ithaca, N.Y.: Cornell University Press, 1969), 203–342.

years researching a book on cold war crises. We scoured archives in four countries, utilized documents collected or declassified by other researchers, and conducted extensive interviews with former American, Soviet, Israeli, and Egyptian policymakers. We accumulated a mass of relevant information but still had no hard evidence about the motives for some of the key decisions made by Kennedy and Khrushchev. We suspect that Khrushchev was never clear in his own mind about the relative importance of the several goals that made a missile deployment in Cuba attractive to him. Given the delicate nature of many crisis decisions, neither leader was willing to share his goals and reasoning with even his most intimate advisers. Khrushchev further complicated the picture by telling various officials what they wanted to hear and thus what he thought was most likely to garner their support.[11]

When we move from the level of analysis of individual actors to small groups, elites, societies, states, and regional and international systems, the balance between evidence and inference shifts decisively in the direction of the latter. Structural arguments assume that behavior is a response to the constraints and opportunities generated by a set of domestic or international conditions. Mark Elvin's elegant study of China starts from the premise that empires expand to the point at which their technological superiority over their neighbors is approximately counterbalanced by the burdens of size. At this equilibrium, imperial social institutions come under constant strain because of the high relative cost of security. Harsh taxation impoverishes peasant cultivators and leads to falling tax revenues. The ensuing decline in the number of free subjects makes military recruitment more difficult, forcing governments to rely instead on barbarian auxiliaries, even for their main fighting forces. To save money, governments also give up active defense policies and try to keep hostile barbarians at bay through diplomacy, bribery, and settlement on imperial lands. The inevitable outcome is a weakened economic base, barbarization from within, and finally partial or total collapse of the empire.[12] Elvin musters considerable evidence in support of his thesis, much of it from primary sources, but it is all by way of illustration. Nowhere is he able to show that Chinese leaders adopted any of the policies he describes for any of the reasons he attributes to them.

For the most part, then, structural arguments are built on a chain of inference that uses behavioral "principles" as anchor points. Empirical evidence, when available, may be exploited to suggest links between

[11] Lebow and Stein, *We All Lost the Cold War* (Princeton: Princeton University Press, 1994).

[12] Elvin, *The Pattern of the Chinese Past: A Social and Economic Interpretation* (Stanford, Calif.: Stanford University Press, 1973).

these principles and behavior. But even in the best of cases these links are indirect and presumptive and can be corroborated only obliquely and incompletely. Readers evaluate these arguments on the seeming "reasonableness" of the inferences drawn, the quality and relevance of the evidence offered in support, and the extent to which that evidence permits or constrains alternative interpretations. Receptivity to arguments is significantly influenced by the appeal of the underlying political and behavioral "principles" in which the inferences are rooted. When these "principles" run counter to the reigning orthodoxy, the arguments may be dismissed out of hand regardless of the evidence.

Good counterfactual thought experiments differ little from "factual" modes of historical reconstruction. If we attempt to evaluate the importance of Mikhail Gorbachev for the end of the cold war by considering the likely consequences of Chernenko being succeeded by someone else, we need to study the career and policies of other possible successors (for example, Grishin, Romanov, Ligachev) and infer their policies on the basis of their past preferences and commitments, the political environment in 1985, and the general domestic and foreign situation of the Soviet Union.[13] There is a lot of evidence about all three questions, evidence that sustains informed arguments about the kind of domestic and foreign policies these leaders might have pursued. Admittedly, unexpected events, like Mathias Rust's Cessna flight to Red Square in May 1987, which Gorbachev exploited to purge the military of many hard-liners, can also have significant influence on policy.

The difference between factual and counterfactual arguments is further blurred when we recognize that, as in the Cuban missile crisis, we often need to understand the factual *and* counterfactual beliefs of historical actors to account for their behavior. In the missile crisis beliefs shaped arguments: in the absence of compelling evidence, the beliefs of officials determined the motives they attributed to Khrushchev for deploying Soviet missiles in Cuba, their estimates of the cost calculations and political conflicts they assumed to being taking place in Moscow, and the likely Soviet responses to a blockade, air strike, or invasion. Some of these beliefs took the form of conditional expectations, and with the passage of time they became historical counterfactuals.

[13] See George Breslauer and Richard Ned Lebow, "Leadership and the End of the Cold War: A Counterfactual Thought Experiment," in Richard Herrmann and Richard Ned Lebow, eds., *Learning from the Cold War* (forthcoming). The authors identify leaders other than Gorbachev, Reagan, and Bush who might have come to power in the Soviet Union and the United States and play out the resulting interaction. They find that different strategies or tactics by either superpower or their allies could have speeded up, slowed down, altered, or derailed the process of accommodation that led to the end of the cold war in 1990–91.

Counterfactuals are frequently smuggled into so-called factual narratives. E. H. Carr, no friend of counterfactuals, did this in his treatment of the Soviet Union when he insisted that the Bolshevik Revolution was highjacked by Stalin. The implication is that socialism would have developed differently without him.[14] After Cuba, former Kennedy administration officials and many scholars maintained that Khrushchev would not have deployed missiles in Cuba if Kennedy had been more decisive at the Bay of Pigs, at the Vienna summit, and in Berlin. There was no evidence to support this interpretation, but it became the conventional wisdom and helped to shape a host of subsequent policy decisions, including the disastrous intervention in Vietnam. The evidence that came to light in the Gorbachev era suggested, to the contrary, that Khrushchev decided to send missiles secretly to Cuba because he *overestimated* Kennedy's resolve. He feared that Kennedy, preparing to invade Cuba, would send the American navy to stop any ships carrying missiles to Cuba to deter that invasion.[15] Counterfactual arguments, like any historical argument, are only as compelling as the logic and "evidence" offered by the researcher to substantiate the links between the hypothesized antecedent and its expected consequences. Every good counterfactual thus rests on multiple factuals, just as every factual rests on counterfactual assumptions—and these assumptions too often go unexamined.

Any sharp distinction between factuals and counterfactuals rests on questionable ontological claims. Many of the scholars who dismiss counterfactual arguments do so because they do not believe they are based on facts. Philosophers have long recognized that "facts" are social constructions. They do not deny that reality exists quite independent of any attempt to understand it by human beings or that some understandings may transcend culture. Physical scientists may be correct in their claims that fundamental concepts like mass, volume, and temperature are essential to the study of nature and that extraterrestrial scientists would have to possess the same concepts to understand the universe. This is not true of social concepts, which vary across and within human cultures. There are many ways of describing social interactions, and the choice and utility of concepts depend largely on the purpose of the "knower."[16]

"Temperature" is undeniably a social construction, but it is a measure of something observable and real: changes in the energy levels of mol-

[14] Carr, *Socialism in One Country, 1924–1926* (New York: Macmillan, 1958–1964), 1:151.
[15] Lebow and Stein (fn. 11) review this argument and the new evidence (chap. 4).
[16] John R. Searle, *The Construction of Social Reality* (New York: Free Press, 1995).

ecules. Social and political concepts do not describe anything so concrete. There is no such thing as a balance of power, a social class, or a tolerant society. Social "facts" are reflections of the concepts we use to describe social reality, not of reality itself. They are ideational and subjective, and even the existence of "precise" measures for them—something we have only rarely—would not make them any less arbitrary. For, as Quine has shown, theoretical concepts insinuate themselves into the "data language" of even the hardest sciences.[17] The construction of "factual" history is therefore entirely imaginary, and its only claim to privilege is that the concepts and categories in terms of which it is constructed tell us something useful or interesting about the social world. The same is true for counterfactual history.

Counterfactuals can be used experimentally to substantiate Quine's claim that there is no conceptually neutral data language. Tetlock and Lebow asked a group of foreign policy experts to assess the contingency of the outcome of the Cuban missile crisis. One of their experiments used a "factual framing" of the question (at what point did some form of peaceful resolution of the Cuban missile crisis become inevitable?) and a counterfactual framing (at what point did all alternative, more violent outcomes become impossible?). From a logical point of view, the two questions are strictly complementary. Knowing the answer to either question, we should be able to deduce the answer to the other. Even though these two measures were obtained almost side by side in our questionnaire, the factual versus counterfactual framings of the historical question elicited systematically different responses, not just random variation that could be attributed to fatigue or boredom. Experts perceived substantially more contingency when they reflected on the counterfactually framed question. This is a good empirical demonstration of the importance of the benchmark against which the outcome is compared and offers support for the constructivist claim that how we pose "purely empirical" questions systematically shapes the answers we find.[18]

Why Counterfactuals?

Counterfactuals can combat the deeply rooted human propensity to see the future as more contingent than the past, reveal contradictions in our

[17] Donald Davidson and Jaakko Hintikka, eds., *Words and Objection: Essays on the Work of W. V. Quine* (Dordrecht: D. Reidel, 1969).

[18] Philip E. Tetlock, "Close-Call Counterfactuals and Belief System Defense: I Was Not Almost Wrong but I Was Almost Right," *Journal of Personality and Social Psychology* 75 (September 1998), 639–52.

belief systems, and highlight double standards in our moral judgments. Counterfactuals are an essential ingredient of scholarship. They help determine the research questions we deem important and the answers we find to them. They are also necessary to evaluate the political, economic, and moral benefits of real-world outcomes. These evaluations in turn help drive future research.

Receptivity to Contingency

International relations theorists seek to understand the driving forces behind events; they usually do so after the fact, when the outcome is known. The process of backward reasoning tends to privilege theories that rely on a few key variables to account for the forces allegedly responsible for the outcomes in question. For the sake of theoretical parsimony, the discipline generally favors independent variables that are structural in nature (for example, balance of power, state structure, size and nature of a coalition). The theory-building endeavor has a strong bias toward deterministic explanations and on the whole downplays understandings of outcomes as the products of complex, conjunctional causality.[19] A recent survey of international relations specialists revealed that those scholars who were most inclined to accept the validity of theories (for example, power transition, nuclear deterrence) and theory building as a scholarly goal were the most emphatically dismissive of plausible-world counterfactuals. They were also most likely to invoke second-order counterfactuals to get developments diverted by counterfactuals back on the track.[20]

In retrospect, almost any outcome can be squared with any theory unless the theory is rigorously specified. The latter requirement is rarely met in the field of international relations, and its deleterious effect is readily observed in the ongoing debate over the end of the cold war. Various scholars, none of whose theories predicted a peaceful end to that conflict, now assert that this was a nearly inevitable corollary of their respective theories.[21] We observe a similar phenomenon in studies of Middle Eastern politics. Developments that seemed almost unthinkable before they happened—Sadat's trip to Jerusalem, the Pales-

[19] On conjunctional causality, see Charles C. Ragin, *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies* (Berkeley: University of California Press, 1987).

[20] Tetlock and Lebow, "Poking Counterfactual Holes in Covering Laws: Alternative Histories of the Cuban Missile Crisis" (Manuscript, Mershon Center, Ohio State University, April 2000); on second-order counterfactuals, see Tetlock (fn. 18).

[21] William Wohlforth reviews this literature and criticizes realists for explaining ex post facto what none of them had predicted ex ante; see Wohlforth, "New Evidence on Moscow's Cold War: Ambiguity in Search of Theory," *Diplomatic History* 21 (Spring 1997).

tinian-Israeli moves toward peace—are subsequently described as having been overdetermined by structural causes, particularly shifts in the relative balance of power.[22]

The disciplinary tendency to privilege structural explanations is reinforced by the "certainty of hindsight bias."[23] Baruch Fischoff has demonstrated that "outcome knowledge" affects our understanding of the past by making it difficult for us to recall that we were once unsure about what was going to happen. Events deemed improbable by experts (for example, peace between Egypt and Israel, the end of the cold war) are often considered "overdetermined" and all but inevitable after they have occurred.[24] By tracing the path that appears to have led to a known outcome, we diminish our sensitivity to alternative paths and outcomes. We may fail to recognize the uncertainty under which actors operated and the possibility that they could have made different choices that might have led to different outcomes.

Many psychologists regard the certainty-of-hindsight effect as deeply rooted and difficult to overcome. But the experimental literature suggests that counterfactual intervention can assist people in retrieving and making explicit their massive but largely latent uncertainty about historical junctures, that is, to recognize that they once thought, perhaps correctly, that events could easily have taken a different turn. The proposed correctives use one cognitive bias to reduce the effect of another. Ross et al. exploited the tendency of people to inflate the perceived likelihood of vivid scenarios to make them more responsive to contingency. People who were presented with scenarios describing possible life histories of post-therapy patients evaluated these possibilities as more likely than did members of the control group who were not given the scenarios. This effect persisted even when all the participants in the experiment were told that the post-therapy scenarios were entirely hypothetical.[25] Philip E. Tetlock and the author conducted a

[22] Steven Weber, "Prediction and the Middle East Peace Process," *Security Studies* 6 (Summer 1997), 196.

[23] This point is also made by Philip E. Tetlock and Aaron Belkin, "Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives," in Tetlock and Belkin (fn. 7), 15–16.

[24] Baruch Fischoff, "Hindsight Is Not Equal to Foresight: The Effect of Outcome Knowledge on Judgment under Uncertainty," *Journal of Experimental Psychology: Human Perception and Performance* 1, no. 2 (1975), 288–99; S. A. Hawkins and R. Hastie, "Hindsight: Biased Judgments of Past Events after the Outcomes Are Known," *Psychological Bulletin* 107, no. 3 (1990), 311–27. The tendency was earlier referred to as "retrospective determinism" in comparative-historical studies by Reinhard Bendix, *Nation-Building and Citizenship* (New York: Wiley, 1964).

[25] L. Ross et al., "Social Explanation and Social Expectation: Effects of Real and Hypothetical Explanations on Subjective Likelihood," *Journal of Personality and Social Psychology* 35 (November 1977), 817–29.

series of experiments to test the extent to which counterfactual unpacking leads foreign policy experts to upgrade the contingency of international crises. In the first experiment one group of experts was asked to assess the inevitability of the Cuban missile crisis. A second group was asked the same questions but was given three junctures at which the course of the crisis might have taken a different turn. A third group was given the same three junctures and, in addition, three arguments for why each of them was plausible. Judgments of contingency varied in proportion to the degree of counterfactual unpacking.[26] There is every reason to expect that scholars exposed to counterfactuals and, better yet, forced to grapple with their theoretical consequences will also become more open to the role of contingency in key decisions and events.

Framing Research

Research questions arise when events strike us as interesting or anomalous. To conceive of an event as anomalous we need a benchmark against which the outcome in question can be compared. Benchmarks can sometimes be derived from well-established laws or statistical generalizations: cold fusion, for example, would have been contrary to several of these laws and thus a truly anomalous event. There are few laws or statistical generalizations applicable to politics, but we hold to theories about how the political world works. They give rise to expectations, and when they are unfulfilled, to counterfactual worlds. These alternative worlds may appear more probable than the actual state of affairs. During the cold war the preeminent question in the security field was the long peace between the superpowers. In international political economy it was the survival of the postwar international economic order in the face of America's decline as a hegemon. Some security specialists considered it remarkable that the superpowers, unlike rival hegemons of the past, had avoided war. Some political economists were equally surprised that neither Germany nor Japan had sought to restructure international economic relations to their advantage in response to the apparent decline of the United States as hegemon.[27] Both

[26] The first of these experiments, involving alternative outcomes for the Cuban missile crisis, is described in Tetlock and Lebow (fn. 20).

[27] Kenneth N. Waltz, *Theory of International Politics* (Reading, Mass.: Addison-Wesley, 1979); John Lewis Gaddis, *The Long Peace: Inquiries into the History of the Cold War Era* (New York: Oxford University Press, 1987); Robert Gilpin, *War and Change in World Politics* (New York: Cambridge University Press, 1981), 232–33; Robert O. Keohane, *After Hegemony* (Princeton: Princeton University Press, 1984); Duncan Snidal, "The Limits of Hegemonic Stability Theory," *International Organization* 39 (Autumn 1985).

research agendas assumed that the status quo was an extraordinary anomaly that required an equally extraordinary explanation. For researchers who started from different premises—who assumed that none of the major powers were so unhappy with the current state of affairs that they were willing to risk war or economic disruption to change it—the seeming robustness of the political and economic orders posed no intellectual puzzle.

Testing and Evaluation

Counterfactuals are fundamental to all theories and interpretations. If we hypothesize that x caused y, we assume that y would not have happened in the absence of x—ceteris paribus. Quantitative research attempts to get around this problem and around the contrapositive form of the fallacy of affirmation by constructing a sample of comparable cases large enough to contain adequate variation on dependent and independent variables.[28] James Fearon rightly observes that this strategy is effective only if there are no causes beyond those considered that vary systematically with the error term. To rule out this possibility, researchers need to pose the counterfactual of what would have happened if variables in the error term were altered. In actual experiments, this problem can only partially be solved by random assignment.[29] In case studies and historical narratives the problem is more pronounced because of the usual "loading up of explanatory variables."

Historians and case-study researchers typically attempt to establish causation by process tracing. They try to document the links between a stated cause and a given outcome in lieu of establishing a statistical correlation. This works best at the individual level of analysis but only when there is enough evidence to document the calculations and motives of actors. Even when such evidence is available, it may still not be possible to determine the relative weight of the several hypothesized causes and which of them, if any, might have produced the outcome in the absence of others or in combination with other causes not at work in the case. To sustain causal inference it is generally necessary to engage in comparative analysis. Within the single-case format comparative analysis can take two forms: intracase comparison and counterfactual analysis.

[28] Robyn M. Dawes, "Counterfactual Inferences as Instances of Statistical Inferences," in Tetlock and Belkin (fn. 7), 301–8. Strictly speaking it does not follow that if x then y; therefore, if not-x, then not-y, because factors other than x may also cause y.

[29] Fearon, "Causes and Counterfactuals in Social Science: Exploring an Analogy between Cellular Automata and Historical Processes," in Tetlock and Belkin (fn. 7).

Intracase comparison breaks down a case into a series of similar interactions that are treated as separate and independent cases for purposes of analysis. Numerous studies of arms control and superpower crises have made use of this technique.[30] Like any form of comparative analysis, intracase comparisons try to show as much variation as possible on dependent (the *explanandum*) and independent (the *explanans*) variables. This is sometimes more difficult to do than in cross-case comparisons. The independence of cases is also more problematic, as the process and outcome of past decisions are likely to have considerable influence on subsequent decisions about similar issues. But intracase comparison confers a singular benefit: it builds variation within a fundamentally similar political and cultural context, controlling better than intercase comparison for many factors that may be important but are otherwise unrecognized. Unfortunately, not every case can be broken down into multiple decision points for purposes of comparison.

When intracase comparison is impossible, variation can be generated within a case by counterfactual experimentation. This latter strategy lies at the core of many simulations where variables are given a wide range of counterfactual values to determine the sensitivity of the outcome to changes in one or more of them. Counterfactual simulation can identity key variables and the range of values at which they will be have the most impact on the outcome. Information obtained this way, especially if it has counterintuitive implications, can guide subsequent empirical work intended to test the model or generate information necessary to make it a better representation of reality. Counterfactual simulation can test theories more directly. Thomas Cusack and Richard Stoll reviewed the realist literature to identify major shared assumptions among and the principal differences between contending interpretations. They formalized a model that would allow a computer-based simulation of some of these variants and focused their analysis on two issues of paramount importance to realists: state survival and system endurance. They found that the behavior realists prescribe as conducive to these ends (for example, maintaining large military forces, seeking to increase

[30] See, for example, Alexander L. George and Richard Smoke, *Deterrence in American Foreign Policy: Theory and Practice* (New York: Columbia University Press, 1974); Alexander L. George and William E. Simmons, eds., *The Limits of Coercive Diplomacy* (Boulder, Colo.: Westview, 1994); Ted Hopf, *Peripheral Visions: Deterrence Theory and American Foreign Policy, 1965–1990* (Ann Arbor: University of Michigan Press, 1994); Robert Jervis and Jack Snyder, eds., *Dominoes and Bandwagons: Strategic Beliefs and Great Power Competition in the Eurasian Rimland* (New York: Oxford University Press, 1991); Thomas Risse-Kappen, *Cooperation among Democracies: The European Influence on U.S. Foreign Policy* (Princeton: Princeton University Press, 1995).

one's power though alliances, especially those aimed at balancing) did not necessarily enhance survival of individual units or the system as a whole.[31] Lars-Eric Cederman developed a more focused critique of realism. If balancing is the inevitable consequence of international anarchy, as neorealism contends, then it follows, he reasoned, that global hegemons should rarely emerge in real or counterfactual worlds. But Cederman found that they appear with regularity in counterfactual simulations, especially under conditions of defense dominance, the best case for neorealism.[32]

Counterfactual experiments and simulations can tease out the assumptions—often unarticulated—on which theories and historical interpretations rest.[33] Apologists for the Soviet system insist that communism would have evolved differently had Lenin lived longer or been succeeded by someone other than Stalin.[34] Attempts to address this question have not resolved the controversy but have compelled historians to be more explicit about the underlying assumptions that guide and sustain contending interpretations of Stalin and the nature of the Communist Party and the Soviet state. Those assumptions have now become the focus of controversy, and scholars have looked for evidence with which to evaluate them. This process has encouraged a more sophisticated historical debate.

Because every causal argument has its associated counterfactual, critics have available to them two generic strategies. They can offer a different and more compelling theory or interpretation—far and away the most common strategy—or they can show that the outcome in question would have happened in the absence of the hypothesized causes. John Mueller's study of the cold war is a nice example of the second strategy. In contrast to the conventional wisdom that attributes the long peace between the superpowers to nuclear deterrence, Mueller argues that Moscow and Washington were restrained by their general satisfaction with the status quo, by memories of World War II, and by the human, economic, and social costs of large-scale conventional warfare. He contends that the unheralded destructiveness of nuclear weapons

[31] Thomas R. Cusack and Richard J. Stoll, *Exploring Realpolitik: Probing International Relations Theory with Computer Simulation* (Boulder, Colo.: Lynne Rienner, 1990). Their evaluation was based on earlier work by Stuart A. Bremer and Michale Mihalka, "Machiavelli in Machina: Or Politics among Hexagons," in Karl W. Deutsch, ed., *Problems of Modeling* (Boston: Ballinger, 1977).

[32] Cederman, "Rerunning History: Counterfactual Simulation in World Politics," in Tetlock and Belkin (fn. 7).

[33] Fearon (fn. 29).

[34] George W. Breslauer discusses this literature in Breslauer, "Counterfactual Reasoning in Western Studies of Soviet Politics and Foreign Relations," in Tetlock and Belkin (fn. 7).

was redundant and possibly counterproductive.[35] This Mueller strategy readily lends itself to simulation. Researchers can build a counterfactual world and look at how actors behave under a wide range of conditions, including those that subtract putative causal factors.

Assessing Outcomes

Counterfactuals are a key component of evaluation. Was the development of nuclear weapons a blessing or a curse for humankind? What about affirmative action, free trade, or the growing economic and political integration of Europe? Serious and thoughtful scholars can be found on all sides of these controversies. Their arguments share one thing in common: the use of counterfactual benchmarks—most often, implicitly—to assess the merits of real-world policies, outcomes, or trends. Proponents of nuclear weapons who claim that nuclear weapons had beneficial consequences during the cold war imagine a superpower war, or at least a higher probability of one, in the absence of nuclear deterrence. Some critics of nuclear weapons, like John Mueller, argue that self-deterrence based on memories of the horrors of conventional war would have kept the peace. Other critics contend that nuclear weapons sustained the cold war and that it would have been less intense and possibly resolved earlier in their absence.

Assessment can be significantly influenced, or even determined, by the choice of counterfactual. The conventional wisdom holds that the allied victory in World War I was a good thing: it prevented an expansionist, continental power from achieving hegemony in continental Europe. This assessment represents the view of the world from the corporate boardrooms and corridors of power in London, New York, and Washington. From the perspective of, say, Polish Jewry, the outcome was a disaster. If Germany had won, there almost certainly would have been no Hitler and no Holocaust. In this case, the choice of counterfactual reflects the different interests of the various groups. As with historical analogies, the interesting and eminently researchable question becomes the extent to which counterfactuals guide evaluation or are chosen to justify positions that people have reached for quite different reasons.[36]

[35] John Mueller, *Retreat from Doomsday: The Obsolescence of Major War* (New York: Basic Books, 1989). And see the debate on this subject between Mueller, "The Essential Irrelevance of Nuclear Weapons: Stability in the Postwar World," and Robert Jervis, "The Political Effects of Nuclear Weapons: A Comment," both in *International Security* 13 (Fall 1988).

[36] For a strong statement of the former position with regard to historical analogies, see Yuen Foong Khong, *Analogies at War: Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965* (Princeton: Princeton University Press, 1992).

Must Counterfactuals Be Realistic?

The several uses of counterfactuals I have described use "plausible" and "miracle" world counterfactuals. Plausible-world counterfactuals are intended to impress readers as realistic; they cannot violate their understanding of what was technologically, culturally, temporally, or otherwise possible. In a recent study of the origins of World War I, I imagine a world in which Archduke Franz Ferdinand and his wife, Countess Sophie, returned alive from their visit to Sarajevo.[37] This counterfactual strikes me as eminently plausible because their assassination was such a near thing. Princip's accomplice missed the royals en route to city hall, and Princip was lamenting his failure when the touring car carrying Franz Ferdinand and his wife came to a stop in front of the bar from which Princip had just emerged to allow the cars at the head of the procession to back up because they had made a wrong turn. Princip stepped forward and fired two shots at point-blank range into their touring car. With only a minimal rewrite of history—the procession is halted after the first attempt or subsequently stays on the planned route—the assassination could easily have been averted.

There are many plausible counterfactuals—historical near misses, if you like—that might have come to pass but probably would not have had any significance for the outcomes in question. Plausible counterfactuals must meet a second test: they must have a real probability of leading to the outcome the researcher intends to bring about. To demonstrate this, the researcher must construct a logical path between the counterfactual change and the hypothesized outcome, and meet other tests that are described in the last section of the review. Plausible-world counterfactuals are thought by some researchers to be the only legitimate kind of counterfactual.[38]

Miracle counterfactuals violate our understanding of what is plausible or even possible.[39] Take the following assertion: "If Bosnians had been blue-nosed dolphins, NATO would not have allowed their slaughter."[40] A landlocked province of aquatic mammals with ethnic identities would be a stretch even for Dr. Seuss—but it is morally provocative. It implies that NATO's belated intervention was somehow related to the

[37] Richard Ned Lebow, "Contingency, Catalysts and System Change," *Political Science Quarterly* (forthcoming).

[38] Jon Elster, *Logic and Society: Contradictions and Possible Worlds* (New York: John Wiley, 1978); G. Hawthorn, *Plausible Worlds: Possibility and Understanding in History and the Social Sciences* (New York: Cambridge University Press, 1991), 31–60.

[39] The term "miracle counterfactual" was coined by Fearon (fn. 29), 60.

[40] Tetlock and Belkin (fn. 23), 14.

Muslim affiliation of much of Bosnia's population and would have come sooner if Bosnians had characteristics that evoked more sympathy in the West. Miracle counterfactuals are useful for purposes of theory building and testing. I could hypothesize that Europe achieved its military advantage because it was the only region of the world where no long-standing hegemony was established and because the resulting prolonged competition among its leading political units made them lean and mean, better armed, and more efficient in the use of large-scale violence.[41] To advance this hypothesis I must consider the counterfactual of a hegemonic Europe—perhaps achieved by a better organized and more astutely led Spain in the sixteenth century. To sustain the hypothesis I need to argue what a hegemonic Europe would have been like, and how it would have differed from the historical Europe. Miracle counterfactuals are particularly useful in evaluating existing interpretations.

Mueller's world without nuclear weapons is a miracle counterfactual because it would require a massive rewrite of a century of scientific and political history to uninvent nuclear weapons, although the timing of their development might be altered by plausible-world counterfactuals. The value of miracle counterfactuals derives not from their realism but from the analytical utility of considering alternative worlds. Counterfactuals of this kind have a distinguished lineage. Euclid used one to prove that there are an infinite number of prime numbers, and Newton, to demonstrate that the universe could not be infinite with regularly distributed and fixed stars. If these conditions held, Newton argued, the sky would not be blue.[42]

Miracle counterfactuals are critical to assessment. To evaluate the relative merits of court-contested versus mediated divorces, we need to know something about the financial consequences of each mode of settlement for women. To do this, we can compare outcomes in states that encourage and discourage mediation in comparable samples of divorced couples.[43] But we confront a serious data problem if we are interested in ascertaining whether making divorce more difficult to

[41] Ludwig Dehio, *The Precarious Balance: Four Centuries of the European Power Struggle* (New York: Knopf, 1995). Dehio argues that competition among many independent units produced "fertile friction" among Greek city-states and in modern Europe.

[42] Newton reasoned that the energy reaching us from an individual star is E/r^2, where r is the distance of the star, and E is the average energy radiated by each star, and if the density (d) of stars in the universe is constant, the number of stars would be d r^3. The total energy produced by these stars would grow in a linear fashion with r, and rise to infinity in an infinite universe—neither we nor the earth would exist. Hence, the density of stars must decrease or the universe must be finite.

[43] Carol Bohmer and Marilyn Ray, "Effects of Different Dispute Resolution Mechanisms on Women and Children after Divorce," *Family Law Quarterly* 28, no. 2 (1994).

obtain would hold families together, as many conservatives allege. Legislation to make divorce incrementally more difficult has been debated in several states but rejected in all of them. Really tough divorce laws nationwide are unrealistic in the current social climate, and the abolition of divorce is even more improbable.[44] But miracle counterfactuals that use either premise as the starting point for research might generate interesting results that could raise the level of the policy debate. This is not dissimilar to the kind of research economists do all the time; they raise or lower prices of commodities well beyond any realistic market expectations to test consumer preferences.

Historical Counterfactuals

Recent publications have sparked renewed interest in counterfactuals. In political science the principal catalysts have been Fearon's 1991 article in this journal and Tetlock and Belkin's edited volume, *Counterfactual Experiments in World Politics*.[45] The earlier publication of Mueller's *Retreat from Doomsday* also drew a lot of attention, but it was focused more on his controversial conclusion than on his counterfactual method.[46] The two works under review have revived interest in counterfactuals in the discipline of history.

Virtual History does not make a good case for counterfactuals. Ferguson's long introduction is a literature review without much purpose. It offers no reasons for engaging in counterfactual history other than sensitizing readers to contingency, and it only briefly addresses the methods by which counterfactual experiments should be conducted. Ferguson criticizes earlier counterfactual works for inferring momentous consequences from "simple, often trivial change[s]." With undisguised scorn, he cites as an example Pascal's intentionally provocative counterfactual about Cleopatra's nose: if it had been ugly, Anthony would not have fallen for her, and the history of Rome might have been different (pp. 11–12). But what is wrong with small changes having big effects? Could anyone seriously doubt that the course of history would have been different if Pharaoh's daughter had not found a child in a basket in the reeds, if the Mongol fleet had not encountered a destructive typhoon en route to Japan, if the duke of Alba had not fallen sick

[44] Deborah Rhode, *Speaking of Sex: The Denial of Gender Inequality* (Cambridge: Harvard University Press, 1997), 184.

[45] James D. Fearon, "Counterfactuals and Hypothesis Testing in Political Science," *World Politics* 43 (January 1991); Tetlock and Belkin (fn. 7). I do not review the latter volume because I am coauthor of one of its chapters.

[46] See fn. 35 for Mueller and the follow-up debate on nuclear weapons between Mueller and Jervis.

in 1572, or if Hitler had died in trenches during World War I or in the near fatal traffic accident he suffered in the summer of 1930? When the duke of Alba took to his bed, his inexperienced and arrogant son took command of the forces laying siege to Haarlem, rejected the town's offer of surrender on terms, and prolonged the Dutch rebellion against Spanish rule. The sustained ability of the Dutch to resist infuriated Philip II and his nephew, Alexander Farnese, duke of Parma. They convinced themselves that the Dutch resisted only because of English support and decided to deal with England directly. Geoffrey Parker suggests that the subsequent defeat of the Spanish Armada laid the American continent open to invasion and colonization by Northern Europeans and made possible the founding of the United States.[47] For want of aspirin, a continent may have been lost.

With reason, Max Weber insisted that the most plausible counterfactuals were those that made only "minimal rewrites" of history.[48] Suppose we want to evaluate Ronald Reagan's role in ending the cold war by considering the likely course of that conflict in the late 1980s if he had not been president. It is more plausible to assume that Hinckley's bullet lodged a few millimeters closer to a vital organ than to concoct a complicated, multistep scenario to deprive Reagan of his 1980 electoral triumph. As a general rule, the fewer and more trivial the changes we introduce in history, the fewer the steps linking them to the hypothesized consequent and the less temporal distance between antecedent and consequent, the more plausible the counterfactual becomes. Not every small change will have significant, longer-term consequences; many, perhaps most, changes are likely to have consequences that are dampened over time. The real problem of counterfactual thought experimentation is trying to determine which minimal rewrites will affect the course of history.

Let us return to Cleopatra's nose. An ugly proboscis is a small change that might have had a big effect. It could have dampened Anthony's ardor, with important consequences for the struggle for power among Caesars' successors. So too might Roman history have been different if Anthony had been gay. We do not object to such counterfactuals because they are trivial but because they are arbitrary and

[47] Henry Ashby Turner, Jr., *Geissel des Jahrhunderts: Hitler and seine Hinterlassenschaft* (Berlin: Siedler, 1989); Geoffrey Parker, *The Grand Strategy of Philip II* (New Haven: Yale University Press, 1998a), 129–34; idem, "The Repulse of the English Fireships," in Robert Cowley, ed., *What If?* (New York: G. P. Putnam's Sons, 1998b).

[48] Weber, "Objective Possibility and Adequate Causation in Historical Explanation," in *The Methodology of the Social Sciences* (1905; Glencoe, Ill.: Free Press of Glencoe, 1949).

contrived. There is no particular reason why Cleopatra should have been less attractive or why Anthony should have had a different sexual orientation. Nor is it clear how these changes would have been brought about. Good counterfactuals arise from the context, and there must be compelling mechanisms to bring them into being.

Ferguson wants to legitimate counterfactual research, but his efforts to do so would put it in a straitjacket. He insists that we consider only those counterfactual scenarios that contemporary actors considered and committed to paper or some other form of record that is accepted by historians as a valid source.[49] Ferguson's criteria would exclude entire categories of plausible-world counterfactuals. It would limit counterfactuals to elites who made written records, to self-conscious decisions in which alternatives are likely to be carefully considered, and to political systems in which leaders and other important actors feel secure enough to write down their thoughts or share them with colleagues, journalists, family members, or friends. It would rule out all counterfactuals that were the result of impulsive behavior (or the lack of it), of human accident, oversight, obtuseness, or unanticipated error, of acts of nature, or of the confluence (or lack of it), or of independent chains of causation. We could not contemplate a world in which the duke of Alba remained healthy and did not relinquish command of his army to his son, Franz Ferdinand's touring car adhered to its planned route, Hitler died in an automobile accident, or Hinckley assassinated Ronald Reagan. Ferguson's criteria would also rule out all miracle counterfactuals.

The nine historical chapters of *Virtual History* are organized chronologically. They consist of one chapter each on the seventeenth and eighteenth centuries—England without Oliver Cromwell and America without a revolution—and seven on the twentieth century, ranging from an Irish Britain following the enactment of Home Rule in 1912 to the survival of the Soviet Union and Warsaw Pact in the absence of Mikhail Gorbachev. Some of the chapters are imaginative and convincingly argued (for example, John Adamson on England without Cromwell and J. C. D. Clark on British America) and sensibly ignore Ferguson's restrictive criteria for plausible-world counterfactuals. Other chapters are less convincing, among them, Ferguson's chapter on 1914. One chapter is downright nasty, partisan, and inappropriate to the volume: Diane Kunz does a hatchet job on John F. Kennedy without any serious counterfactual analysis. Ferguson's conclusion is as rambling as

[49] A similarly restrictive definition is offered by P. Nash, "The Use of Counterfactuals in History: A Look at the Literature," *SHAFR Newsletter* (March 1991).

his introduction and makes no attempt to tie the historical chapters together or draw any conclusions from their counterfactual explorations. Neither do any of the contributors in the absence of any guidance from the editor.

The Pity of War is a big book about big questions — ten of them to be precise. Ferguson wants to know if World War I was inevitable, why Germany's leaders gambled on war in 1914, why Britain intervened, whether the war was really popular at the outset, the extent to which propaganda kept it going, why the Schlieffen Plan failed, why Britain's economic advantages did not result in an earlier victory, why men kept fighting, why they stopped fighting, and who won the peace. To answer these questions, Ferguson relies heavily on counterfactual arguments on the grounds that they help recapture the uncertainty of the actors, to whom the future was merely a set of possibilities, and that they allow us to assess the extent to which they made the right choices.

Ferguson's questions are good ones; his answers are not. He appears to be following the A. J. P. Taylor approach to the writing of history: make outrageous arguments that stand history on its head, infuriate the critics, gain publicity, and with any luck, sell a lot of books. Taylor's *Origins of the Second World War* was a masterpiece of the genre.[50] It tried to shift much of the burden for the outbreak of war in 1939 from Adolf Hitler and Germany to Neville Chamberlain and Britain. Ferguson attempts to do the same—much less cleverly—for World War I.

Most historians assign primary responsibility for the First World War to Austria-Hungary and Germany. The dominant interpretation holds that Austria-Hungary exploited the assassination of Franz Ferdinand as the pretext for war with Serbia, and that Germany encouraged—even pushed—Austria toward decisive action.[51] Students of

[50] A. J. P. Taylor, *The Origins of the Second World War* (London: Hamish Hamilton, 1961).

[51] Fritz Stern, "Bethmann Hollweg and the War: The Limits of Responsibility," in Fritz Stern and Leonard Krieger, eds., *The Responsibility of Power: Historical Essays in Honor of Hajo Holborn* (Garden City, N.Y.: Doubleday, 1967); Konrad J. Jarausch, "The Illusion of Limited War: Chancellor Bethmann Hollweg's Calculated Risk, July 1914," *Central European History* 2, no. 1 (1969); idem, *The Enigmatic Chancellor: Bethmann Hollweg and the Hubris of Imperial Germany* (New Haven: Yale University Press, 1973); Karl Dietrich Erdmann, "War Guilt 1914 Reconsidered: A Balance of New Research," in H. W. Koch, ed., *The Origins of the First World War*, 2d ed. (Oxford: Oxford University Press, 1984); Wolfgang J. Mommsen, *Das Zeitalter des Imperialismus* (Frankfurt: Fischer Bucherei, 1969); idem, "Domestic Factors in German Foreign Policy before 1914," *Central European History* 6, no. 1 (1972); Andreas Hilgruber, "Riezlers Theorie des kalkulieren Risikos und Bethmann Hollwegs politische Konzeption in der Julikrise 1914," *Historische Zeitschrift* 202 (April 1966); Egmont Zechlin, *Krieg und Kriegrisiko: Zur deutschen Politik im Ersten Weltkrieg* (Düsseldorf: Droste, 1979); David Kaiser, "Germany and the Origins of the First World War," *Journal of Modern History* 55 (December 1983); Hartmut Pogge von Strandmann, "Germany and the Coming of War," in R. J. Evans and Hartmut Pogge von Strandmann, *The Coming of the First World War* (Oxford: Oxford University Press, 1988); Holger H. Herwig, *The First World War: Germany and Austria-Hungary, 1914–1918* (London: Arnold, 1998),

Austria-Hungary argue that its leaders acted out a combination of closely related foreign and domestic concerns. Since the publication of Kurt Riezler's diaries, something of a consensus has emerged among German historians that Chancellor Theobald von Bethmann-Hollweg did not seek to provoke a European war but recognized that an Austrian conflict with Serbia would be difficult to contain. He also believed that a continental war was inevitable and that Germany's chance of winning it would only diminish with every passing year.[52] Hence, sooner was preferable to later.

To the extent that Britain has been criticized for its policy in 1914, it has usually been for Sir Edward Grey's naïveté. The foreign secretary recognized only belatedly that Germany was playing a double game, and for most of the crisis he had tried to get Berlin to restrain Austria and cosponsor a conference to resolve the problem by diplomacy. It is sometimes alleged that Britain might have prevented war if only Grey had put Germany on notice early on in the crisis that Britain would come to the aid of France if Germany violated Belgian neutrality.[53] Ferguson propounds a different idea in *The Pity of War*. Germany was on the defensive and had been put there by other powers, most notably Britain: "The Committee of Imperial Defence meeting of 23 August (rather than the notorious meeting between the Kaiser and his military chiefs sixteen months later) was the real 'war council' which set the course for a military confrontation between Britain and Germany" (pp. 64–65). This meeting acquiesced in the strategy of sending an expeditionary force to the Continent if Britain sided with France in a war with Germany. Britain's plans for an aggressive land and naval policy, its naval agreement with France in the Mediterranean, and secret discussions with Russia in 1914 made "German fears of encirclement seem less like paranoia than realism" (p. 68).

Ferguson contends that the British rode roughshod over Germany's interests because "they realized she did *not* pose a threat." The Germans

chap. 1; John C. G. Röhl, "Germany," in Keith Wilson, ed., *Decisions for War, 1914* (New York: St. Martin's Press, 1995). On Austria, see Samuel R. Williamson, Jr., "Influence, Power, and the Policy Process: The Case of Franz Ferdinand," *Historical Journal* 17, no. 2 (1974); idem, "The Origins of World War I," *Journal of Interdisciplinary History* 18 (Spring 1988); idem, *Austria-Hungary and the Coming of the First World War* (London: Macmillan, 1990); and R. J. Evans, "The Habsburg Monarchy and the Coming of War," in Evans and Pogge von Strandmann, eds., *The Coming of the First World War*.

[52] Riezler, *Tagebücher, Aufsätze, Dokumente: Eingeleidet und hrsg. Von Karl Dietrich Erdmann* (Göttingen: Vandenhoeck & Ruprecht, 1972); Koch (fn. 51), introduction.

[53] See, for example, Sean M. Lynn-Jones, "Détente and Deterrence: Anglo-German Relations, 1911–1914," *International Security* 11 (Fall 1986). German scholarship is very skeptical about such claims. Karl Dietrich Erdmann, "War Guilt 1914 Reconsidered: A Balance of New Research," in Koch (fn. 51); Zechlin (fn. 51); Jarausch (fn. 51, 1973), 167, 170.

were not aggressive but were pushed toward war by the military preparations of their neighbors, including Britain, and resolved to fight while they still had a chance to win (pp. 151–52). This argument leads to the paradoxical conclusion—and the key counterfactual of the book—that "a *more* militaristic Germany might have averted the First World War" (p. 142). The counterfactual rests on three assumptions: (1) that the financial constraint on Germany's military capability was "perhaps *the* crucial factor in the calculations of the German General Staff in 1914"; (2) that General Erich von Ludendorff's 1913 "maximum plan" to add three hundred thousand men to the army was financially feasible; and (3) that a more powerful Germany would have been more restrained in its foreign policy (pp. 141–42).

The German general staff was deeply concerned about the strategic implication of Russian and French military preparations, but this was because of the infamous Schlieffen Plan. It committed Germany to an all-out offensive in the West to defeat France before the more slowly mobilizing Russia could bring its forces to bear against Germany in the East. If the general staff had adopted a defensive strategy, a more appropriate strategy for an allegedly status quo power, they could have regarded French and Russian military preparations with more equanimity. The Ludendorff plan may have been financially feasible for the reasons Ferguson adduces: additional taxes could have been levied, or the Reichsbank, which was hoarding gold, could have purchased a substantial issue of treasury bills. But the counterfactual violates one of the criteria Ferguson proposed in *Virtual History:* the German chancellor never seriously considered the Ludendorff plan as an option. He did not, because increased military expenditures in 1913 were, in Ferguson's own words, "politically impossible" (p. 141). This admission pulls the rug out from under the counterfactual. The pièce de résistance is Ferguson's bald assertion that a more powerful Germany would have been a more peaceful Germany—while providing no grounds whatsoever for believing this could have been. It seems at least as likely that with a larger army in 1914, German leaders would have been *more* confident of victory and *more* intent on going to war. In this connection, we need to remember that an important catalyst for the German decision to risk war in the July crisis was the perceived need to bolster Austria-Hungary, and this consideration would still have been present and equally important.

Ferguson fails to consider any additional counterfactuals that might have followed from his counterfactual. If Germany had increased its armed forces, how would France and Russia have responded? If they

had each increased their armed forces by only half as much, the German general staff would have been right back where it started. Expectation of tit-for-tat escalation—a characteristic of prewar arms races—may well have been an additional reason why Bethmann-Hollweg did not greet Ludendorff's proposal with enthusiasm. Any number of other changes are possible as a consequence of Ferguson's counterfactuals, and a good historian should attempt to identify the most important and explore their implications for Germany policy and the likelihood of war.

Ferguson's counterfactual is politically and methodologically unsound. Most of his other counterfactuals are similarly incomplete and unconvincing. Consider the alternative strategies by which Germany might have achieved a military victory. Ferguson rejects out of hand the *Ostaufmarsch* plan, by which the German offensive would have been directed against Russia instead of France, and the related suggestion that Falkenhayn should have refrained from attacking Verdun, remained on the defensive in the West, and concentrated on defeating Russia. He suggests that Ludendorff might have obtained a negotiated peace if he had offered to relinquish Belgium after his spring 1918 offensive stalled on April 5. Instead of pressing on with the attack, Ludendorff should have withdrawn German forces to the Hindenburg Line and asked the allies for an armistice. This is not a course of action that seems to have crossed Ludendorff's mind at the time, and Ferguson gives no reasons for believing that the allies would have responded favorably (pp. 315–17).

The mother of all World War I counterfactuals is Ferguson's outlandish suggestion that had Britain stood aside in 1914, even for a matter of weeks, "continental Europe could therefore have been transformed into something not wholly unlike the European Union we know today" (pp. 443–44, 457–61). This benign outcome assumes that Germany would have defeated France without the rapid intervention of the British Expeditionary Force—an argument about which Ferguson is only lukewarm elsewhere in the volume. It also assumes a long string of second-order counterfactuals (for example, Russia is also defeated, Berlin organizes the Continent in a manner that allows other nations to benefit, these states and people become reconciled to German dominance, Germany is a sated power and conducts no further wars in Europe and makes no additional colonial demands, and its political culture and institutions evolve to become something akin to the liberal, peaceful, Europe-oriented country that it is today). Ferguson does not address these questions and offers no chain of logic to connect

the counterfactual of a German victory to the peaceful and consensual Europe he posits as the long-term outcome. Ferguson is reacting to the arguments of Grey and "other Germanophobes" that a German victory would have been disastrous for British interests and goes full tilt—and foolishly so—in the opposite direction (pp. 166–68).

"Plausible World" Counterfactuals

More thoughtful historians and social scientists have pondered the problem of plausible-world counterfactuals and appropriate criteria for using them. There is no consensus about what constitutes a good counterfactual, but there is a common recognition that it is extraordinarily difficult to construct a robust counterfactual—one whose antecedent we can assert with confidence could have led to the hypothesized consequent. There are three reasons for this well-warranted pessimism: the statistical improbability of multistep counterfactuals, the interconnectedness of events, and the unpredictable effects of second-order counterfactuals.

Compound Probability

The probability of a consequent is a multiple of the probability of each counterfactual linking the hypothesized antecedent to it. Suppose I contend that neither world war nor the Holocaust would have occurred if Mozart had lived to the age of sixty-five.[54] Having pushed classical form as far as it could go in the Jupiter Symphony, his last three operas, and the requiem, Mozart's next dramatic works would have been the precursors of a new, "postclassicist" style. He would have created a viable alternative to romanticism that would have been widely imitated by composers, writers, and artists. Postclassicism would have kept the political ideas of the Enlightenment alive and held romanticism in check. Nationalism would have been more restrained, and thus Austria-Hungary and Germany would have undergone very different political evolution. This alternative and vastly preferable world has at least five counterfactual steps linking antecedent to consequent: Mozart must survive to old age and develop a new style of artistic expression; subsequent composers, artists, and writers must imitate and elaborate it; romanticism must become to some degree marginalized; and artistic developments must have important political ramifications. This last

[54] Richard Ned Lebow, "What If Mozart Had Died at Your Age?" (Manuscript, Mershon Center, Ohio State University, April 2000).

counterfactual presupposes numerous other enabling counterfactuals about the nature of the political changes that will lead to the hypothesized consequent (for example, internal reforms that resolve or reduce the threat of internal dissolution of Austria-Hungary, German unification under different terms, or at least a Germany satisfied with the status quo, no First World War, no Hitler and no Holocaust without Germany's defeat in World War I). Even if every one of this long string of counterfactuals had a probability of at least 50 percent, the overall probability of the consequent would be a mere .03 for five steps and a frighteningly low .003 for eight steps. This particular counterfactual may appear far-fetched, but most interesting counterfactuals are no less improbable statistically. They may start with a tiny and plausible alteration of the real world but then infer numerous follow-on developments to end up with a major change in reality.

Interconnectedness

Scholars not infrequently assume that one aspect of the past can be changed and everything else kept constant. Mueller's cold war counterfactual is a case in point. He analyzes postwar history as it actually happened, including the Cuban missile crisis, to see if any major outcome would have been different in the absence of nuclear weapons. But what incentive would Khrushchev have had to deploy conventionally armed missiles in Cuba? The missiles could not have deterred an American invasion and might well have invited one, and this was the very event Khrushchev hoped to prevent.[55] Without a Cuban missile crisis, which had significant consequences for the future course of Soviet-American relations, the cold war would have evolved differently, and the course of future superpower relations, benign or malign, is impossible to predict.

"Surgical" counterfactuals are no more realistic than surgical air strikes. Causes are interdependent and have important interaction effects. Even minimal rewrites of history may alter the context in such a way as to render the consequent moot or to undercut the chain of events or logic leading to it. Consider another missile crisis counterfactual: if Richard Nixon had won the 1960 presidential election—and he lost by the narrowest of margins—he would have ordered an air strike against the Soviet missiles in Cuba. It is reasonable to assume that Nixon would have preferred an air strike to a blockade because he was more hawkish than Kennedy and would not have had a secretary of defense like Robert McNamara to make the case for restraint. But for

[55] For this criticism of Mueller, see Cederman (fn. 32), 253, 255.

these same reasons, Nixon might well have committed American forces to the faltering Bay of Pigs invasion in April 1961. If Castro had been overthrown, there would have been no communist Cuba to which Khrushchev could send missiles a year later.

The Nixon example invokes a counterfactual arising from the antecedent—but one outside the chain of logic leading from it to the consequent—to render the missile crisis moot. If Nixon had been elected president in 1960, the world would have been different in many ways, some of them with implications that are impossible to trace. He would have appointed a different defense secretary, who in turn would have appointed a different chairman of the joint chiefs of staff. Personnel changes at the top would have had amplifying consequences for promotions and appointments further down the line. To the extent that the behavior of individual officers could have important foreign policy implications—and it certainly did during the Berlin and Cuban missile crises—all kinds of possibilities open up.

Second-Order Counterfactuals

The problem of prediction is further complicated by the fact that the clock of history does not stop if and when the hypothesized consequent is reached. Subsequent developments can return history to the course from which the antecedent was intended to divert it.[56] Colin Martin and Geoffrey Parker show that the defeat of the Spanish Armada was a near event; they suggest that better communication, different decisions by local commanders, or better weather might have allowed the Spanish to land an invasion force in England. If Spain had put an army ashore, it almost certainly would have conquered the country. Martin and Parker go on to consider what would have happened next: Philip II was succeeded by Philip III, a far less capable ruler, who would have had enormous difficulty in maintaining an already overextended empire. In relatively short order, they believe, England would have overthrown the Spanish yoke.[57]

Some counterfactuals, like the "butterfly effect," introduce small changes that have major, lasting, long-term effects. Others, like a Spanish victory in 1588, are big changes that appear to have big consequences, but the changes they introduce may be damped down over time and end up having little lasting effect. It is also possible that second counterfactuals arising from a Spanish occupation of England,

[56] Tetlock, ""Distinguishing Frivolous from Serious Counterfactuals" (Manuscript, 1999).
[57] Parker (fn. 47, 1998a), 281–96.

however brief that occupation, could have had dramatic and long-lasting changes for European politics or in other realms that Marin and Parker did not consider. None of these outcomes is predictable, and the butterfly effect may not be knowable in advance or even in retrospect.[58]

Criteria for Use

Recognition that counterfactual arguments often have indeterminate consequences has prompted scholars to impose restrictive criteria on their use. Fearon proposes a proximity criterion. We should consider only those counterfactuals in which the antecedent appears likely to bring about the intended consequent and little else. Counterfactuals, he suggests, must be limited to cases where "the proposed causes are temporally and, in some sense, spatially quite close to the consequents."[59] Dawes argues that counterfactual inferences are warranted "if and only if they are embedded in a system of statistical contingency for which we have reasonable evidence."[60] Kiser and Levi suggest that counterfactuals are best used as substitutes for direct empirical analysis when data are limited or unavailable. Such counterfactuals should be based on a general deductive theory with clear microfoundations and scope conditions.[61] Elster, who also insists that good counterfactuals are derived from good theories, believes that there is only a narrow window for such experimentation: "The theory must be weak enough to admit the counterfactual assumption, and also strong enough to permit a clear-cut conclusion."[62]

Criteria that tie counterfactuals to established laws and statistical generalizations and attempt to limit second-order counterfactuals are superficially appealing. In practice, they are generally unworkable or would rule out some of the most important uses of counterfactual experimentation. The Fearon proximity criterion suffers from both defects. The requirement of a minimal cause that produces only a minimal effect is extraordinarily restrictive. Weber rightly observes that

[58] See M. Maruyama, "The Second Cybernetics: Deviation-Amplifying Mutual Causal Processes," *American Scientist* 51, no. 2 (1963). Maruyama contends that the very essence of the "butterfly effect" is that it cannot be discovered. He offers the example of a city built on the American plains because the first white settler awoke on an especially beautiful morning and took it as a sign to put down his stakes here rather than somewhere else. The coincidence of weather and intent was unpredictable and would be invisible as a cause to researchers a century later attempting to discover why the city grew up where it did.

[59] Fearon (fn. 29), 50.

[60] Dawes, "Counterfactual Inferences as Instances of Statistical Inferences," in Tetlock and Belkin (fn. 7)

[61] Edgar Kiser and Margaret Levi, "Using Counterfactuals in Historical Analysis: Theories of Revolution," in Tetlock and Belkin (fn. 7).

[62] Elster (fn. 38), 184–85.

"it rewards the psychologically easy and comfortable task of generating counterfactuals close to the margins of existing theories. It predisposes toward varying only the familiar variables, the ones that we think we know are tied into causal paths that we feel we know well." It also presupposes that we know what "minimal" really means and that we have a rather complete understanding of the behavior in question and its likely consequences. "If we knew this," Weber continues, "we would no longer need counterfactuals."[63] Counterfactuals almost always have multiple consequences, and a counterfactual powerful enough to test a theory that makes only one small change in reality is probably an oxymoron.

Elster, Kiser and Levi, and Tetlock and Belkin favor counterfactuals that are derived deductively from good theories.[64] This may be possible in a data-rich and reductionist field like cognitive psychology. But as Breslauer contends, it is hardly a realistic standard for counterfactuals in history and most of the social sciences, where after fifty years of intensive behavioral research there are no "established" theoretical laws or generalizations.[65] In political science, where researchers do not agree on the meaning of "cause" or "fact," it is unlikely that they will ever agree about what constitutes a good theory and, by extension, a useful or valid counterfactual.[66]

In the absence of established theory, Tetlock and Belkin worry that competing schools of thought will invent "counterfactuals of convenience." Breslauer and Herrmann and Fischerkeller report that this was a common practice in Soviet studies during the cold war. The widespread appeal of the expansionist "theory" of Soviet foreign policy made American policymakers and scholars receptive to counterfactuals that did not meet the Tetlock-Belkin criteria. Tetlock and Belkin would eliminate this problem by insisting that counterfactuals "must stimulate testable propositions that hold up reasonably well against new data." Herrmann and Fischerkeller point out that even if good evidence had been available, neither the expansionist not the defensive "theories" of Soviet behavior were formulated in ways to make them falsifiable.[67]

[63] Weber, "Counterfactuals, Past and Future," in Tetlock and Belkin (fn. 7), 278, 272.

[64] Elster (fn. 38); Kiser and Levi (fn. 61); and Tetlock and Belkin (fn. 23) also believe this is a reasonable standard for judging the plausibility of counterfactuals; Tetlock and Belkin acknowledge that "we should expect disagreement about what counts as well-established theory in world politics" (p. 27).

[65] Breslauer (fn. 34), 73; Steven Bernstein et al., "G-d Gave Physics All the Easy Problems: Adapting Social Science to an Unpredictable World," *European Journal of International Relations* 6 (January 2000).

[66] Weber (fn. 63), 271.

[67] Breslauer (fn. 34); Richard K. Herrmann and Michael P. Fischerkeller, "Counterfactual Reasoning in Motivational Analysis: U.S. Policy toward Iran," in Tetlock and Belkin (fn. 7); Tetlock and Belkin (fn. 23), 27.

Dawes's criterion of well-established statistical regularities is open to the same criticism. Based on an examination of the American literature on Iran, Herrmann and Fischerkeller conclude that "too often in world politics what is taken as a base rate for a generalization about the motives of another country is too much an ideological conviction and too little a product of deductive and empirical behavioral science."[68] Lebow and Stein have documented the same phenomenon with respect to deterrence; data sets used to test the strategy of deterrence were patently ideological in the cases they recognized as deterrence encounters and coded as successes for the West.[69]

There are other fundamental problems with attempts to derive counterfactuals from theory. Most theories rely on "structural" variables for their analytical power. To introduce variation in structure, it is necessary to make major changes in reality, changes that violate not only Fearon's proximity criterion but also any conception of a "minimal" or plausible rewrite of history. In international relations, most of our theories are at the levels of either the state (for example, democratic peace) or the system (for example, balance of power, power transition). To manipulate the relevant variables counterfactually, it would be necessary to make structural changes in the constitutions or capabilities of states or in the polarity of the system. It is inconceivable that any change of this kind could be brought about by a plausible-world counterfactual, unless it is introduced at some temporal remove and expected to have a long-term, cumulative effect. Multiple-step counterfactuals with unknowable second-order counterfactuals are unacceptable to theory-driven researchers. To change governments, capabilities, or polarity we must rely on miracle counterfactuals.[70] They require additional counterfactuals to bring them about, counterfactuals that researchers cannot in good conscience ignore because of their possible implications for the consequent. If we accelerated the decline of Great Britain in the nineteenth century relative to Germany to evaluate power-transition theory or if we made post-1918 Russia a democracy to test the relative importance of capabilities versus regime type, we would need to ask what could cause Britain's decline or Russia's emergence as a democracy. We would probably have to transform the economic development, class structure, and ideology of these states, and perhaps the outcomes of

[68] Herrmann and Fischerkeller (fn. 67), 146–47.

[69] Richard Ned Lebow and Janice Gross Stein, "Deterrence: The Elusive Dependent Variable," *World Politics* 42 (April 1990); and idem, *When Does Deterrence Succeed and How Do We Know?* (Ottawa: Canadian Institute for International Peace and Security, 1990).

[70] Fearon (fn. 29) acknowledges the utility of miracle counterfactuals but not the contradiction between their use and his requirement that counterfactuals be deduced from theories (pp. 60–65).

wars in which they had engaged. Any of these changes could significantly alter their foreign policies and perhaps the rules and norms of the international system.

None of the miracle counterfactuals we could use to alter regime type, capability, or polarity can be derived from existing international relations theories. They would all be ad hoc. The same holds true for most plausible-world counterfactuals. Assassinations, ill-health, last-minute changes in travel plans, undelivered messages, misunderstandings, accidents, and coincidences—all the most credible vehicles of plausible-world counterfactuals—are random events, often unconnected with politics and certainly outside of any theory of politics. Why should this be considered a methodological weakness? In the health and physical sciences, where counterfactual experimentation is routine, counterfactuals are almost always independent of theory. They take the form of imaginary, often knowingly unrealistic parameters for key variables. Epidemiologists have robust equations to describe the spread of infectious diseases. In studying HIV, they might assign a range of arbitrary values to its virulence, the time that elapses between infection and the onset of symptoms, the rate of infection (by altering sexual practices or illicit drug use), or the nature and success rate of treatments, all with the goal of determining the combinations of social and medical practices that will be most effective in limiting the spread of the virus. So too in social science, the appropriate criteria for good counterfactuals have nothing to do with their origins, but rather concern how they are used.

This brings us to the nub of the problem. The scholars most troubled by the inherent unpredictability of counterfactual outcomes are those like Elster, Fearon, and Dawes who want to use counterfactuals to test propositions and theories. The less demonstrable the consequent, the less useful a counterfactual is for this purpose. Fearon is willing to consider a fallback position. Although it may be impossible to prove that x caused y, it may be possible, he suggests, to demonstrate that without x, whatever happened would not have been y.[71] The author employed this strategy to critique structural theories that attempt to explain transformations of the international system. I used counterfactual arguments to show that these transformations were contingent in a double sense. They were the result of independent chains of causation that produced nonlinear effects when they came into confluence. The transformations were also dependent on catalysts, whose presence was problematic and whose causes were independent of any of the underly-

[71] Fearon (fn. 29), 54.

ing causes of the transformations. Both the confluence and the catalysts could easily have been prevented by minimal rewrite counterfactuals. At best, therefore, structural theories can identify conditions under which transformations become possible.[72]

Scientific validation of a counterfactual is important only if the goal is to test a proposition or theory. Testing is only one part of theory building and by no means the most important part. Even within the neopositivist tradition, the prior steps of identifying important questions or anomalies and formulating theories to explain or resolve them are generally recognized as critical. I argued earlier that counterfactuals serve these ends admirably by making scholars more sensitive to contingency, helping them work through the implications of existing theories and identify gaps and inconsistencies in them. Theory building is only one goal of social science. It also aims to broaden our intellectual horizons and to provide methods relevant to assessing the relative benefits and value of policy outcomes, real and hypothetical. Counterfactual experimentation is essential to these tasks and can be used effectively without requiring certainty that antecedents will lead to specific consequents. Steven Weber has rightly observed that counterfactuals are better used as "mind-set changers" and "learning devices" rather than as data points in explanation."[73] We should worry less about the uncertainty of counterfactual experimentation and think more about its mind-opening implications.

For most of these purposes described above, the clarity, completeness, and logical consistency of the arguments linking antecedent to consequents are more important than their external validity. I accordingly propose eight criteria for plausible-world counterfactuals. Numbers 1, 2, 4, and 5 are drawn from or are variants of the Tetlock-Belkin list, number 6 was recently proposed by Tetlock, and numbers 3, 7 and 8 are the author's.[74]

1. *Clarity.* All causal arguments should define as unambiguously as possible what is to be explained (the consequent in counterfactual arguments), what accounts for this outcome (the antecedent), and the principle(s) linking the two. Good counterfactuals should also specify the conditions that would have to be present for the counterfactual to occur. Some historians have argued that timely public health measures could have significantly reduced the mortality in Europe associated with the Black Death pandemic of the fourteenth century. For Euro-

[72] Lebow (fn. 37).
[73] Weber (fn. 63), 270.
[74] Tetlock and Belkin (fn. 23), 16–32; Tetlock (fn. 56); Lebow and Stein (fn. 8), 146–47.

pean communities to have implemented these measures, they would have had to recognize that human intervention could affect the spread of disease, and they would have needed the authority and will to impose draconian measures on travel and trade over the likely objections of the wealthy and merchant classes.[75] Both additional conditions are unrealistic given the values, knowledge, and political structure of the age; large-scale quarantines would not be implemented to combat plague until the eighteenth century.[76] Plausible-world counterfactuals not only require realistic antecedents, but the antecedents themselves must not require other, implausible conditions or counterfactuals.

2. *Logical consistency or cotenability*. Every counterfactual is a shorthand statement of a more complex argument that generally requires a set of connecting conditions or principles. The hypothetical antecedent should not undercut any of the principles linking it to the consequent. A case in point is Robert Fogel's famous argument that if railroads had not existed, the American economy in the nineteenth century would have grown only slightly more slowly than it actually did because a strong incentive would have existed to invent the internal combustion engine sooner.[77] Elster has rightly objected that if the technology were present to invent and produce automobiles, it would almost certainly have also led to the development of railroads.[78]

3. *Enabling counterfactuals should not undercut the antecedent*. Counterfactuals may require other counterfactuals to make them possible (for example, for Richard Nixon to have been president at the time of the Cuban missile crisis, he would have had to have won the 1960 election, and that would have required significant changes in the political context at home and possibly abroad. These changes might have had significant implications for both American and Soviet foreign policy.) Researchers need to specify all important enabling counterfactuals and consider their implications for the consequent.

4. *Historical consistency*. Max Weber insisted that plausible counterfactuals should make as few historical changes as possible on the grounds that the more we disturb the values, goals and contexts in which actors operate, the less predictable their behavior becomes.[79] Counterfactual arguments that make a credible case for a dramatically

[75] Hawthorn (fn. 38), 31–60.

[76] William H. McNeill, *Plagues and People* (Garden City, N.Y.: Doubleday, 1976), 239–40, 254.

[77] Fogel, *Railroads and American Economic Growth: Essays in Econometric History* (Baltimore: Johns Hopkins University Press, 1964).

[78] Elster (fn. 38), 204–8; Tetlock and Belkin (fn. 23), 22–23.

[79] Weber (fn. 48).

different future on the basis of one small change in reality are very powerful, and the minimal rewrite rule should be followed whenever possible. The *nature* of the changes made by the experiment are nevertheless more important than the *number* of changes. A minimal rewrite that makes only one alteration in reality may not qualify as a plausible-world counterfactual if the counterfactual is unrealistic or if numerous subsequent counterfactual steps are necessary to reach the hypothesized consequent. A counterfactual based on several small changes, all of them realistic, may be more plausible, especially if they lead more directly to the consequent.

5. *Theoretical consistency*. There are few, if any, generally accepted theories in the social sciences, and none in international relations, comparative politics, or history. For purposes of counterfactual analysis, it is nevertheless useful to reference any theories, empirical findings, historical interpretations, or assumptions on which the causal principles or connecting arguments are based. This will provide readers with a more explicit perspective from which to evaluate the counterfactual's plausibility.

6. *Avoid the conjunction fallacy*. There are good statistical grounds for the "minimal rewrite" rule, as the probability of a consequent is the multiple of the probability of each counterfactual step linking the antecedent to it. We nevertheless need to recognize the conservative bias inherent in statistical reasoning. The laws of statistics indicate that the probability of *any* compound counterfactual is exceedingly low. This does not mean that the current state of affairs was overdetermined, only that it is very unlikely that hypothesized antecedents will produce *specific* consequences at any temporal distance. Social and political developments are highly contingent, and the future is undetermined—as was the past before it became the present. The long-term consequences of change are unpredictable. If Mozart had lived to sixty-five, today's world could well have turned out to be strikingly different from the world we know. But many alternative worlds are possible, and the probability of any one of them coming to pass is exceedingly low. Counterfactuals might have changed the world but in ways that become exponentially more difficult to track over time because of the additional branching points that enter the picture. As the probabilities associated with these outcomes will vary enormously, researchers accordingly need to specify if their counterfactuals are intended to produce a specific world, a set of worlds with particular characteristics, or *any* world (on a specific dimension) other than the one that actually came to pass.

7. *Recognize the interconnectedness of causes and outcomes*. Surgical counterfactuals are unrealistic because causes are interdependent and have important interaction effects. Changes we make in the past may require other changes to make them possible and may also produce additional changes beyond those we intend to lead to the consequent. History is like a spring mattress: if one of the springs is cut or simply subjected to extra pressure, the others will also to varying degrees shift their location and tension.[80] Earlier we considered the counterfactual that "President" Nixon would have ordered an air strike and follow-up invasion of Cuba, which in turn would have triggered a nuclear war. But if Nixon had been president, he probably would have committed American forces to save the faltering Bay of Pigs invasion and Castro would have been overthrown. A subsequent Soviet missile deployment would have become moot. Good counterfactuals must specify what else might change as a result of a hypothesized antecedent, and they must consider how the most important of these changes might interact and influence the probability of the consequent.

8. *Consider second-order counterfactuals*. Even when there is good to reason to believe that the antecedent will produce the desired consequent, the possibility remains that subsequent developments will return history to the course from which it was initially diverted by the antecedent. This might be the long-term result of enabling counterfactuals necessary to bring about the antecedent, of follow-on counterfactuals produced by the antecedent, of counterfactuals arising from the consequent, or of interaction among any combination of these counterfactuals. Interaction effects among second-order counterfactuals might be considered third-order counterfactuals, and they too can have profound consequences for the subsequent course of development.

Attempts to identify and analyze *all* of the counterfactuals arising from the antecedent and consequent would quickly lead to an infinite regress. Researchers should nevertheless try to identify what in their view is the most likely course of events that could unravel their consequent or negate its value as an outcome. The last point entails the recognition that we choose a consequent because of some larger effect it is intended to have. If other developments make it unlikely that the consequent will have that effect, it may lose its attractiveness. No counterfactual argument is complete without some argument about "alter-

[80] Friedrich Engels suggested something similar. History was a "parallelogram of forces." If one person shook his arms to move one corner of the parallelogram, it affected parts of the figure far away and far removed from intentions of the actor. Richard J. Evans, *In Defense of History* (New York: Norton, 1999), 118.

native" alternative futures and some assessment of their likelihood and implications for both the consequent and its value as a consequent.

These criteria will not allow researchers to validate plausible-world counterfactuals, but they will help them weed out poor counterfactuals primarily on the basis of clarity and logical and substantive completeness. Most of the criteria are not applicable to miracle-world counterfactuals, which, by definition, are not required to meet any real-world tests. The value of such a counterfactual is based entirely on its ability to provoke or, better yet, to compel researchers to think about issues and problems they would not otherwise address, or to look at them in a new light. For a field where careful, technical work is increasingly valued over imagination, miracle-world counterfactuals can refocus our attention on important, big questions.