

CHAPTER TWO

**COUNTERFACTUAL THOUGHT EXPERIMENTS:  
A NECESSARY RESEARCH TOOL**

To appear in Phil A. Tetlock, Richard Ned Lebow  
and Geoffrey Parker, eds.

**UNMAKING THE WEST: COUNTERFACTUAL  
THOUGHT EXPERIMENTS IN HISTORY**

Richard Ned Lebow  
The Mershon Center  
The Ohio State University  
lebow.1@osu.edu

April 2001

Counterfactuals are routinely used in the physical and biological sciences to develop and evaluate sophisticated, non-linear models.<sup>1</sup> They have been used with telling effect in the study of economic history and American politics.<sup>2</sup> But for most historians counterfactual arguments have no scholarly standing. They are flights of fancy, fun over a beer or two in the faculty club, but not the stuff of serious research.<sup>3</sup> This dismissive attitude may be encouraged by the emergence and popularity of counterfactual historical as a fictional genre, and the uncomfortable similarities between some recent works of counterfactual scholarship and such fiction. Our volume rests on the assumption that counterfactuals can be used as an effective research tool, but this requires a clear understanding of the nature of counterfactuals, the circumstances to which they are best suited and appropriate protocols for conducting counterfactual thought experiments.

With these ends in mind, I begin my chapter with an exploration of the differences between counterfactual and so-called “factual” arguments and offer the proposition that the difference between is greatly exaggerated; it is one of degree, not of kind. I go on to discuss different uses of counterfactual thought experiments for historians. In the course of the discussion, I distinguish between miracle and plausible world and scholarly and folk counterfactuals and their respective uses. I critique a recent historical work based on counterfactuals and some of the relevant literature in social science. On the basis of these evaluations, I propose nine criteria to guide plausible world counterfactuals. I conclude by looking at the special problems of applying counterfactual analysis to a problem as broad as the rise and success of the West.

## COUNTERFACTUALS VERSUS FACTUALS

Counterfactuals are “what if” statements, usually about the past. Counterfactual experiments vary aspects of the past and analyze how these changes would have affected the course of events. In history, such experiments always have uncertain outcomes because we can neither predict the future nor rerun the tape of history to see what might actually happen.

The speculative nature of counterfactuals makes many scholars wary of them. But counterfactual analysis can often be conducted in an evidence rich environment. In the aftermath of Aldrich Ames’ arrest as a Soviet spy, the Central Intelligence Agency convened a team of counter-intelligence 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. Their knowledge of Ames’ personality, motives and behavior and of the *modus operandi* of his Soviet spy masters allowed them to conduct their inquiry with what they believed to be a high degree of precision.<sup>4</sup> In New York City authorities recently used the fear of widespread Y2K computer failures to test the ability of City agencies to respond to a complex emergency. Officials were confident that such breakdowns would not occur, and used the scenarios as “future counterfactuals” to provide useful data to identify bottlenecks and improve performance.<sup>5</sup>

The controversy surrounding the strategy of deterrence offers a history-international relations example. 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.<sup>6</sup> 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 another evidence-rich environment, and one in which counterfactuals drove policy and can now be tested. Nikita Khrushchev's decision to send and remove missiles from Cuba, and Kennedy's decision to impose a blockade were contingent upon hypothetical antecedents. Kennedy believed, incorrectly, as it turned out, that Khrushchev sent missiles to Cuba because he doubted his resolve, and would not have done so if he had 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. Recent evidence from Soviet and American archives and interviews with former officials make it possible to explore the validity of most of these counterfactuals, and thus to evaluate the choices of Soviet and American leaders and subsequent scholarly analyses of the crisis.<sup>7</sup>

Quantitative counterfactual analysis is another possibility. Jay Winter exploits counterfactual projections of mortality rates based on pre-war 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 by of five year cohorts for each year of the war.<sup>8</sup>

Real cases can be framed as counterfactuals, and their evidentiary base can be exploited to make counterfactual comparisons. Charles Tilly uses imperial China to construct a counterfactual Europe to discover the conditions under which post-Westphalian Europe, like China, might have been dominated by a single political unit that excluded the majority of the population from any political participation.<sup>9</sup> Tin-bor Victoria Hui employs this strategy to examine state formation in China and Europe. She identified a series of processes common to both regions and has uses China as a “real world counterfactual” to evaluate state formation in Europe, and vice versa.<sup>10</sup> In our volume Kenneth Pomeranz uses Europe as a real counterfactual case to explain why China did not have an independent industrial revolution.<sup>11</sup> All these experiments are based on the assumption that similar causal mechanisms may operate in diverse contexts and thus have different consequences. By examining the differential impact of these mechanisms, and how they were determined by local conditions, we can learn a lot about different trajectories of historical development.<sup>12</sup>

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 going 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 to its olive orchards and vineyards. After a few years of frustration, Pericles expected the Spartans to recognize the futility of waging war against Athens, the peace faction, led by King Archidamus, to regain power, and the two hegemons to reach a more lasting accommodation.<sup>13</sup> This scenario is purely speculative, and intended to explain away behavior that would otherwise appear 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 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 of 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 advisors. Khrushchev further complicated the picture by telling various officials what they wanted to hear and thus what he thought most likely to garner their support.<sup>14</sup>

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 that 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, and governments 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.<sup>15</sup> Elvin musters considerable evidence in support of his thesis, much of it from primary sources, but it is all in the way of illustration. Nowhere is he able to show that Chinese leaders took 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 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 made, 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 (e.g., 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.<sup>16</sup> 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 hardliners, 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 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 be 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 (if I don’t withdraw the missiles, Kennedy will attack Cuba), and with the passage of time they became historical counterfactuals (Kennedy would have

attacked Cuba if Khrushchev had not agreed to withdraw the Soviet missiles).

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 hijacked by Stalin. The implication is that socialism would have developed differently without him.<sup>17</sup> After Cuba, former Kennedy administration officials and many scholars maintained that Khrushchev would not have deployed missiles in Cuba if Kennedy had behaved more decisively at the Bay of Pigs, 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 became available in the Gorbachev era suggested that Khrushchev decided to send missiles secretly to Cuba because he *overestimated* Kennedy's resolve; he feared that Kennedy was preparing to invade Cuba and would send the American navy to stop any ships carrying missiles to Cuba to deter that invasion.<sup>18</sup> 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 factials 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 the existence of reality quite independently of any attempt to understand it by human beings, or that some understandings may transcend culture.

Physical scientists may be correct in their claim 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.”<sup>19</sup>

“Temperature” is undeniably a social construction, but it is a measure of something observable and real: changes in the energy levels of molecules. 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 only rarely have -- 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.<sup>20</sup> The construction of “factual” history is therefore 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. If we know 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.<sup>21</sup>

\

### **WHY COUNTERFACTUALS?**

Counterfactuals can combat the deeply rooted human propensity to see the future as more contingent than the past, reveal contradictions in our 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:** There is a strong bias in the scholarly community against the contingency of important historical events. Two distinct but reinforcing processes account may be at work here. R. H. Tawney observed that historians give “an appearance of inevitableness” to an existing order by dragging into prominence the forces which have

triumphed and thrusting into the background those which they have swallowed up.”<sup>22</sup> According to Charles Tilly, inevitability is an artifact of the “retrospective” approach that examines phenomena by working backwards from them to discover their causes. When applied to the study of state formation it ignores the “hundreds of states that once flourished but then disappeared.” For this reason, Tilly proposes the “prospective” method to combat this bias; researchers would begin at the formative stages in history and search forward for alternative paths and outcomes.<sup>23</sup>

Historical interpretations, no matter how well-documented or convincingly presented, rarely achieve a consensus. They provoke critiques and contending interpretations, and the resulting controversy stimulates a search for additional documents. New evidence may require historians to reformulate their interpretations but it rarely discredits them outright. As the historiography of the First World War indicates, new evidence often provides the basis for still more interpretations. Decades of controversy, or even centuries of controversy about the French Revolution or the collapse of the Roman Empire, give rise to a plethora of interpretations that attribute these developments to a wide range of intellectual, social, economic and political causes.<sup>24</sup> Remove one or several putative causes and a half-dozen others remain. Historical debate encourages the belief that most major events and developments were massively “overdetermined.”<sup>25</sup>

The disciplinary tendency to privilege structural explanations is reinforced by the “certainty of hindsight bias.”<sup>26</sup> Baruch Fischhoff 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 (e.g., peace

between Egypt and Israel, the end of the Cold War), are often considered “overdetermined” and all but inevitable after they have occurred.<sup>27</sup> 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 eliminate. 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, Lepper, Strack and Steinmetz exploited the tendency of people to inflate the perceived likelihood of vivid scenarios to make them more responsive to contingency. People they 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.<sup>28</sup> Philip E. Tetlock and the author conducted a 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 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 three arguments for why each of them was

plausible. Judgments of contingency varied in proportion to the degree of counterfactual unpacking.<sup>29</sup> 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 that establishes what is normal. In the hard sciences, benchmarks can sometimes be derived from a well-established laws or statistical generalizations; cold fusion would have been contrary to several of these laws and thus a truly anomalous event. There are few laws or statistical generalizations applicable to the social world, but we all hold to theories, or at least strong views, about how that 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 had avoided war, unlike rival hegemonies of the past. 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.<sup>30</sup> Both 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.

From social science and history alike, the West has too often been the starting point for our examination of the rest of the world. To return to our example of state formation, early Modern Europe was composed of multiple political units, many with internal checks, something many Western scholars tend to see as a highly likely, if not inevitable outcome.<sup>31</sup> China developed into a unitary empire, and during periods of fragmentation, a universal empire was still recognized as the norm. Western scholars have sought to explain why China evolved differently. Not surprisingly, many historians in China assume that the development of a universal empire was inevitable, and that the Chinese experience is the norm.<sup>32</sup> Chang Kwang-chih suggests that it is time to consider building theories of history on the Chinese experience, whose history and sources are “as formidable and massive as Western history.”<sup>33</sup>

All these examples suggest how ideology and culture provide the unspoken assumptions that shape our understanding of social reality. These understandings generate expectations, which if they are not met, whether in the future or the past, are regarded as anomalies and perhaps, interesting research questions. Ideological and cultural biases readily assume the guise of science and lend themselves to tautological confirmation. This only tightens their hold on the minds of scholars and public opinion. Counterfactuals are a useful means of exposing the subjective foundations of these beliefs and providing a possible means for scholars to step outside their own cultures and belief structures. Suppose we consider if there would have been a Cold War if the Soviet Union had been a liberal democracy in 1945, or if there had been no nuclear weapons? What if China had accessible coal reserves and iron deposits close to major

commercial centers? These counterfactuals are entirely unrealistic but they compel us to examine the unspoken assumptions that guide our expectations. Perhaps the most imaginative use of counterfactuals for this purpose has been by Stanislaw Lem who has published a book of reviews of books that were never written!

**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 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.<sup>34</sup> James Fearon rightly observes that this strategy is only effective 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.<sup>35</sup> 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, 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. This is difficult to do because it requires extraordinary breadth on the part of the researcher, or a collaborative enterprise.

Within a single case comparative analysis is also possible and can take two forms: intra-case comparison and counterfactual analysis. Intra-case 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.<sup>36</sup> Like any form of comparative analysis, intra-case comparisons try to show as much variation as possible on dependent (*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 intra-case comparison confers a singular benefit: it builds variation within a fundamentally similar political and cultural context, controlling better than inter-case comparison for many factors that may be important but otherwise unrecognized. Unfortunately, not every case can be broken down into multiple decision points for purposes of comparison.

When intra-case 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 identify key variables and the range of values in which they will have the most impact on the outcome. Information obtained this way, especially if it has counter-intuitive 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. Lars-Eric Cederman used this method to test the realist assumption that balancing is the inevitable consequence of international anarchy. It follows, Cederman reasoned, that global hegemons should rarely emerge in real or counterfactual worlds. However, he found that they appear with regularity in counterfactual simulations, especially under conditions of defense-dominance, the best case for neo-realism.<sup>37</sup>

Counterfactual experiments and simulations can tease out the assumptions -- often unarticulated -- on which a theories and historical interpretations rest.<sup>38</sup> Apologists for the Soviet system insist that communism would have evolved differently if Lenin had lived longer or had been succeeded by someone other than Stalin.<sup>39</sup> 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 two generic strategies open to them. They can offer a different and more compelling theory or interpretation — far and away the most common strategy -- or 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, by memories of World War II and the human, economic and social costs of large-scale, conventional warfare. He contends that the unheralded destructiveness of nuclear weapons was redundant and possibly counter-productive.<sup>40</sup> 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 human kind? 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: they use 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 they 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 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.<sup>41</sup>

An example of the latter is provided by Alessandro Portelli in his study of Italian communists. In interviews he conducted in the small city of Terni in Umbria he found that workers often told stories of how history *could* and *should* have gone. These narratives describe a better world that workers believe might have come to pass if their leadership had caused specific events to occur different between 1991 and 1925 and again between 1943 and 1953.<sup>42</sup>

**Must counterfactuals be realistic?** The several uses of counterfactuals I have described use “plausible” and “miracle” world of 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 significance 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.<sup>43</sup> 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 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.<sup>44</sup>

“Miracle” counterfactuals violate our understanding of what is plausible or even possible.<sup>45</sup> Take the following claim: “If Bosnians had been blue-nosed dolphins, NATO would not have allowed their slaughter.”<sup>46</sup> Even Dr. Seuss might find a landlocked province of aquatic mammals with strong ethnic identities something of a stretch, but the analogy suggests a series of morally provocative counterfactuals. If Bosnians had were Roman Catholics or Protestants — and not Muslims — would NATO have intervened sooner? Miracle counterfactuals are useful for purposes of theory building and testing. If I hypothesize that Europe achieved its military advantage because it was the only region of the world where no long-standing hegemony was established, and 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.<sup>47</sup> To advance this hypothesis I must have considered 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 consider what a hegemonic Europe would have been like, and how it would have differed from historical Europe. Miracle counterfactuals are particularly useful in evaluating existing interpretations.

John 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 utility of miracle counterfactuals does not depend on their realism, but on 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.<sup>48</sup>

We might also distinguish between scholarly and folk counterfactuals. The former are those we invent and use in attached experimental way for any of the research purposes I have described. They may be plausible or miracle world counterfactuals. Folk counterfactuals are embedded in popular culture and often influence how people frame and respond to problems. They may provide pithy and powerful statements of widely-held beliefs that mobilize people for action. I recently heard a prominent religious leader proclaim on a TV talk show that the Columbine shootings would never have happened if Hollywood and the other media had not produced so many violent movies, television shows, popular music and video games. This is a good representative of the "where did we go wrong?" question, and assumes a prior turning point with a better counterfactual outcome if only we had followed it. Regret and desire appear to be the principal motivating agents of such counterfactuals, and the latter may reach its apotheosis in

Andrew Marvel's immortal lines: "Had we but World enough and time, This coyness, Lady, were no crime. . . ." <sup>49</sup>

Scholarly and counterfactuals may interact. Blaise Pascal invented the counterfactual of Cleopatra's nose to illustrate the power of human vanity, and it has since become folk wisdom. <sup>50</sup> A volume of *Astérix* is devoted to Cleopatra and her nose, and there is even a Turkish saying, "to have a nose like Cleopatra's." <sup>51</sup> The counterfactual sensitizes and socializes listeners to a particular causal connection that we can state in the form of an English-language proverb: love makes the world go round. I noted earlier that the strategy of deterrence rests on the counterfactual that Hitler might have been overthrown, or at least made more cautious, if the democracies had stood firm earlier in the 1930s. This counterfactual belief not only shaped postwar American security policy, but became so entrenched that in instances when deterrence failed scholars and politicians did not blame the strategy but those who had carried it out. John F. Kennedy's threats failed to deter a Soviet missile deployment in Cuba, and academic and government analysts subsequently reasoned that deterrence would have worked if the president had not been too young for Khrushchev to take seriously, had committed troops to the faltering Bay of Pigs invasion, behaved more forcefully at the June 1961 Vienna summit and during the Berlin crisis that followed, and had put Khrushchev on notice several months earlier that offensive weapons in Cuba were unacceptable. <sup>52</sup> The absence of evidence, then or now, for these counterfactuals has not dissuaded many scholars or members of the foreign policy elite of their validity.

The discussion of scholarly and folk counterfactuals points to a broader truth about the relationship between counterfactuals and "factual" history. History consists of acts of labeling,

and its core activity is framing. Straight history and counterfactual history, popular and academic, are rhetorical strategies to get people to view the world and critical aspects of the present in particular ways. Counterfactual history can be a particularly powerful way of achieving this end, and even those who do not accept its status as a scientific research tool can not afford to ignore it.

### **HISTORICAL COUNTERFACTUALS**

Two recent publications by Niall Ferguson have sparked renewed interest in counterfactuals in the history discipline.<sup>53</sup> Unfortunately, neither work makes a good case for counterfactuals. In a long introduction to his edited volume, *Virtual History* Ferguson offers no reasons for engaging in counterfactual history other than sensitizing readers to contingency, and only briefly addresses methods for conducting counterfactual experiments. He 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.<sup>54</sup> 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 invasion fleet had not encountered a destructive typhoon en route to Japan, if the Duke of Alba had not fallen sick 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 only resisted 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.<sup>55</sup> 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.<sup>56</sup> 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, multi-step, 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 contrived. There is no particular reason why Cleopatra should have been less attractive or Anthony 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 into a straight jacket. He insists that we only consider 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.<sup>57</sup> Ferguson's criteria would exclude entire categories of plausible world counterfactuals. It would limit counterfactuals to elites who made written records, to contemplative 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); human accident, oversight, obtuseness or unanticipated error; acts of nature or the confluence (or lack of it) of independent chains of causation. We could not contemplate a world in which the Duke of Alba had remained healthy and not relinquished command of his army to his son, Franz Ferdinand's touring car adhered to its planned route, Hitler had died in an automobile accident or in which Hinckley had assassinated Ronald Reagan. Ferguson's criteria would also rule out all miracle counterfactuals.

### **“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 will lead to the hypothesized consequent. There are three reasons for this well-warranted pessimism: the statistical improbability of multi-step 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 neither World War nor the Holocaust would have occurred if Mozart had lived to the age of sixty-five.<sup>58</sup> Having pushed classical form as far as it could go in the Jupiter Symphony, Mozart's last three operas and the Requiem, his next dramatic works would have been the precursors of a new, "post-classicist" style. He would have created a viable alternative to romanticism that would have been widely imitated by composers, writers and artists. Post-classicism would have kept enlightenment political ideas 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, 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 counterfactual presupposes numerous other enabling counterfactuals about the nature of the political changes that will lead to the hypothesized

consequent (e.g., 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 fifty 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. John 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.<sup>59</sup> 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 to render the consequent moot or 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 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 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.<sup>60</sup> 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 over-extended empire. In relatively short order, they believe, England would have overthrown the Spanish yoke.<sup>61</sup>

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 dampened down over time and end up having little lasting effect. It is also possible that second counterfactuals arising from a Spanish occupation of England, however brief that occupation, could have had dramatic and long-lasting changes for European politics or in other realms that Martin and Parker did not consider. None of these outcomes are predictable, and the butterfly effect may not be knowable in advance or even in retrospect.<sup>62</sup>

**Criteria for use:** Recognition that counterfactual arguments often have indeterminate consequences has prompted scholars to impose restrictive criteria on their use. James 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.”<sup>63</sup> Robyn Dawes argues that counterfactual inferences are only warranted “if and only if they are embedded in a system of statistical contingency for which we have reasonable evidence.”<sup>64</sup> Edgar Kiser and Margaret Levi suggest

that counterfactuals are best used as substitutes for direct empirical analysis when data is limited or unavailable. Such counterfactuals should be based on a general deductive theory with clear micro-foundations and scope conditions.<sup>65</sup> Jon 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.”<sup>66</sup>

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. Steven Weber rightly observes that “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 requires 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.”<sup>67</sup> 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.

The scholars most troubled by the inherent unpredictability of counterfactual outcomes are those who want to use counterfactuals to test propositions and theories. The less demonstrable the consequent, the less useful counterfactual is for this purpose. Fearon is willing to consider a

fall-back 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.<sup>68</sup> The author employs this strategy in his chapter on the origins of World War I to show why that conflict was contingent in a double sense. It was the result of three largely independent chains of causation that produced non-linear effects only when they came into confluence. War was also dependent on a special kind of catalyst, whose presence was problematic and whose causes were independent of any of the underlying causes of war. Both the confluence and the catalyst could easily have been prevented by minimal rewrite counterfactuals.<sup>69</sup>

Scientific validation of a counterfactual is only important 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 neo-positivist tradition, the prior steps of identifying important questions or anomalies and formulating theories to explain or resolve them are generally recognized as critical. Counterfactuals can 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. Social science and history ultimately aim to broaden our intellectual horizons and to provide insight into contemporary problems and policy choices. Counterfactual experimentation is essential to these tasks, and can be used effectively without the same degree of confidence 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 data points in explanation.”<sup>70</sup> We should worry less about the uncertainty of counterfactual experimentation and think more from 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 nine criteria for plausible world counterfactuals.

Numbers 2,3, 5 and 6 are drawn from or variants of the Tetlock-Belkin list, number 6 has been recently proposed by Tetlock, and numbers 1, 4, 8 and 9 are the author's.<sup>71</sup>

1. **Realism:** Ferguson rightly ridiculed, although for the wrong reason, Pascal's counterfactual of Cleopatra's nose. It is objectionable because it is arbitrary and contrived. There is no particular reason why Cleopatra should have been less attractive or why Anthony should have a different sexual orientation. Nor is it clear how we could bring these changes about. Good counterfactuals must arise from the context, and we must have compelling mechanisms to bring them into being that themselves require only minimal rewrites of history.

2. **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 European communities to have implemented these measures, they would have had to recognize that human intervention could affect the spread of disease, and had the authority and will to impose draconian measures on travel and trade over the likely objections of the wealthy and merchant classes.<sup>72</sup> 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.<sup>73</sup> Plausible world counterfactuals require not only realistic antecedents, 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 the 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.<sup>74</sup> John Elster has rightly objected that if the technology was present to invent and produce automobiles, it would almost certainly have also led to the development of railroads.<sup>75</sup>

3. **Enabling counterfactuals should not undercut the antecedent:** Counterfactuals may require other counterfactuals to make them possible (e.g., for Richard Nixon to have been president at the time of the Cuban missile crisis he would have 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.<sup>76</sup>

Counterfactual arguments that make a credible case for a dramatically 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 *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. According to the laws of statistics, 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 future. The long-term consequences of change are nevertheless 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 characteristic 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 to varying degrees shift their location and tension.<sup>77</sup> 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 Richard 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 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, or 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 develops.

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 “alternative” 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, 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 a such counterfactuals 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.

## UNMAKING THE WEST

Having developed appropriate protocols for the use of counterfactual thought experiments let us now turn to the subject of this volume: the development of Western civilization and its remarkable success in the modern era. How contingent were both developments? Could Western civilization or its political, economic, technological and military primacy have been prevented by minimal rewrites of history? At what point did the success of the West become difficult to prevent? When did it become impossible to prevent? And what about the parallel development of other competitor civilizations — China, India, the Arabs or the Ottomans? Could minimal rewrites of their histories have made them more successful competitors with the West? At what point did that become unlikely or impossible?

To answer these questions, our authors identified some critical turning points of Western and non-Western history and seek to probe their contingency through counterfactual experiments. To evaluate their individual efforts and our broader collective goals, we need protocols for assessing the contingency of each turning point and the aggregate implications of these experiments for the course of Western history. Toward this end in mind, I propose six tests, all of which take the form of questions:

1. **What do we have to do to negate a turning point?** Turning points are important because of their consequences. Constantine's embrace of Christianity and victory at Milvian Bridge in 312 and subsequent victories over Licinius in 323-24 made him sole ruler of the Roman Empire and able to impose Christianity on the Empire and repress "heretical" sects. A common

religion and political tradition may have been essential for the long-term development of Western civilization. Had Constantine been decisively defeated in any of these campaigns, the religious life of Rome would have developed differently and perhaps removed one of the foundation stones of Western culture.

Some of the events that we consider turning points may not need to be prevented. It may only be necessary to alter them in minor ways that render them innocuous much the way small mutations in the genes of pathogens may decisively alter their virulence, one way or the other. Let us suppose we allow the Greeks to win their crushing naval victory at Salamis — the turning point addressed by Hanson and Strauss in their point and counterpoint chapters. But introduce a small change in the events that followed. Sparta was asked to name a general to lead the Greek alliance against the Persians, and nominated their King Pausanias for the position. They recalled him about a year later, tried him on charges of corruption and decided to withdraw from the alliance for fear that exposure to the life of richer city states was too threatening to their way of life. Sparta's withdrawal opened the way for Athens to assume leadership of the alliance, and ultimately to transform it into a powerful empire. A different and uncorrupt Spartan leader might have forestalled Athenian leadership of the Delian League, and with it the wealth and self-confidence that produced the fifth century cultural explosion that is widely acknowledged as another foundational stone of Western civilization.

Mere postponement of a turning point may sometimes deprive it of its most important effects. In my chapter on the origins of the First World War, I argue that it, or any conflict like it, might have been prevented if Franz Ferdinand's assassination could have been forestalled for as little as three years. The underlying conditions that moved Europe toward war were rapidly

evolving, and any one of a number of changes would have made Austria, Germany or Russia — or all three — war averse. In the interim, it is unlikely that any other event could have served as catalyst for war because Sarajevo met an unusual set of conditions without which neither Austrian or Germans would have been prepared to risk war.

Our first task, therefore, is to determine what kinds of changes we have to make in the events identified as turning points. Once we know whether they need to be prevented, altered or merely forestalled, we can search for minimal rewrites capable of achieving these ends.

**2. How many credible minimal rewrites can be found that might prevent, alter or stall the turning point?** Earlier I noted how easy it would have been to avert Franz Ferdinand's assassination if only his cavalcade had followed the planned route. Numerous other minimal rewrites could also have kept the archduke and his wife alive. Princip might have obeyed the order to abort the assassination sent to him by the military conspirators in Belgrade; Austrian authorities in Bosnia might have taken security as seriously as they did the menu and music for the banquets they planned in the archduke's honor; Franz Ferdinand might have canceled his trip in response to multiple warnings and his wife's fears; he might have followed the advice of his advisors and left Sarajevo directly after the ceremony at city hall or have raced down Appel Quay past Princip. These minimal rewrites suggest contingency by virtue of their variety. The first counterfactual removes the assassins, two others remove the target, and two more leave him on the scene but make him much more difficult to kill. As a general rule, the more different components of a turning point that can be removed by minimal rewrite counterfactuals, the more contingent the turning point.

**3. How far back must we go to find credible minimal rewrites?** We should try to

avoid the conjunction fallacy and the best way to do this is to keep the intervening step between antecedent and consequent to a minimum. Temporal proximity, while not the same thing, is a good rough measure. Most of the minimal rewrites I suggest in the example above meet this test. They would have prevented Franz Ferdinand's assassination with few intervening steps, and this is one reason why they appear plausible. The further back in time we go to find a minimal rewrite counterfactuals, the more steps there are likely to be between antecedent and consequent, and we must remember that the probability of any counterfactual is the multiple of the probabilities of every step in the chain.

Minimal rewrite counterfactuals at temporal remove invite a second problem: they allow more parallel chains of causation, one or more of which may lead by alternative routes to the outcome the minimal rewrite is intended to prevent. Early public cancellation of the archduke's visit to Sarajevo might have encouraged dedicated assassins to find some way of getting close him during the army maneuvers scheduled outside the city. This would have been difficult but not impossible given the poor state of Austrian military security.

4. **At what level of analysis are our minimal rewrites?** Minimal rewrites of history require small, plausible changes in reality that are likely to have big consequences. Surgical interventions at close proximity to the hypothesized consequent seriously constrain the choice of counterfactuals. For these reasons our authors and other practitioners of counterfactual history most often invoke changes in personnel due to assassinations (Hinkley kills Reagan), failed assassinations (Franz Ferdinand lives), the fortunes of war (William of Orange and Hitler die in battle), disease (Pericles survives the plague) and accidents (Alexander the Great lives to a ripe old age. Equally popular are changes in policy where the chosen course of action was a near

thing (Greek naval strategy at Salamis) where another strategy appeared just as feasible (the Ottomans make an all-out effort to capture Vienna) or unintended deviations from plans arising from what Carl von Clausewitz called “friction” -- the “countless minor incidents” including undelivered, garbled or misinterpreted messages, the failure of key players or equipment to appear on time and a host of other problems that degrade performance.<sup>78</sup> Other minimal rewrites rely on changes in the weather (the Mongol invasion fleet has a smooth crossing to Japan), timing (British forces advance in a more timely way capture George Washington and most of his army before they can flee Long Island), or combinations of these changes.

To help assess contingency it is useful to introduce the concept of level of analysis. This is where we look for explanations -- or counterfactuals in this instance --- and it may be a system, unit, sub-unit or individual levels, or the environment in which the system functions. The latter may be the easiest place to look for minimal rewrite counterfactuals. It includes such things as the weather, timing and much of what comes under the rubric of friction. This is followed by the level of leaders, and here too, there are often relatively simple and plausible ways of changing, or at least removing, critical personnel. Almost all of the counterfactuals in the paragraph above are at one of these two levels. Changes in strategies may require intervention at the level of elites, bureaucracies or domestic politics, and more elaborate arguments linking antecedents to consequences; Richard Bulliet invokes a combination of such factors to arrange for the Turkish capture of Vienna.

When we change ideas, state structures and the balance of power, the latter requiring intervention at the system level, minimal rewrites are out of the question unless we go back to a point in time when those ideas, structures and balances had not jelled and might be significantly

affected by small, plausible changes at levels one and two. We could readily imagine a very different balance of power and set of alliances in seventeenth or eighteenth century Europe if we prevented or rearranged some of the marriages that Habsburgs, Bourbons and other royal families used to extend or consolidate their influence. But the further back we introduce changes, the less plausible the consequent because our antecedent is likely to introduce other changes with unknown interaction effects and consequences. So for our purposes, a turning point should be considered contingent to the extent that it can be untracked by numerous, plausible counterfactuals at levels one and two.

5. **How redundant is the turning point?** Different causal paths have different implications for contingency. A turning point described by the simple, linear pathway of  $A \rightarrow B \rightarrow C$  (and only  $A \rightarrow B \rightarrow C$ ) might be prevented by severing any link in the three step chain. If A, B are themselves the products of other chains of causation, there may be many possible way of using minimal rewrite counterfactuals to prevent the turning point by preventing preconditions A and B. There are few turnings that are the outcomes of simple linear chains. Others are likely to have one or more paths leading to them. If we prevent  $A \rightarrow B \rightarrow C$ , the possibility  $G \rightarrow H \rightarrow C$ , and perhaps of  $M \rightarrow N \rightarrow C$ , remain. To prevent turning point with we need to know all the chains leading to it something about their probability. In situations of equifinality, where multiple paths lead to the same outcome, estimates of probability are likely to be complicated by interaction effects. The removal of any causal chain may significantly change the probability (in either direction) of other paths. Death is an extreme example of this phenomenon. If we eradicate child diseases, or inoculate people against them, reduce the incidence of other infectious diseases through public health measures, and gain the upper hand in the struggles against heart

disease and cancer, the likelihood of dying from other causes (e.g., accidents, stroke, kidney failure), including new ailments (e.g., HIV) will increase. One way or the other, the grim reaper will come for all of us.

Even multiple chains may be inadequate to capture the causal complexity of some of the events identified by our authors as turning points. Another model to consider is the confluence, where a stream of independent causal chains come together to make an outcome possible. A house goes up in smoke. Investigation reveals that the fire spread from a lighted candle that was left unattended on a window sill. The window was not completely sealed, and a draft blew one of the curtains close enough to the flame for it to catch fire. The smoke alarm, connected to the house security system, did not function because its battery was dead, and the fire department failed to receive the timely warning that might have permitted it to save the dwelling. What caused the house to burn down? The candle was the source of the fire, but it would not have been lit or placed on the window sill if it had not been the holiday season and had its owners not been following a neighborhood custom. If the window had not been warped, or the insulation around it had provided a better seal, the candle would not have started a fire. If the owners had been home, or if smoke alarm had a charged battery, the house would not have burned down. No single factor was responsible for this disaster; it took a combination of them interacting in a particular way.<sup>79</sup> The Eire, Yates, Pomeranz and Mokyr chapters, all of which touch on the causes of the industrial revolution and why it began in Western Europe make implicit use of the concept of confluence. They do not attribute industrialization to any single cause, but to complex interactions among a confluence of factors, most or all of which were necessary for this turning point.

An outcome that requires the confluence of many independent causes, but could be

prevented, or transformed in magnitude, by removing any one of them with a minimal rewrite — like the fire in the house -- is highly contingent. But other confluences may have multiple pathways that lead to them — and require multiple interventions to prevent. Their contingency would depend on how many minimal rewrites were necessary to halt or deflect each possible pathway. Some of our contributors appear to suggest that Western industrialization was somewhat redundant in this way.

My chapter on the origins of World War I also makes use of confluence, but introduces the further complication of nonlinearity. In nonlinear systems, the value (or effect) of one variable (or cause) depends on the presence of another, making it impossible to understand the consequences of either in isolation.<sup>80</sup> In nonlinear systems, the rules that govern the system can change as actors transform themselves and their understandings of their strategic interactions in the course of those interactions. I contend that World War I is best understood as a nonlinear confluence of three separate chains of causation. These chains produced independent but more or less simultaneous *gestalt* shifts in Vienna and Berlin, and a slightly earlier one in Russia. Had the timing of the Austrian and German shifts been off by as little as two years, Austrian leaders would not have felt so intent on destroying Serbia, and German leaders would not have been so willing to encourage them to do so. For this reason alone, World War I was highly contingent.

6. **What about second order counterfactuals?** Up to this point we have tried to prevent, alter or stall turning points. But we must also think about how they might come to pass despite our best efforts. Second order counterfactuals, either by themselves or in interaction with one another, might produce the turning point, or some variant of it at a later date. Ross Hassig reasons that in some circumstances Cortés' defeat might have served as the catalyst for a renewed

Spanish attempt at conquest. If disease had decimated the Aztec population, the Spanish might have succeeded the second time around, and the history of Meso-America might have evolved along strikingly similar lines.

There are two other routes to turning points via the counterfactuals we have created to prevent them. The enabling counterfactuals necessary to bring about the antecedent can set in motion a chain of events that leads to the turning point, as can events arising from the antecedent itself [see figure 1]. A rigorous attempt to assess contingency compels us to work through these possible pathways in search of the most likely ways that turning points could still come about. If secondary routes to turning points can be found, researchers need to find ways to stop them with additional counterfactuals. Roughly speaking, the more alternatives that are found, and the more likely they appear, the less contingent the turning point.

In this volume, we break down the creation and rise of the West into a set of smaller, more manageable problems through the use of turning points. By factoring our two big questions this way we allow a more fine-grained analysis of key components of a causal chain. When we work back to the larger picture we confront the same set of problems we did with the individual turning points. There is no reason to suppose that Western civilization and its remarkable success in the modern era was the result of a single causal chain and could be stopped by preventing any of its component turning points? It seems likely that outcomes so complex and with so many attributes were the result of multiple chains and are at least in part equifinal. Barry Strauss appears to subscribe to this position; he argues that Western civilization would still have developed even if Persia had occupied all of Greece. In the final chapter of this volume, Geoffrey Parker will consider this possibility of multiple paths and equifinality for the subsequent success of the West.

The inescapable conclusion is that counterfactuals are like fractals. The defining characteristic of a fractal is that its structure is the same at every level of magnification. So it is with counterfactuals. Enabling counterfactuals, turning point counterfactuals and the use of turning points to prevent a larger outcome involve the same set of problems and require the same set of procedures to address them. And like fractals, the chain of connections could be extended further (but not infinitely as in the case of fractals) in either the micro or macro direction. We could look at levels beneath enabling counterfactuals, or beyond the success of the West if we made it an antecedent to some other consequent.

### **KASPAROV VS. DEEP BLUE**

Historical counterfactual research confronts two fundamental problems that can often be solved or finessed in engineering and the sciences: there is no rigorous way to establish probability, or reasonably to track the long-term consequences of counterfactuals.

Through this chapter I have made casual use of the concept of probability, largely for heuristic purposes. Probability is the frequency of a particular outcome when all outcomes are known. This can be determined logically, as in the case of a coin toss. If there are no imperfections in the coin and the height and trajectory of each toss varies randomly, the probability of either a head or a tail should be fifty percent. Probability can also be determined empirically; we can observe a large number of the same class of event and keep track of the frequency of different outcomes. If the sample is large enough, and the events fully comparable but independent, we can make a good estimate of the probability of any given outcome. Empirical tests of this kind can sometimes be carried out in laboratories, or in real world in the

case of some kinds of iterative processes. Simulations are sometimes also possible. For better or worse, history is a one time event. The tape of time cannot be rewound and run again and again. The best we can do is assess probability and contingency by means of counterfactual experimentation, and here we encounter a catch-22. Estimates of contingency require some knowledge of the range of possible outcomes, real and counterfactual, and of the probability of each. But probability can not be measured; if we could form good estimates we would not need counterfactual experiments.

The protocols for counterfactual experimentation and those for assessing the contingency of turning points put extraordinary demands on researchers. They must determine the best strategy for negating a turning point, and this involves consideration of numerous possible counterfactuals. Even the most promising counterfactuals may require additional, enabling counterfactuals, and researchers must then consider the most important consequences of all of these counterfactuals. They must also ask if the turning point could be produced by some other set of causal chains and, if so, find ways of preventing this from happening. They must then direct their attention to second order counterfactuals, and search for and untrack any that might lead to the turning point at some future time. Thorough execution of each of these tasks would make for more convincing counterfactuals but it is simply not feasible given the time and effort it would take.

In February 1996 and May 1997 world chess champion Gary Kasparov played two matches against Deep Blue, a bank of refrigerator size IBM SP 2 computers. Kasparov won the first match 4-2, but lost the second, 3.5-2.5 in what was hailed as a dramatic victory for artificial intelligence. The idea of a chess playing machine dates back to the 1760s, when Wolfgang von

Kempelen exhibited the Maelzel Chess Automaton throughout Europe. His machine infuriated Napoleon when it beat him in nineteen moves. The emperor never guessed the nature of its mechanism: a diminutive Turkish chess master hidden in a cabinet inside who operated the turbaned and mustachioed marionette who moved the pieces on the board above. Modern computers rely on elaborate algorithms to identify possible moves and the most likely responses of opponents, work through each line of play for a given number of turns and evaluate all the outcome to reach an informed choice.<sup>81</sup>

The principal advantage that all computers have over people is their awesome ability to crunch numbers. Deep Blue works through a billion positions a second and this is why it can explore so many variations over so many turns. Counterfactual research in history demands a similar search pattern to chess. Historians must identify the most promising “moves,” look at how reality might respond and how the interaction of counterfactuals and existing context might play out even beyond the hypothesized consequent. The game of counterfactuals may even be more complicated than chess. In the latter, two players engage in move and countermove. Each new ply in the branching tree of analysis encompasses roughly thirty-eight times as many positions. In counterfactuals, there can be *multiple* interactions at every move, as antecedents, enabling counterfactuals, and the chains they set in motion interact with one another. In chess, the definition of a move is clear; each player moves a piece in turn. In counterfactuals, chains of consequence unfold simultaneously and continually.

Supercomputers are routinely used to conduct counterfactual simulations in the sciences. Until such time as the programmers of Deep Blue turn their attention to historical counterfactuals, historians must behave more like chess masters than computers. Gary Kasparov has the merest

fraction of the processing speed and capability of the simplest computer, but beat Deep Blue in their first tournament and lost the second by only one game. The most challenging task in chess is determination of which of the multitudinous branching points are most worth examining because no computer can analyze them all. Deep Blue was an improvement on its predecessors, and made use of more sophisticated tuning mechanisms to select promising lines of play. Human beings, especially chess masters, still have a considerable comparative advantage in this respect. They use their intuition and judgment to select one or two possible moves and project their likely consequences out to a limit of about eight turns ahead. At each turn, they are ruthless in eliminating possibilities, and thus stay within the limiting data processing capabilities of the human mind by tracking only a few chains of moves. The best computer algorithm is crude in comparison to the still uncanny human ability to zero in on promising lines of play.

Historians have no recourse but to follow the strategy of Gary Kasparov and substitute judgment for computing power and limited for more comprehensive search. And this is what our contributors do; they use their imagination to invent counterfactuals they think capable of preventing or adding turning points, and their professional judgment to select a few chains of counterfactual interaction to analyze in detail and estimate the probabilities associated with each.

In chess, there is a straight-forward way to estimate performance: players win, lose or draw their games. In counterfactual research this is not possible. The best players can do is to be persuasive by making absolutely explicit the logic connecting antecedent and consequent and the assumptions on which all the chains of causation they consider are based. Such an exercise makes no pretense to be a science, only to harness human intelligence with the goal of making

informed individual and collective judgments about interesting and important problems. This is no different than the claims made by more conventional forms of history.

## REFERENCES

- 
1. Examples from physics include, R. Penrose, *Shadows of the Mind: A Search for the Missing Science of Consciousness* (Oxford: Oxford University Press, 1994); A. C. Elitzur and L. Vaidman, "Quantum-Mechanical Interaction-Free Measurement," *Foundations of Physics* 23, no. 7 (1993), pp. 987-97, which use information from non-events to test nuclear weapons; P. G. Kwiat, H. Weinfurter, H. Herzog, A. Zeilinger and M. A. Kasevich, "Interaction Free Measurements," *Physics Review Letters* 74, no. 12 (12 June 1995), pp. 4763-66, which uses interaction free measurements with a test particle to determine the presence of an object; Graeme Mitchison and Richard Josza, "Counterfactual Computation," *Quantum-Ph/9907007*, 2 July 1999, on counterfactual computation.
  2. Electoral studies are a case in point. See Raymond E. Wolfinger and Benjamin Highton, "Can More Efficient Purging Boost Turnout," Paper presented at the 1994 annual meeting of the American Political Science Association; 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. Timpane, "Past History Reverses: Counterfactuals and the Impact of Registration Reform on Participation and Representation," Unpublished paper.
  3. According to A. J. P. Taylor, "a historian should never deal in speculation about what did not happen." *The Struggle for the Mastery in Europe, 1848-1918* (London: Oxford University Press, 1954); E. H. Carr, *What is History?* (London : Macmillan, 1961), p. 127; E. P. Thompson, *The*

---

*Poverty of Theory and Other Essays* (New York: Monthly Review Press, 1978), p. 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*, 2nd ser. 22 (August 1969), pp. 195-207; See also David Hackett Fischer, *Historians’ Fallacies* (New York: Harper Colophon Books, 1970), pp. 15-21; Peter McClelland, *Casual Explanation and Model-Building in History, Economics, and the New Economic History* (Ithaca: Cornell University Press, 1975).

4. Private communication to the author.

5. *New York Times*, 27 November 1999, p. A1.

6. 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), pp. 95-118.

7. Richard Ned Lebow and Janice Gross Stein, “Back to the Past: Counterfactuals and the Cuban Missile Crisis,” in Tetlock and Belkin, *Counterfactual Thought Experiments in World Politics*, pp. 119-148.

8. Jay M. Winter, *The Great War and the British People* (London: Macmillan, 1986), 76-83.

9. Charles Tilly, “Westphalia and China,” unpublished manuscript, Columbia University.

10. Tin-bor Victoria Hui, *Rethinking War, State Formation, and System Formation: A Historical Comparison of Ancient China (659-221 BC) and Early Modern Europe (1495-1815 AD)*, Ph. D. dissertation, Columbia University, 2000. See also R. Bin Wong, *China Transformed: Historical Change and the Limits of the European Experience* (Ithaca: Cornell University Press, 1997).

- 
11. Kenneth Pomeranz, "China, Europe and Industrialization."
  12. Jon Elster, *Political Psychology* (New York: Cambridge University Press, 1993), p. 5; Sidney Tarrow, "Expanding Paired Comparison: A Modest Proposal," American Political Science Association, *Comparative Politics Newsletter*, (Summer 1999), pp. 9-12; Doug McAdam, Sidney Tarrow and Charles Tilly, "Comparisons, Mechanisms and Episodes," in *Dynamics of Contention*, forthcoming.
  13. Donald Kagan, *The Outbreak of the Peloponnesian War* (Ithaca: Cornell University Press, 1969), pp. 203-342.
  14. Richard Ned Lebow and Janice Gross Stein, *We All Lost the Cold War* (Princeton: Princeton University Press, 1994).
  15. Mark Elvin, *The Pattern of the Chinese Past: A Social and Economic Interpretation* (Stanford: Stanford University Press, 1973).
  16. George Breslauer and Richard Ned Lebow, "Leadership and the End of the Cold War: A Counterfactual Thought Experiment," to appear in Richard Herrmann and Richard Ned Lebow, eds., *Learning from the Cold War*, 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.
  17. E. H. Carr, *Socialism in One Country, 1924-1926*, 3 vols. (New York: Macmillan, 1958-

---

1964), vol. 1, p. 151.

18. Lebow and Stein, *We All Lost the Cold War*, ch. 4 reviews this argument and the new evidence.

19. John R. Searle, *The Construction of Social Reality* (New York: Free Press, 1995).

20. Donald Davidson and Jaakko Hintikka, eds., *Words and objections. Essays on the work of W. V. Quine*. (Dordrecht, D. Reidel, 1969).

21. 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 (1998), pp. 230-42

22. R. H. Tawney, *The Agrarian Problem in the Sixteenth Century* (London: Longmans, Green, 1912), p. 177.

23. Charles Tilly, ed., *The Formation of National States in Western Europe* (Princeton, N.J. : Princeton University Press, 1975), pp. 9, 14-15.

24. Alexander Demandt, *Der Fall Roms: die Auflösung des römischen Reiches im Urteil der Nachwelt* (Munich: Beck, 1984), p. 695, found two hundred and ten causes for the fall of Rome in the historical literature since 1600.

25. Paul W. Schroeder, "World War I as Galloping Gertie: A Reply to Joachim Remak," *Journal of Modern History* 44 (September 1972), pp. 319-345, reasons: "The fact that so many plausible explanations for the outbreak of war have been advanced over the years indicates on the one hand that it was massively overdetermined."

26. 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, *Counterfactual Thought Experiments in World Politics*, pp. 15-16.

<sup>27</sup> Baruch Fischhoff, “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), pp. 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), pp. 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).

28. L. Ross, M. R. Lepper, F Strack and J. Steinmetz, “Social Explanation and Social Expectation: Effects of Real and Hypothetical Explanations on Subjective Likelihood,” *Journal of Personality and Social Psychology*, 35 (1977), pp. 817-29.

29. The first of these experiments, involving alternative outcomes for the Cuban missile crisis, is described in an as yet unpublished paper, Philip E. Tetlock and Richard Ned Lebow, “Poking Counterfactual Holes in Covering Laws: Alternative Histories of the Cuban Missile Crisis.”

30. 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), pp. 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), pp. 579-614.

31. Leften Stavros Stavrianos et al, *A Global History of Man* (Boston: Allyn and Bacon, 1966).

32. Hui, *Rethinking War, State Formation, and System Formation*, p. 7.

33. Kwang-chih Chang, *Art, Myth and Ritual: The Path to Political Authority in Ancient China* (Cambridge: Cambridge University Press, 1983), pp. 128-29.

34. Robyn M. Dawes, "Counterfactual Inferences as Instances of Statistical Inferences," in Tetlock and Belkin, *Counterfactual Experiments in World Politics*, pp. 301-08. 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.

35. Fearon, "Counterfactuals and Hypothesis Testing in Political Science."

36. 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, Co.: 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).

37. Lars-Eric Cederman, "Rerunning History: Counterfactuals Simulation in World Politics," in

---

Tetlock and Belkin, *Counterfactual Thought Experiments in World Politics*, pp. 247-67.

38. Fearon, "Counterfactuals and Hypothesis Testing in Political Science."

39. George W. Breslauer, "Counterfactuals Reasoning in Western Studies of Soviet Politics and Foreign Relations," in Tetlock and Belkin, *Counterfactual Thought Experiments in World Politics*, pp. 69-94, discusses this literature.

40. John Mueller, *Retreat from Doomsday: The Obsolescence of Major War* (New York: Basic Books, 1989) and 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," *International Security* 13 (Fall 1988), pp. 55-90.

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

42. Alessandro Portelli, "Uchronic Dreams: Working-Class Memory and Possible Worlds," in *The Death of Luigi Trastulli and Other Stories: Form and Meaning in Oral History* (Albany: State University of New York Press, 1991), pp. 99-116.

43. Richard Ned Lebow, "Franz Ferdinand Found Alive: World War I Unnecessary," to appear in Tetlock, Lebow and Parker, *Unmaking the West*.

44. John 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), pp. 31-60.

45. The term miracle counterfactual was coined by Fearon, “Causes and Counterfactuals in Social Science,” p. 60.

46. Tetlock and Belkin, *Counterfactual Thought Experiments in World Politics*, (fn 23), p. 14.

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

48. 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 a infinite universe -- neither we nor the earth would exist. Hence, the density of stars must decrease or the universe must be finite.

49. Andrew Marvell, “To His Coy Mistress,” Hug MacDonald, ed., *Andrew Marvell* (London: Routledge & Kegan Paul, 1952), pp. 21-22.

50. Francis Kaplan, ed., *Les Pensees de Pascal* (Paris: Les Éditions du Cerf, 1982), p. 209.

51. René de Doscigny and Albert de Uderzo, *Astérix et Cleopatre* (Neuilly-sur-Seine: Dargaud, 1965). Is it a coincidence that this publishing house is located at 12, rue Blaise-Pascal? The Turkish saying was brought to my attention by Prof. Orhan Tekelioglu of Bilkent University.

- 
52. Lebow and Stein, *We All Lost the Cold War*, ch. 4.
53. Niall Ferguson, ed. *Virtual History: Alternatives and Counterfactuals*. (New York: Basic Books, 1999) and *The Pity of War: Explaining World War I*. (New York: Basic Books, 1999).
54. Ferguson, ed. *Virtual History*, pp. 11-12.
55. 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, 1998), pp. 129-34; Geoffrey Parker, "The Repulse of the English Fireships," in Robert Cowley, ed., *What If?* (New York: G. P. Putnam's Sons, 1998), pp. 139-54.
56. Max Weber, "Objective Possibility and Adequate Causation in Historical Explanation," in *The Methodology of the Social Sciences* (Glencoe, Ill.: Free Press of Glencoe, 1949 [1905]), pp. 164-88.
57. P. Nash, "The Use of counterfactuals in history: A Look at the Literature," *The SHAFR Newsletter* (March 1991), pp. 2-12, offers a similarly restrictive definition.
58. Richard Ned Lebow, "If Mozart Had Died at Your Age?," to appear in Katherin Aaslestad and David Murphy, *Re-envisioning Power and Security: The Culture of Diplomacy, Essays in Honor of Paul W. Schroeder*.
59. Cederman, "Rerunning History," pp. 253, 255, for this criticism of Mueller.
60. Tetlock, "'Distinguishing Frivolous from Serious Counterfactuals.'"
61. Parker, *The Grand Strategy of Philip II*, pp. 281-96.

- 
62. M. Maruyama, "The Second Cybernetics: Deviation-amplifying Mutual Causal Processes," *American Scientist* 51, no. 2 (1963), pp. 164-79, 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.
63. Fearon, "Causes and Counterfactuals in Social Science," p. 50.
64. Dawes, "Counterfactual Inferences as Instances of Statistical Inferences."
65. Edgar Kiser and Margaret Levi, "Using Counterfactuals in Historical Analysis: Theories of Revolution," in Tetlock and Belkin, *Counterfactual Thought Experiments in World Politics*, pp. 187-210.
66. Elster, *Logic and Society*, pp. 184-85.
67. Weber, "Counterfactuals, Past and Future," pp. 278, 272.
68. Fearon, "Causes and Counterfactuals in Social Science," p. 54.
69. Richard Ned Lebow, "Frans Ferdinand Found Alive. World War I Unnecessary."
70. Steven Weber, "Counterfactuals, Past and Future," p. 270.
71. Tetlock and Belkin, "Counterfactual Thought Experiments in World Politics," pp. 16-32; Philip Tetlock, "'Distinguishing Frivolous from Serious Counterfactuals,'" unpublished paper; Lebow and Stein, "Back to the Past," pp. 146-47.
72. Hawthorn, *Plausible Worlds*, pp. 31-60.

- 
73. William H. McNeill, *The Rise of the West: A History of the Human Community* (Chicago: University of Chicago Press, 1970), p. \_\_\_.
74. Robert Fogel, *Railroads and American Economic Growth: Essays in Econometric History* (Baltimore: Johns Hopkins University Press, 1964).
75. John Elster, *Logic and Society: Contradictions and Possible Worlds* (New York: Wiley, 1978), pp. 204-08; Tetlock and Belkin, “Counterfactual Thought Experiments,” pp. 22-23.
76. Weber, “Objective Possibility and Adequate Causation in Historical Explanation.”
77. 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), p. 118.
78. Carl von Clausewitz, *On War*, trans. Michael Howard and Peter Paret (Princeton: Princeton University Press, 1976), pp. 119-21.
79. J. L. Mackie, “Causes and Conditions,” in Myles Brand, ed., The Nature of Causation (Urbana: University of Illinois Press, 1976), pp. 308-44; Roy Bhaskar, The Possibility of Naturalism. A Philosophical Critique of the Contemporary Human Science (Brighton: Harvester Press, 1979); Rom Harré and Peter Secord, The Explanation of Social Behaviour (Oxford: Basil Blackwell, 1973); Heikki Patomäki, “How to Tell Better Stories about World Politics,” European Journal of International Relations 2 (March 1996), pp. 105-34.

---

80. Robert Jervis, System Effects: Complexity in Political and Social Life (Princeton: Princeton University Press, 1997).

81. Feng-hsing Hsu, Thomas Anantharaman, Murray Campbell and Andrew Nowatzyk, “A Grandmaster Chess Machine,” *Scientific American* (October 1990) 263, pp. 44-50; “The Deep Blue Team Plots Its Next Move,” [www.research.ibm.com/deepblue/home/may11/story\\_1.html](http://www.research.ibm.com/deepblue/home/may11/story_1.html), a web site that includes more articles, the games and commentary on the play.