Why Historians (And Everyone Else) Should Care About Counterfactuals

Abstract

There are at least eight good reasons practicing historians should concern themselves with counterfactual claims. Furthermore, four of these reasons do not even require that we are able to tell which historical counterfactuals are true and which are false. This paper defends the claim that these reasons to be concerned with counterfactuals are good ones, and discusses how each can contribute to the practice of history.

"Counterfactuals" are a particular kind of conditional statement. Sometimes, when we wonder what happens if something else happens, we are genuinely unsure whether that something else has happened, or will happen. Other times, we remain interested in conditionals even when we are sure the antecedent condition is not met. "How would I feel now if I had gone for a run first thing this morning?" is a question I might be interested in if I am trying to weigh up whether to take up running in the mornings. Since I think it would be helpful to think about a particular case (today), in order to shed light on my future options, I remain interested in the question of how I would feel if I *had* gone for a run even after I know that I did *not* go for a run first thing this morning.

The particular group of conditional statements classed as counterfactuals are characterised by several features. The paradigm cases often have "would" in them, concern antecedents in the past, and have antecedents that are false. Here are some examples:

If the bus had not arrived, I would not have made it to my appointment.

If Germany had not attacked the USSR in World War Two, they would have kept France longer.

If Earth had not had a moon, there would have been milder tides.

If I had become a farmer, I would now be a billionaire.

These sorts of constructions are sometimes called "subjunctive conditionals", though whether this is a good name for them is an issue that is entangled with the issue of how to apply the category of "subjunctive" found in traditional grammar.

There are less standard cases of conditionals that might be counted as counterfactual. Some theorists think that conditionals with antecedents partially or entirely about the future (or the present) can fall in that category:

If the USA were to pull out from South Korea tomorrow, that would embolden the North Koreans.
If I were to shoot you right now, I would be a criminal.

Though it is controversial how one might map the common distinction between "counterfactuals" and "indicative conditionals" to the future. The way "counterfactual" is used in philosophy, it applies also to conditionals constructed the same way but with true antecedents:

If Xerxes were to have attempted to invade Greece, he would have been unsuccessful.
If France had an independent nuclear deterrent during the cold war, it would have complicated US/USSR relations.

"Might" conditionals that are otherwise constructed like counterfactuals are counted as counterfactuals, though some think their treatment should be quite different from "would" counterfactuals:

If I had gone for a run this morning, I might have been feeling sore now.
If Germany had tried to invade Britain during World War Two, it might have captured London.

All of these constructions (and their equivalents in other languages) are very common in everyday talk. But their use in history is contested. There is a tradition in twentieth

century historiography that counterfactuals are useless for serious history—they may be entertaining as flights of fancy, but evaluating them is no part of the job of a serious historian. E.H. Carr dismisses discussing counterfactuals about the Russian Revolution as "a parlour game with the might-have-beens of history" (Carr 1986 p 91), for example, and D.H. Fischer dubs the attempt to establish the truth of historical counterfactuals the "fallacy of fictional questions" (Fisher 1970 pp 15-21), to just cite two of the more colourful condemnations of counterfactuals in history. In part, this might be a reaction to *certain* counterfactuals, which the historian might feel are no part of their serious business. The most suspicious ones, I imagine, are those that have antecedents that require radical differences from our world, and/or those which rely on changes which ramify through many complicated processes. Evaluating counterfactuals such as the following are ones that many serious historians would regard as irrelevant to the tasks of history:

1. If the Aztecs had entered the industrial revolution in the fourteenth century, they would have conquered Europe.

2. If Germany had conquered Europe by 1946 and then landed in North America, its armies would have been stopped before they reached Chicago.

3. If Newton had been Russian, Cartesian physics would have been orthodoxy in the nineteenth century.

Those who are inclined to think of dabbling with counterfactuals as engaging with the more fanciful "what ifs" that can capture the imagination might understandably think the professional history ought to be more sober.

But of course many counterfactuals are far more sober, and settling them depends much more on finding out facts about the real world. This is familiar from everyday life. If I say in an email "If I had been in at the office this morning, I would be able to find that reference for you straightaway", then we can easily see what sorts of facts are relevant to

evaluating it: whether the information is in my office, my habits of searching, whether there was anything else going on at the office this morning that would have consumed my attention, and so on. Such a counterfactual is a world away from Nazi armies in the mid-West or Russian Newtons.

Recently, more and more historians and philosophers of history have been arguing that counterfactuals have a place in the practice history, even an important place, so it seems worthwhile to produce a list of some of the main reasons they can be important for practicing historians. Many of these are mentioned elsewhere, and I should particularly mention the work of Richard Ned Lebow (e.g. Lebow 2000), from which I draw some of these suggestions. Since many of the apparent advantages of using counterfactuals I will discuss have been mentioned elsewhere, at least in passing, some aficionados may worry that my message is stale news. I think there are three reasons why it is not. The first is that it is useful to have a list of advantages all in one place: and while this is not the only list (See Lebow 2000 pp 557-564, Tetlock and Parker 2006 pp 17-28) this paper does bring together advantages that are not all included on such lists.

The second is that my argument for the benefit of counterfactuals is in several places more nuanced than that in the current literature. For example, many theorists seem to presume that commitment to causal claims and commitment to counterfactual claims are virtually the same thing, while I suggest that this connection is rather more nuanced than those thinkers allow.

The third distinctive contribution I want to make is to stress something about the reasons to employ counterfactual thinking. Some reasons for historians to be concerned with counterfactuals are largely independent of which counterfactuals are true, or correct. Some hostility to using counterfactuals in history is based on the suspicion that we cannot tell which counterfactuals are correct, if indeed any are. So the fact that theorising using counterfactuals can be useful *even if such scepticism is warranted* is very interesting, and as far as I can tell not sufficiently appreciated.

My topic is the use of counterfactual conditionals in history, but I should point out that the debate in history and the social sciences about dealing with "counterfactuals" is sometimes broader. Sometimes that debate not only concerns these conditionals, but any consideration of non-actual possibilities. This not only includes what would happen if.., but also what might have happened, and what other possibilities there are, even if there is no simple antecedent that would yield them. Much writing of "alternate histories", even if they are treated as "what ifs", are probably best seen not as spelling out a counterfactual conditional, but an entertaining and perhaps informative exploration of a mere possibility not too unlike what in fact occurs. Considering "what could have been?" questions can be valuable for some of the reasons "what would have been?" questions are valuable, but they are not my focus here.

While the focus of this paper is on the use of counterfactuals in history, the benefits of employing counterfactuals to be discussed are much broader. Much of our concern about the past and the past behaviour of our fellow human beings arises in contexts outside academic history. Most of the social sciences takes past activity as part of its subject matter and most, or all, of its evidence. Everyone in their own lives has a concern with their past and the past activity of others. A concern with history, to a greater or lesser extent, is everyone's business. If reasoning about counterfactuals and taking counterfactuals seriously is helpful in historical inquiry, especially in the ways I will outline, reasoning with counterfactuals and about counterfactual subject matters is something that would benefit us all.

I will begin with four reasons for taking counterfactuals seriously that are rather independent of the issue of the truth or correctness of counterfactuals. Then I will discuss four reasons to use counterfactuals that *do* suppose we can distinguish between correct and incorrect counterfactuals, with some degree of accuracy at least.

Reasons to Use Counterfactuals Independent of Truth

Considering what would have happened if various contrary-to-fact conditions obtained—Britain did not enter World War I, the Ostrogoths had not attacked Italy, Beethoven had

never been born, or whatever—has a variety of psychological effects. At the risk of lumping together too many disparate psychological effects in one category, let me first mention various respects in which thinking about counterfactuals "broadens the mind". It can expand the historical imagination: thinking about hypothetical cases might give one ideas to apply elsewhere. For example, if we think about what would have happened if Julius Caesar had attempted an all-out invasion of Britain more seriously, we might gain some insight into the problems that were probably faced when Claudius's troops tried to turn Britannia into a province. If we think about how the early Japanese shogunate would have dealt with a deep split in religious authority, we might gain some insight into how the "feudalism" of the early Shogunate compares and contrasts with the "feudalism" of Latin Christendom during the Investiture Conflict. This need not depend on us getting it *right* about how the Shogunate would have been—merely considering the question in this way (and a grounding, of course, in the non-counterfactual facts about medieval Japan and Europe) could lead us to consider the right sorts of issues. So considering counterfactuals can be "mind expanding" both by expanding our imaginations and suggesting interesting hypotheses for further investigation.[1]

Another way consideration of counterfactuals can be of value, even if we do not get them right, is by clarifying positions and driving debate about areas of conflict. We might accept that a variety of things influenced a particular historical event: but we can highlight assumptions about what were important or controlling factors by considering counterfactuals about situations where influences were absent or different. If it turned out that we could come up with access to the truth of counterfactuals, we could also use this process as part of discovering what *are* important and controlling factors (see below): but even without that, we can at least illuminate our *assumptions* about such factors. Illuminating assumptions is something we can do by engaging in counterfactual thinking, especially when we find our counterfactual judgements disagreeing. This, in turn, can

---

[1] Stimulating the imagination is mentioned by Tetlock and Belkin 1996a p 15, and Tetlock and Parker 2006 p 25 recommend counterfactuals as a remedy of the narrow-mindedness resulting from exclusively factual "framing".

drive debate—we can get clearer on what disagreements about non-counterfactual matters there are when we unpack *why* there are particular counterfactual disagreements.[2]

Yet another way that considering counterfactuals can be valuable, regardless of their truth, is stressed by writers such as Ned Lebow and Philip Tetlock (Lebow 2000 pp 558-560, which includes a number of references to unpublished work by Lebow and Tetlock; Tetlock and Belkin 1996 pp 15-16; and Tetlock and Parker 2006 pp 22-28). People are known to suffer from *hindsight bias*—in hindsight, events can seem inevitable and often overdetermined, even if they were considered unlikely or highly chancy beforehand. Lebow, for example, argues, with empirical work to back him up, that this hindsight bias is drastically reduced after subjects are encouraged to consider various counterfactual versions of a given scenario. Perhaps it might be thought that World War II was more or less inevitable (because of Hitler's warmongering, or intensive nationalism on all sides, or the growing tension between fascism and bolshevism, or whatever). But, if Lebow is right, we should expect that this judgement of inevitability will be reduced if people think carefully about counterfactual scenarios. What if France had sent their military into the Rhineland when Hitler tried to unilaterally remilitarise it? What if the United States had joined the French and British announcement that an attack on Poland would lead to war? What if the Soviet Union had refused to agree to the Molotov-Ribbentrop pact? Connected to this phenomenon of reversing hindsight bias is the phenomenon of increasing awareness of contingency. Again, Lebow argues, that people who are invited to seriously consider counterfactuals about a given scenario are inclined to see the outcome as more contingent than they would have otherwise.

Of course, counteracting hindsight bias and increasing belief in historical contingency are only good things if hindsight bias is a pernicious bias and the appearance of a lack of historical contingency is a mere appearance. We would also need to be careful about counterfactual thinking producing excesses the other way—a belief in too much

---

[2] Lebow 2000 p 563 briefly mentions this. Fearon 1996 talks about the use of counterfactuals as "spotlights" for other issues.

contingency and lack of overdetermination. That said, my suspicion is that both hindsight bias and an initial lack of appreciation of contingency *are* both tendencies that it is worthwhile counteracting, though making the case for that would be difficult. In any case, it is worth noticing that these effects do not depend on correctly determining which counterfactual suppositions are *correct* (e.g. about what would have happened if France had used force to oppose the remilitarisation of the Rhineland). They have their effect, apparently, simply as part of the psychology of engaging in counterfactual suppositions of various sorts.

The final way of employing counterfactual thinking that is of value, regardless of the truth of the counterfactuals themselves, should be of especial concern to those historians who think that gaining understanding of historical subjects "from the inside" is worthwhile. (History as *Verhesten* would be one example of this.) Whatever a sceptical historian might do, many historical subjects themselves entertained counterfactual suppositions and were concerned with which counterfactuals are true. A good example of trying to understand historical actors' counterfactual beliefs in a historical investigation is the work done in Herrmann and Fisherkeller 1996, which examines US policy towards Iran, especially in 1946 and 1953. To take one example, Herrmann and Fisherkeller discuss the influence of the view that "If President Truman had not threatened the Soviet Union with forceful action in 1946 [over Soviet occupation of northern Iran], Stalin would not have retreated from northern Iran and Moscow would have kept satellite governments in Azerbaijan and Kurdistan". Acceptance of this counterfactual encouraged US policy makers to "stand up to" the Soviet Union in future confrontations (Herrmann and Fisherkeller p 156). Another example, cited by Martin Bunzl 2004 for a slightly different purpose, is Greif 1998, which explains some of the dynamics of Genoese society in the eleventh and twelfth centuries, despite the lack of documentary evidence of the views of important figures in Genoese society, in terms of the rational strategies to pursue, presumably implicitly relying on the ability of Genoese citizens to engage in counterfactual reasoning about what courses of action would serve them best, and our ability to understand them as so reasoning.

Consider another example. In the nineteenth century, Bismark presumably considered carefully what the effect would have been if the UK had formed the impression Germany planned to compete for naval equality, or if Turkey had reached a secret agreement with Russia about the fate of various Balkan Slavs.

A lot of Bismark's conditional thinking would have been future tense—if Germany were to do *this*, what will the response be by the other powers, if *that* sort of opportunity presents itself, what should the Kaiser order, and so on. Those conditionals are not paradigm counterfactual judgements, and so we might want to set them aside. But some of his thinking, no doubt, would have been counterfactual thinking directed to the past. When Bismark was trying to decide on the lessons of the Franco-Prussian war, I expect he would have thought about how different decisions by Prussia or France would have influenced the course of the conflict and the final outcome. Now, *whether or not Bismark was correct* in *any* of his counterfactual judgements, a historian trying to understand Bismark's point of view had better take into account Bismark's attitudes towards such counterfactuals, and try to speculate about what Bismark's attitudes to various counterfactuals might have been even when there was not specific evidence. This would not apply to any old counterfactual—what Bismark's view of the counterfactual "If Russia had been as industrialised as Britain..." might have been is probably of little interest to the political historian. But counterfactuals about how Britain would have reacted to different announcements about German naval intentions, for example, are a different matter. A historian interested in Bismark, or the personalities involved in great power conflict in Bismark's day, might quite reasonably feel the need to consider such counterfactuals *as they would have seemed* to Bismark.

This last reason to take counterfactuals seriously might seem to miss the point. Surely *we* do not need to take counterfactuals seriously to theorise about historical actors who did, any more than we need to take the prospect of punishment by the Greek gods seriously to understand historical actors in Ancient Greece. The difference between counterfactual thinking, on the one hand, and belief in Zeus, on the other, it seems to me, is that we have to engage in the practice of counterfactual reasoning (or at least simulate such an

engagement) to properly judge what counterfactuals were accepted by historical actors.[3] Of course, it would have to be counterfactual reasoning starting from how our actors took things to be, not how we now took things to me, but that is a very familiar part of trying to see things, at least partially, from a historical actor's perspective. It is true that evaluating a "what if" from a historical actor's perspective, with the main aim of determining what that actor might have thought about it, is a different matter from thinking that it is part of the historian's business to pronounce on "what ifs" in her own voice. But it should give pause to those who think that historians should never, or rarely, entertain counterfactuals.

"Mind expanding" through invigorating historical imagination and suggesting new hypotheses; identifying assumptions and driving debate; counteracting hindsight bias and increasing appreciation of contingency; and enabling a better appreciation of historical actors' situations gives us a far from trivial list of benefits. This would already counter some of the animus shown by historians to engaging with counterfactuals. However, I think counterfactual reasoning can give us much more of historical value. For the next set of rewards, though, we would need to suppose that we can usefully sort the "good", accurate, correct counterfactuals (in my book, the true ones), from the others. Let us therefore turn to the issue of what this would do for historical inquiry.

Using True Counterfactuals

A lot of concerns about the use of counterfactuals centre on the difficulty with finding true counterfactuals. There are at least two sources of this concern. One is epistemic.

---

[3] Specific recommendations about how to simulate such counterfactual reasoning may invoke counterfactuals in another way. R.G. Collingwood suggests, for example, in coming to understand the battle of Trafalgar one may ask oneself "what should I have done, if I had been in Nelson's place?" (Collingwood 1939 p 113). If part of "understanding" of historical actors proceeds by working out what reactions I would have had if I were in that actor's situation with that actor's character, then reasoning with counterfactuals is at the core of the method. Carrying this out successfully, of course, would require coming up with the *correct* counterfactuals about what I would have thought or done, and so may better be included in the next section. I thank Stephen Davies for discussion of Collingwood's method.

With so many factors potentially relevant to the truth of a historical counterfactual, and with so many of them relatively inaccessible to researchers working today, we might fear that historians will usually not have the information that would be needed to evaluate a counterfactual. The other source is more metaphysical or perhaps semantic: one variety is the worry is that many interesting counterfactuals just lack determinate truth-conditions, or alternatively that the wealth of possibilities that incorporate the antecedent plus other constraints leave too much unsettled, which might make most, or all, interesting historical counterfactuals false. (See Lewis 1973 pp 77-83 for a discussion about whether this is best seen as yielding indeterminate counterfactuals or false ones, albeit in the context of the machinery of Lewis's theory of counterfactuals.) Another variety is that counterfactual conditionals do not have truth-conditions at all: they are not in the business of being true or false, but serve some other purpose (see for example Edgington 1995). We could have a distinction between "correct" and "incorrect" conditionals that did not require truth and falsehood but was intersubjective enough and stable enough to serve many of the purposes to be discussed below. But a number of non-factualist approaches to conditionals will not be suitable for these purposes.

Both the epistemic challenge and the metaphysical/semantic challenge raise big issues, and not ones that will be settled in this paper. Responding to scepticism about any general area of human inquiry is difficult, and the metaphysical and semantic issues about counterfactuals will be best resolved, I think, by comparing the overall performance of general theories of counterfactuals. So while there are probably sceptics about our ability to discover the truth of counterfactuals in particular, and there are definitely those who think most counterfactuals of the sort we are concerned with are false (see Hajek, in preparation), the rest of this paper will rely on the supposition that counterfactuals are in good shape epistemically, metaphysically and semantically. (I hope to defend that conviction at greater length in future work.) Those who do not agree can see the remaining argument of the paper as somewhat conditional: on the assumption that counterfactuals can be true and we have reasonable epistemic access to whether they are true, at least in a range of interesting historical cases, it will be argued that the historian can derive a range of valuable benefits from employing them.

The first reason for a historian to be concerned about counterfactuals is one of the most obvious, but nevertheless it is worth saying a few words in its defence. Information about counterfactuals is interesting in its own right.[4] We would like to know what would have happened if Louis the Pious had had one surviving male heir, or Germany had not attacked the Soviet Union in World War Two, or Copernicus had not proposed his heliocentric model of the planets. Of course not everyone is interested in these questions, and perhaps not every historian is very interested in the counterfactuals connected to the part of the subject they focus upon. But many find these questions about history as interesting as many others. Furthermore, it is not just historians who might find these questions of interest. The general public has an appetite for considering these questions—an appetite being catered to by the increasing number of "What If" books written by historians but aimed at more general audiences.[5] Students of history are also often intrigued by counterfactual questions, at least until they are indoctrinated to think that these questions were no matter of concern for the serious historian.

What are we to say to those who think that despite historical counterfactuals being of interest to some historians, significant sections of the public, and many students, they are nevertheless unworthy of serious attention by historians? I suppose that the question of the "aims of history" is an endlessly controversial one, but I presume historical inquiry can properly have more than one aim. And I would have thought satisfying historical curiosity was one legitimate aim of the historian, even if it is not as lofty as some that have been proposed, like revealing the ways of God to man or providing the material for political revolution. Historical information is valuable in its own right, I think, and I hope many historians would agree: and this seems especially so when that historical information is particularly interesting. Information about what counterfactuals are true

---

[4] Editors of collections of "What If?" books often mention this, and Scalmer 2006 mentions the "fun" of considering counterfactual history as one of its attractions.

[5] Macintyre and Scalmer 2006 is a very recent example. Rosenfeld 2005 in chapter 1 provides an interesting overview both of counterfactual history and recent alternate history fiction.

seems to inherit this justification as much as any other interesting information about the past.

However, there are more particular, and perhaps deeper, reasons for historians to concern themselves with discovering which historical counterfactuals are true. One use of counterfactuals that has received some attention is the role of counterfactuals in causal judgements.[6] Historians can hardly avoid speculation about causation if they are to manage anything more than mere chronologies of events. Historical works are full of claims about certain events producing or causing others, as well as the use of other "causal verbs": actors respond to previous events, some factors influence others, or give rise to others, and there are killings, creations, consequences expected and unexpected. As well as the target of historians' inquiry often being causal, examining historical evidence unavoidably requires evaluations in causal terms. An archaeologist, today, finds some items buried in the ground. To work out what these items tell us about the past, we need to have a good idea of what caused them to be found as they were, and what processes are responsible for their present condition. Historians go from information about the find to hypotheses about the cause of the find being the way it was. Or when a historian relies on a written "record", she necessarily has to consider what the causal links are between the text in front of her and events in the past. Is the text a result of a chain of testimony through reliable reporters? Or is it simply the result of an active imagination of a historical writer, or the deceptive intentions of a contemporary hoaxer? Or the effect of some other process altogether? History, as it is practiced, cannot do without causal judgements.

Causation is closely connected to counterfactual matters. It is controversial among philosophers exactly what the connection is. It is fairly clear that it is not as simple as A

[6] This is cited by several defenders of examining counterfactuals in history, including Gould 1969, Fearon 1996, Tetlock and Parker 2006 pp 18-22, and testing causal hypotheses by Lebow 2000 p 563. Bunzl 2004, pp 7-8 on the other hand, seems to claim that the direction of support is from causal information to counterfactual information: we use causal information to discover counterfactuals, not the other way around.

causing B always going together with it being the case that A and B in fact occurred, and if A had not occurred, B would not have occurred either. This is not always appreciated by defenders of the use of counterfactuals: Tetlock and Parker 2006, for example, say on p 17 that "Whenever we draw a cause-effect lesson from the past, we commit ourself to the claim that, if key links in the causal chain were broken, history would have unfolded otherwise." and "whenever we make the apparently factual claim that factor *x* made a critical causal contribution to outcome *y* we *simultaneously* make the counterfactual claim that, in a logical shadow universe with factor *x* deleted, outcome *y* would not have occurred". (p 18).

One sort of case where there is causation without counterfactual dependence of this sort is overdetermination: it could be that an ugly building might cause a skyline to be spoiled, even though the skyline would have been spoiled without that particular building, perhaps because of another building soon to be built or a new program of belching industrial waste into the atmosphere. And there are arguably several other ways to have causation without this sort of direct counterfactual dependence. The relationship between causation and counterfactual dependence is somewhat subtle—nevertheless, in a wide range of cases, we are prepared to accept that without A, B would not have occurred only in the same situations where we would be prepared to accept that A was a cause of B.

Some philosophers have proposed that causal talk can be analysed in terms of claims about counterfactuals. Many of the early versions of these counterfactual theories of causation faced severe difficulties, though this has not dissuaded some philosophers from continuing the attempt to analyse causation counterfactually. (See for example Lewis 2000 for a sophisticated attempt to analyse causation counterfactually.) If causal claims are equivalent to counterfactual claims—especially if this equivalence is something like synonymy—then any historian who investigates a causal hypothesis or renders a causal judgement is thereby engaging in some sort of counterfactual judgement, albeit in language that is not explicitly counterfactual.

However, it is not so easy to see what lesson we should draw from historians' use of causal judgements about the acceptability of counterfactual judgements if causation is not analysed counterfactually. And it is probably still the majority view of philosophers working on causation that causation is not susceptible to a counterfactual analysis. For some it is because they hold that, on the contrary, the truth conditions of many counterfactual judgements have to be understood in terms of a prior conception of causation (see Jackson 1977), and for some it is for other reasons (see e.g. Armstrong 1999). If causal judgements are not equivalent to counterfactual judgements, but just tend to be associated with them, then what is to stop a historian interested in some causal hypothesis from just reasoning directly about the causal relationships involved and ignoring any associated counterfactuals?

Tetlock and Belkin (1996 pp 8-10) talk of "nomothetic theory-testing" using counterfactuals and Lebow talks of "testing" causal assumptions by considering counterfactuals (Lebow 2000: 561-564): for example, we can create models that reflect causal links we believe in but vary parameters from their actual values to see whether the model behaves as we would expect. It is unclear how this "tests" our original causal assumptions, unless it is through checking for consistency or plausibility: if variations in parameters produce outcomes that conflict with other beliefs we have about the causal structure (e.g. they allow something to happen we are sure is impossible), then that might give us reason to reject one or other of the assumptions that went into the model.[7] If the assumptions work well together and the model yields results that are intuitively "realistic", then that might speak in favour of the assumptions used to construct the model.

If we are using models with counterfactual values for parameters merely for consistency checking, then while we are reasoning about non-actual set-ups, it is not clear that counterfactuals need play any role. We can set up conditions on models, reason generally

---

[7] Lebow tells me (private communication) that something like this is what he had in mind: we can use these counterfactual interventions in models to "evaluate the feasibility of the causal claims".

about the properties of classes of such models, and draw conclusions about whether any of these models are unacceptable without using any explicitly counterfactual constructions—we can restrict our attention to what models there in fact are that meet certain constraints, and what those models are in fact like. An analogy will suggest itself to philosophers: when we are examining a logic, we can work out facts about what follows from what by reasoning explicitly about what models the logic has—if all the models in which the premises of an argument are true are ones were the conclusion of that argument also is, then the argument is valid, according to one standard treatment. This does not mean we have to worry about whether the conclusion *would* be true *if* all of the premises were. There may well be that sort of counterfactual dependence with many valid arguments, but it is practically speaking beside the point when we reason about the models of a logic to prove a meta-logical result.

So it is not obvious to me that "testing" causal hypotheses by considering models with some parameters that take other than their actual values is any good as a vindication of historians using counterfactuals.

I think the close link between counterfactual judgements and causal judgements indicates a different sort of reason why historians considering causal hypotheses should take counterfactuals seriously. We could imagine a purified historical science where historians reasoned rigorously from causal data to causal conclusions, ignoring any judgements in counterfactual form. But real historical practice is more of an art than this imaginary science. In evaluating many historical causal chains, we have to rely on our implicit sense of what is likely to have depended on what. When we ask why the Persian Empire was conquered by Alexander, the causes we will look at will be partly military: how the Macedonian phalanx and Companion cavalry fare against different sorts of Persian units, for example. They will be partly social: what the political and economic relationship between satrapies and the central government in the Persian empire functioned, and how that caused reactions in different parts of the empire when Alexander won battles in other parts. They will be partly psychological: we would try to understand the Persian administration's attitudes to Greeks and the Macedonian

administration's attitude to Persians, and what impact those attitudes had on behaviour on both sides. We might try to understand the character of Alexander and of Darius III, to see how that will have influenced interactions within and between Macedon and Persia. None of this is standardly done with explicit causal models, or testing formal causal models by changing the values of various parameters.

What historians do employ is their implicit understanding of human processes, their awareness of how the case at hand resembles and does not resemble other historical processes that seem analogous, and in general the art of good historical judgement. In relying on this capacities, rather than explicit algorithms or reasoning about classes of formally specified causal models, the historian should respect all sorts of judgements they are naturally inclined to make in these circumstances. That is not to say these judgements are incorrigible or must be the last word, but they cannot be safely ignored either. And some of these judgements will, I believe, be counterfactual ones.

A historian of Alexander's war against Persia may well be inclined to believe that Alexander's success in the battle of Issus did not depend on the attitude or actions of the satrap of Egypt. That should be taken into account when determining the causes of success at Issus—*prima facie*, such a story can leave out information about the satrap of Egypt. Or to take a less obvious example, historians might wonder whether Alexander's victory at Issus was caused, in whole or part, by Darius's behaviour on the battlefield. One reasonable approach to this question, once our evidence is before us, is to consider what would have happened if Darius had made different decisions. Perhaps Darius's decision to flee the battle well before it was over was crucial. But whether that decision was crucial in part seems to depend on what would have happened if Darius had not fled—if his staying would have just meant he would have been killed or captured shortly after, his fleeing would seem less likely to be a cause of Alexander's victory than if Darius would have been able to continue successful resistance if he had not fled. Here I am not saying that considering counterfactuals about Darius's decisions is *necessary* to determine the causes of Alexander's victory at Issus—but in fact if a historian's process of consideration of the case goes through counterfactuals, then that historian might not

have an equally good way of coming to a considered judgement about the causal question without them. Causal reasoning is at least a lot more difficult, in complicated situations, if we must avoid any counterfactual thoughts.

Perhaps not all reasoning about causal hypotheses important to history are like this—Bunzl 2004 pp 4-5 cites an example of a historical question of why a certain bridge collapsed when it was first built. Some of the story here has to do with the practices of contracting, the political pressures that resulted in the sort of bridge that was built being constructed, and so on: but some of the story is the physical explanation of why girders in such-and-such a configuration undergo enough stress in such-and-such winds so that the structure will collapse. This part of the story need not rely on a historical sense for causal influences so much as applying an explicit, quantitative model from the natural sciences and engineering. Historians may of course have occasion to employ explicit quantitative models and causal reasoning that does not need to be sensitive to the counterfactual judgements informed experts are inclined to make. But this sort of causal reasoning remains somewhat an exception rather than the rule when examining the causal relationships usually of interest to historians. When we come to judge the causes of the decline and fall of the Western Roman empire, our methods will resemble those of Gibbon more than those of engineers investigating a bridge collapse.

While causal reasoning in history remains a matter of art, the fact that we go naturally from counterfactual questions to causal questions and back, even if these transitions are non-deductive, ought to be treated with respect. We face a mass of information, about the specific event to be evaluated, other similar events, and a general grasp of human psychology, cultural influences, our understanding of the constraints put on historical actors by the physical environment, and so on. Provided our ordinary competent reasoning about causation employs counterfactuals, and especially if it does so in a way that cannot be codified easily and is not via explicit algorithms, then a historian relying on that capacity should not also reject reasoning using counterfactuals, on pain of their principles not lining up with their practice. I think this is the best way to argue that the close links between causal and counterfactual judgements mean that historians interested

in causation (and most, if not all, should be) should not reject a role for counterfactuals in historical reasoning.

Closely connected with reasoning about causation will be reasoning about other topics in what some philosophers call the "nomic family" of cause-related topics. The question of what *dispositions* historical actors or institutions had, for example. Or we might be interested in whether certain outcomes (actual or non-actual) had a high or low *chance* of occurring: chance in some fairly objective sense, not merely the question of how confident we should be that it occurred, or how confident observers at the time should have been that it would occur. Issues about the necessity or contingency, in context, of a historical outcome can arise. Was a certain event *inevitable*, or relatively inevitable, for example?[8] In all of these kinds of cases our ordinary reasoning may invoke counterfactuals. When we want to know how fragile a peace was, for example, we would naturally consider what sorts of events would have led to the peace breaking *if* they had occurred. When we want to know whether one outcome was very likely, we would naturally consider what sorts of events would have prevented it *if* they had happened, and how likely they were. When we claim a development was inevitable, we often think it would have occurred under a range of counterfactual conditions as well as the actual one, and so on. Exactly what dispositions, chance and contingency have to do with causation are difficult questions. But that our reasoning about these things is in fact caught up with our causal and counterfactual reasoning seems very plausible.

Another use of counterfactuals, still connected their causal applications though perhaps more distantly, is the use of counterfactuals in determining what *explains* what. While some historians might have suspicions about "causation", surely every historian thinks that explanation is important: they aspire to more than listing what happened when, but they also want to provide explanations of important or surprising historical events.[9]

---

[8] Thanks to Aness Webster for a suggestion here.

[9] The role of using counterfactuals when searching for explanations has not received much attention in this literature. Tetlock and Parker 2006, for example, "assert the impossibility of avoiding counterfactual history if we hope to go beyond bare description of what happened to explanations of why certain things

Explanations are also wanted when dealing with evidence: a historian would like an explanation of why a chronicler wrote what he did, or why a particular distribution of artefacts is found in an archaeological dig. Some philosophers argue that an explanation of an event often just is a specification of what that event counterfactually depended upon: for example, James Woodward at one point says "I have defended the idea that explanations must answer what-if-things-had-been-different questions" (Woodward 2003 p 226), in the course of his defence of what he calls a "counterfactual theory of causal explanation". If philosophers like Woodward are right, to engage in causal explanation is to engage in work intimately connected to discovering the relevant counterfactuals.

Even if counterfactual "analyses" of explanation are no good, however, judgements about counterfactuals will still have an important role to play. It is a commonplace that many different explanations of the same event are possible. (The lit match being dropped into the oily rags explains why the fire started, as does the intentions of the arsonist, as does the negligence of the parole board which released him despite strong evidence he would reoffend.) Some explaining events are more like trigger events, however, and others are more like citing a general background that made events of the sort to be explained more likely. However, one kind of explanation we are often particularly interested in consists of detailing the conditions without which the event to be explained would not have happened. The particular meeting where Hitler's cabinet decided to invade Poland is certainly a cause of Germany's invasion, and indeed a cause of World War II (or at least that part of World War II that occurred in Europe). But that meeting was more in the

---

rather than others happened." (p 17, italics suppressed). But this is followed immediately by a discussion of causation instead. Fearon 1996 is aware they are different issues (p 39 n1), but drops into "explanation" language from time to time even though his topic is determining causation. Geoffrey Hawthorn, who does think reasoning about counterfactuals is important to explanation, also runs explanation and causation together, but he is at least explicit that he thinks that the word "'cause' (and 'reason' and suchlike terms) serve to say that it is an *explanation* we are offering" (Hawthorn 1991 p 25). Perhaps it is because the issue of determining causes and the issue of determining explanations are not always properly distinguished that the role of counterfactual reasoning in discovering explanations is relatively neglected compared to the role of counterfactuals in considering causal hypotheses.

nature of a trigger: to just be told about that meeting and its immediate consequences would not be satisfying, in part because World War II most probably would have happened without it. Germany would have invaded Poland a bit later, perhaps, or Nazi expansionism would have produced armed conflict with Britain and France at some later stage. Or at least we cannot be sure that it would not have.

If we want to give a satisfying explanation of the outbreak of war in the European theatre of World War II, a better set of facts to cite than the facts of Hitler's particular plans for Poland would be some set of facts that World War II would not have happened without, and ideally were sufficient or close to sufficient (i.e. those facts obtaining would be enough). A more general story about the relationships between the major European powers, the history of German revanchism, the Munich agreement, and so on, together perhaps with some information about the individuals in the governments of Germany, the United Kingdom, France, and probably the Soviet Union as well.

Explanations which specify conditions without which the phenomenon to be explained would not have happened seem to be of particular interest. If they also specify phenomena which are such that they are enough to ensure the relevant phenomenon would occur, they are desirable in a different way, but also one that seems to involve the same sort of conditional as in paradigm counterfactuals. So historians interested in offering satisfying explanations of historical phenomena may need to attend to the counterfactual relationships between the phenomena they cite in explanation and the phenomenon to be explained.

One use of counterfactuals for historical purposes that has rarely been focused on in the recent literature is in the attribution of responsibility, and in the determination of the appropriateness of regret and pride, and to a lesser extent praise and blame. Plausibly, to determine whether someone was responsible for an outcome, part of what we would like to know is whether that outcome would have occurred anyway. This is by no means the only thing that matters – I can be responsible for tracking mud onto a carpet even if it was inevitable that mud would soon be tracked onto the carpet anyway. But it does make

some difference to responsibility, we often think – how much responsibility the assassins of Ferdinand have for World War 1 depends in part on whether or not a great European war was inevitable and would have eventually been sparked by something. If it had turned out, for example, that Austria would have created a pretext for war with Serbia in any case within the next year or two, the share of the blame borne by the assassins is arguably less than if the international situation would have stayed peaceful without such an extreme intervention.

Counterfactual considerations also enter into whether regret or praise is appropriate, and the extent of regret or praise that might be appropriate. How much we should regret the lack of first-world intervention to stop the Rwandan genocide in 1994 depends, in part, on whether realistic intervention would have made a difference. In this case many people think that it would have made a substantive difference. On the other hand, other actors may have made less of a difference. My guess is that Malawi would have not been able to make an important difference even if it had tried to intervene in the genocide, and intervention by some other closer neighbours may have even exacerbated trouble. If I am right about Malawi's inability to prevent much of the massacres, then it seems less appropriate for Malawi to regret its lack of intervention than it might for France or the USA.

Of course, what the outcome would have been, or what the probable outcome would have been, is not the only factor relevant in regret. One might regret the failure to intervene because one thinks it shows a lack of the right sort of intentions, whether or not the intervention would have been successful. How much regret one might legitimately feel depends in part on the degree of moral responsibility for a region or an outcome, that is not immediately tied to how much counterfactual control one had. Even if we discovered that the United States could have stopped the genocide in Rwanda as easily as it could have stopped acts of mass-murder in one of its client states, we might still think the US had more legitimate reason for regret in the latter case because of their influence in creating the situation in the first place, or the responsibilities they incur when they become heavily involved in another nation's affairs.

Another factor other than counterfactual ones that plays some role in responsibility and regret but plays an especial role in the case of praiseworthiness or reasonable pride is the question of what was intended in a course of behaviour.[10] If some happy outcome was an intended outcome of a policy or a decision, that is often more of a source of pride than if it was an unintended consequence. Obviously what the exact determinants of responsibility, praise and blame, for historical occurrences are is a difficult question, and no doubt a variety of factors play a role, whether we are considering the responsibility of individuals, nations, or other entities. The factors are probably even more complicated when we come to issues of praise and blame—not only are there the sorts of factors already mentioned, there may be an extra dimension of pragmatics to decisions to assign praise or blame—sometimes it is better to let sleeping dogs lie, or to employ more or less charity, and so on. But provided information about counterfactuals is *one* component in the appropriate evaluation of responsibility, legitimate pride and legitimate regret, praise and blame, then they have a role.

Some might object that it is not the historian's job to make value judgements, especially of actors or institutions in the distant past. I suspect this is not a very popular position at the moment, and it is hard to see how it could be correct in general. *Even if* some historians are pursuing a project that does not require judgements, who is to say that it is no part of *any* historian's job to engage in them: different historians have different jobs, even as historians, and see their roles in different ways, and sometimes this is just because different working historians occupy different roles. In any case, even if it was no part of the historian's job to make value judgements like attribution of responsibility, or recommendations for pride or regret, the need for counterfactuals in those judgements might still matter to historians. For even if historians are not to engage in such value judgements *qua* historians, surely it is plausible that one of their tasks is to provide those

---

[10] The attribution of responsibility, praise and blame is mentioned briefly in Tetlock and Belkin 1996a p 8. Lebow 2000 p 564 claims that counterfactuals are needed for "evaluation of outcomes" as good or bad, pointing out that we often employ a counterfactual benchmark against which we measure the actual outcome.

who do make such value judgements with the historical facts available that bear on those judgements. "Consumers" of history will still want to know what about their institutions or countries they can legitimately feel pride in or they should legitimately regret, for example. The view that historians, qua historians, should not only not make value judgements, but they should not be concerned with getting the factual basis right for others to make value judgements, seems bizarre. So the role of counterfactuals in judgements of responsibility, regret and pride, praise and blame, give us an important reason why discovering some counterfactual truths, when we can, is worthwhile for the historian.

To sum up: eight reasons have been discussed for historians to interest themselves in historical counterfactuals and the sort of reasoning that involves those counterfactuals. The final four presuppose that we can, with some reliability, sort the correct counterfactuals from the incorrect ones, but the first four do not even suppose that.

The eight are:
1) "Mind Expanding". Considering counterfactuals can invigorate the historical imagination and suggest new, non-counterfactual, hypotheses for investigation.
2) Bringing out Disagreement. Discussion of counterfactual scenarios can make underlying assumptions about the course of events clearer and bring to light disagreements that may not so easily come out if historians stick to only discussing what in fact did happen.
3) Mitigating Hindsight Bias and Increasing Appreciation of Historical Contingency. These advantages are particularly stressed by Lebow, and to the extent that hindsight bias is distorting, or people have trouble recognizing the full extent of historical contingency, surely correcting for these errors is worthwhile.
4) Understanding from the inside. Historical actors worried about counterfactual matters and tried to assess counterfactuals—to understand events from their point of view seems to require the historian to enter into taking those counterfactual issues seriously.

5) The value of counterfactuals in their own right. Many people are curious about counterfactual questions and interested in counterfactual judgements: these seem to provide legitimate topics of inquiry for their own sake.
6) Causation. Counterfactuals are intimately linked to causation, and assessing counterfactuals is an important part of our practice of making causal judgements. Likewise with related judgements such as those about dispositions, objective chances, and contingency or its lack.
7) Explanation. Some explanations of particularly valuable sorts line up with counterfactual dependence of the explanadum on the explanans.
8) Informing value judgements. Counterfactual information is relevant to assessments of responsibility, of the legitimacy of pride or regret, and of praise or blame.

Not every historian will find every application useful, of course. I would also not want to give the impression that there are no pitfalls in the investigation of counterfactuals, or be seen as a partisan of untrammelled counterfactual speculation in history. Of course there are interesting and difficult issues about how to go about investigation of which counterfactuals are correct in particular historical situations, and both art and skill must be brought to bear. In this respect adjudicating the correctness of historically interesting counterfactuals does not differ from any interesting theoretical enterprise. There is more, much more, to be said about the best use of counterfactuals in historical inquiry. I hope in this paper to have shown that at least there *are* significant uses for counterfactuals in the historical enterprise, and indicated what some of the benefits might be of successfully carrying out counterfactual investigations.[11]

Daniel Nolan
Department of Philosophy
University of Nottingham
University Park, Nottingham
NG7 2RD
United Kingdom

---

[11] Thanks to Richard Ned Lebow for the initial stimulation for writing this paper and for discussion, and thanks to an audience at the University of Auckland for discussion. Thanks to the Leverhulme Trust for financial support of the research of this paper through a Philip Leverhulme Prize.

References

Armstrong, D.M. 1999. “The Open Door: Counterfactual vs. Singularist Theories of Causation”. In Sankey, H. (ed) *Causation and Laws of Nature*. Dordrecht: Kluwer, 175-85

Bunzl, M. 2004. “Counterfactual History: A User’s Guide”. *The American Historical Review* 109.3: pp 1-12: page numbers are from the online version.

Carr, E. H. 1986. *What is History*, 2nd edition (1st edition 1961). Macmillan, London.

Collingwood, R.G. 1939. *An Autobiography*. Oxford University Press, Oxford.

Edgington. D. 1995. "On Conditionals". *Mind* 104: 235-329

Fearon, J.D. 1996. “Causes and Counterfactuals in Social Science: Exploring an Analogy between Cellular Automata and Historical Processes” in Tetlock and Belkin 1996.

Fischer, D. H. 1970. *Historians’ Fallacies: Toward a Logic of Historical Thought*. Harper & Row, New York.

Gould, J.D. 1969. “Hypothetical History”. *The Economic History Review*. 22.2: 195-207

Greif, A. 1998. “Self-Enforcing Political Systems and Economic Growth: Late Medieval Genoa” in Bates, R., Greif, A. Rosenthal, J. and Weingast, B. eds. 1998. *Analytic Narratives*. Princeton University Press, Princeton, pp 28-63

Hajek, A. in preparation. “Most Counterfactuals are False”.

Hawthorn, G. 1991. *Plausible Worlds*. Cambridge University Press, Cambridge.

Herrmann, R.K., Fisherkeller, M.P. 1996. "Counterfactual Reasoning in Motivational Analysis: U.S. Policy toward Iran". in Tetlock and Belkin 1996, pp 149-167

Jackson, F.C. 1977. "A Causal Theory of Counterfactuals". *Australasian Journal of Philosophy* 55: 3-21

Lebow, R. N. 2000. "What's So Different About a Counterfactual?". *World Politics* 52: 550-583

Lewis, D. 1973. *Counterfactuals*. Harvard University Press, Cambridge MA.

Lewis, D. 2000. "Causation as Influence". *Journal of Philosophy* 97: 182-97

Macintyre S. and Scalmer, S. 2006. *What If? Australian History As It Might Have Been*. Melbourne University Press, Melbourne.

Rosenfeld, G. D. 2005. *The World Hitler Never Made: Alternate History and the Memory of Nazism*. Cambridge University Press, Cambridge.

Scalmer, S. 2006. "Introduction" in Macintyre and Scalmer 2006 pp 1-11

Tetlock, P. E. and Belkin, A. 1996a. "Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives" in Tetlock and Belkin 1996b pp 1-38.

Tetlock, P. E. and Belkin, A. (eds). 1996b. *Counterfactual Thought Experiments in World Politics*. Princeton University Press, Princeton.

Tetlock, P.E. and Parker, G. 2006. “Counterfactual Thought Experiments: Why We Can’t Live Without Them and How We Must Learn to Live With Them” in Tetlock, Lebow and Parker 2006 pp 14-44.

Tetlock, P.E., Lebow, R.N. and Parker, G. (eds) 2006. *Unmaking the West: “What-If”? Scenarios that Rewrite World History*. University of Michigan Press, Ann Arbor.

Woodward, J. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, New York.