

# What Should We Require from an Account of Explanation in Historiography?

AQ

*Dr Veli Virmajoki*

Finland Futures Research Centre, University of Turku

*vevirm@utu.fi*

## Abstract

In this paper, I explicate *desiderata* for accounts of explanation in historiography. I argue that a fully developed account of explanation in historiography must explicate many explanation-related notions in order to be satisfactory. In particular, it is not enough that an account defines the basic structure of explanation. In addition, the account of explanation must be able to explicate notions such as *minimal explanation*, *complete explanation*, *historiographical explanation*, *explanatory depth*, *explanatory competition*, and *explanatory goal*. Moreover, the account should also tell how *explananda* can be chosen in a motivated way. Furthermore, the account should be able to clarify notions that are closely connected with *explanation* such as *historical contingency*. Finally, it is important that the account is able to recognize when explanation-related notions and issues are so closely intertwined that we are in danger of not seeing the differences between them. In other words, I argue that a satisfactory account of explanation in historiography must have the power to explicate central explanation-related notions and to clarify discussions where the differences between the notions are obscure. In order to explicate these *desiderata*, I formulate a (version of the) counterfactual account of explanation and show how that account is able to explicate explanation-related notions and clarify issues that are connected with historiographical explanations. The success of the counterfactual account suggests that historiographical explanations do not differ fundamentally from explanations in many other fields.

## Keywords

explanation in historiography – historical counterfactuals – historical contingency – explanatory depth – explanatory goal – the choice of *explanandum*

## 1 Introduction

In the philosophy of historiography,<sup>1</sup> the analyses of explanation have focused on what kind of connection must hold between some Z and X in order for Z to explain X. The thing to be explained is called *explanandum* (X in our case) and the thing that explains (Z and some additional information) is called *explanans*, and the point of investigation is to explicate the structure of *explanans* in historiography. For example, it has been suggested that Z and X must be connected by universal laws;<sup>2</sup> that the connection is that of reason and action;<sup>3</sup> that there is an (ephemeral) mechanism;<sup>4</sup> that the connection is that of difference-making;<sup>5</sup> and, as a limiting case, that the connection requires no analysis,<sup>6</sup> to mention some prominent suggestions.

While all these suggestions – and the discussions they have generated – have deepened our understanding of historiographical explanation, there is still something missing in the debates concerning the nature of historiographical explanation: the initial suggestions concerning the connection between *explanans* and *explanandum* have not led to *fully developed accounts of explanation*. A fully developed account is one that is able to explicate many explanation-related notions such as *minimal explanation*, *complete explanation*, *historiographical explanation*, *explanatory depth*, *explanatory competition*, and *explanatory goal*. Moreover, a fully developed account should also tell how *explananda* (things to be explained) can be chosen in a motivated way. Furthermore, the account should be able to clarify notions that are closely connected with *explanation* such as *historical contingency*. Although some accounts of explanation in historiography have discussed some of these issues,<sup>7</sup>

1 By “historiography” I mean “the study of history” and by “history” I mean the past events and processes.

2 Carl G. Hempel, “The function of general laws in history”. *Journal of Philosophy* 39:2 (1942). 35–48.

3 See R. G. Collingwood, *The Idea of History* (Clarendon Press, 1946); W. H. Dray, *Laws and explanations in History*. (Oxford University Press, 1957).

4 Stuart Glennan, “Ephemeral Mechanisms and Historical Explanation”. *Erkenntnis* 72:2 (2010). 251–266.

5 See Max Weber, *On the Methodology of Social Sciences*, (The Free Press of Glencoe, 1949); Yemima Ben-Menahem, “If Counterfactuals Were Excluded from Historical Reasoning”. *Journal of the Philosophy of History* 10:3 (2016). 370–381.

6 Aviezer Tucker, *Our Knowledge of the Past: A Philosophy of Historiography*. (Cambridge University Press, 2004).

7 For example Glennan, “Ephemeral Mechanisms and Historical Explanation” discusses historical contingency and Hempel “The function of general laws in history” can be interpreted as making a distinction between complete explanations (that mention a universal law) and historiographical explanation (that are “explanatory sketches”).

there has been no systematic and unified treatment of all the issues involved. The lack of a fully developed account of explanation is a problem since without such an account (1) there is no systematic analysis of the important notions mentioned above, (2) it is difficult to assess the relative merits of different suggestions concerning the structure of explanation since the explicatory and clarificatory power of the suggestions (i.e. to what extent explanation-related notions and issues can be explicated and clarified) is a very important dimension in such assessments, and (3) without being able to assess the relative merits of the suggestions, it is difficult to assess whether historiographical explanations differ from explanations in other fields.<sup>8</sup> Seen from another angle, we need an analysis of the basic explanatory structure because such analysis makes it possible to provide further explications of explanation-related notions. The function of analyzing the basic structure of explanation is to provide the basis of further explications and clarifications. While it is probably true that historians are able to arrive at and justify explanations without analyzing the structure of explanation,<sup>9</sup> the function of a philosophical account of historiographical explanation is not to allow historians to explain but to clarify explanatory claims when those claims are “confused, unclear, and ambiguous” and to suggest “how these limitations might be addressed” as Woodward elegantly puts it.<sup>10</sup>

In this paper, I formulate a counterfactual account of explanation in historiography (based on the interventionist account by James Woodward<sup>11</sup>) and, on the basis of the account, explicate the above-mentioned explanation-related notions and discuss issues that are connected with historiographical explanations, such as the nature of explanatory counterfactuals and the choice of *explananda*. The main aim of this paper is to explicate *desiderata* for fully developed accounts of historiographical explanation and, at the same time, to formulate a (version of the) counterfactual account that is applicable to historiography. These goals serve each other: by introducing an account of explanation that is able to explicate explanation-related notions, those notions and their role can be better understood in the philosophy of historiography. At the same time, the fact that the account is able to explicate many important

8 In my view, we cannot debate the differences between historiographical explanations and explanations in other fields before we have at least one fully developed and powerful account of historiographical explanations. Otherwise it is difficult to understand on what our judgements about the differences should be based. (See section Conclusion.)

9 As Tucker argues in *Our Knowledge of the Past* (Chapter 5).

10 James Woodward, *Making Things Happen. A Theory of Causal Explanations* (Oxford University Press, 2003), 7.

11 Woodward, *Making Things Happen*.

notions and to clarify unclear issues can be seen as a reason to accept the account.

In what follows, I will first describe the core idea of the counterfactual account of explanation and then discuss explanation-related notions and issues in historiography on the basis of the account. I tailor the interventionist framework to the nature of historiography. As we move forwards, the details, basic ideas, and motivation behind the interventionist account of explanation are developed further. We will see how those details, ideas and motivations are helpful when thinking about explanation in historiography. To achieve this, I develop further and locally modify certain notions of the counterfactual account.<sup>12</sup> In particular, I develop new notions such as *historiographical explanation*, *competing explanations*, and *explanatory goal*, modify the notion of *explanatory depth* from the one formulated by Hitchcock and Woodward,<sup>13</sup> and discuss historiographical counterfactuals from the point of view of interventionism. Woodward's theory has proven its philosophical worth in many fields.<sup>14</sup> Therefore, there is no worry that the account solves historiographical problems in an *ad hoc* manner. Rather, the applicability of the interventionist account on historiography indicates that historiographical explanations do not fundamentally differ from explanations in many other fields.

## 2 A Counterfactual Account of Explanation

In James Woodward's account of explanation, "the underlying or unifying idea in the notion of causal explanation is the idea that an explanation must answer what-if-things-had-been-different questions, or exhibit information about a pattern of dependency."<sup>15</sup> To put it simply, explanations answer questions of the form "Why X rather than Y?" by pointing out factors Z and W such that "had W rather than Z been the case, Y rather than X would have been the

<sup>12</sup> This also means that we are not merely passively applying the account for the purposes of historiography but also developing the account further on the basis of historiography.

<sup>13</sup> Christopher Hitchcock and James Woodward, "Explanatory generalizations, part II: Plumbing explanatory depth", *Noûs* 37, no 2 (2003): 181–199.

<sup>14</sup> E.g. James Woodward, "Causation in biology: Stability, specificity, and the choice of levels of explanation", *Biology and Philosophy* 25:3 (2010), 287–318; Lauren N. Ross and James Woodward, "Koch's postulates: An interventionist perspective". *Studies in History and Philosophy of Science Part C*. 59 (2016), 35–46; M. Rescorla, "An interventionist approach to psychological explanation". *Synthese* 195:5 (2017). 1909–1950. Jaakko Kuorikoski and Petri Ylikoski, "Explanatory relevance across disciplinary boundaries: the case of neuroeconomics", *Journal of Economic Methodology* 17:2 (2010), 219–228.

<sup>15</sup> Woodward, *Making Things Happen*, 201.

case.”<sup>16</sup> Explanations provide information about counterfactual dependencies between *explanans* and *explanandum* and are contrastive in nature.<sup>17</sup> An explanation “must enable us to see what sort of difference it would have made for the explanandum if the factors cited in the explanans had been different in various possible ways.”<sup>18</sup>

Let’s take an example. We ask

Why did scientists come to believe that atoms exist?

In order to be able to answer this question, we need to know what contrast is assumed by the questioner. For a moment (I will return to this in Sections 3 and 6) we may assume that the question is

Why did scientists come to believe that atoms exist rather than believe that atoms do not exist?

The answer to this question is

Scientists believe that atoms exist rather than believe that atoms do not exist because Einstein *formulated an explanation* of the Brownian motion and Perrin confirmed this explanation with his *experimental work*. Had there not been such explanation or experimental work, scientists would not believe in atoms.<sup>19</sup>

16 Woodward’s account defines causal and explanatory relations as holding between variables (i.e. the value *explanandum*-variable depends on the value of *explanans*-variable). I simplify the discussion and therefore I do not write in terms of variables. How the notions here can be written in such terms see Veli Virmajoki, *Cementing Science*, (University of Turku), Chapter 6.

17 The idea of contrastive explanations is not a novelty of Woodward’s account. There has been a lot of discussion of contrastive explanations (see e.g. H. L. A. Hart and Tony Honoré, *Causation in the Law*, (Oxford University Press, 1959); B. van Fraassen, *The Scientific Image* (Oxford University Press, 1980); A. Garfinkel, *Forms of Explanation*, (Yale University Press, 1981); P. Lipton, “Contrastive Explanations,” in Knowles (ed.), *Explanation and its Limits* (Cambridge University Press, 1990), 247–266.

18 Woodward, *Making Things Happen*, 11. Notice that the idea that historiographical explanations are based on difference-making is not a novelty by any means: See Weber, *On the Methodology of Social Sciences*; Ben-Menahem, “If Counterfactuals Were Excluded from Historical Reasoning”.

19 See e.g. J. Renn, “Einstein’s invention of Brownian motion”. *Annalen der Physik* 14 (2005), 23–37, and Stathis Psillos, “Moving Molecules Above the Scientific Horizon: On Perrin’s Case for Realism”, *Journal for General Philosophy of Science*, 42:2 (2011), 339–363, discussing this topic.

Notice that here we have a very simple explanation. The explanation shows how the belief in the existence of atoms *depends on* the work of Einstein and Perrin. Of course, such explanation does not tell much. It is rather minimal in the quite literal sense that it cites only two counterfactual situations such that had one of those occurred, the outcome would have been different: one where no explanation of the Brownian motion was formulated, and one where the explanation was not confirmed by experimental work. A much more satisfactory explanation would also describe what the assumptions, methods, and values of the relevant scientific community at the time, and what kind of processing was required before the results of Einstein and Perrin became accepted. We could try to establish these issues and provide a more complete and deeper (see Section 4) explanation. However, once we know the counterfactual dependencies between the works of Einstein and Perrin and the belief in atoms, we have an explanation. We are able to answer what-if-things-were-different questions. That the threshold for explanatory status not being set very high is an advantage in historiography. We should not set the bar too high and end up saying that there are very few explanations in historiography. Clearly, historiography has been able to explain many events and processes. Moreover, even though the threshold is not set very high, we can still seek deeper and deeper explanations. This is a fruitful way to look at historiographical explanations: we start from a minimal explanation that answers one what-if-things-were-different question and strive towards an explanatory account that gives a more detailed picture and answers more what-if-things-were-different questions, towards a more complete explanation.

However, complete explanations are never achievable. This is not a problem in itself, since complete explanations, even if they could be formulated, always incorporate rather pointless information. For example, it is true that *scientists not being brainwashed* explains why they came to believe in atoms; had they been brainwashed to believe that atoms do not exist, they would not have believed that atoms exist.<sup>20</sup> As complete explanations are unreachable, it is obvious that historiographical explanations cannot be equated with such explanations. In practice, *historiographical explanations* lie between minimal and complete explanation. They cite some set of counterfactual situations where the outcome would have been different. Historiographical explanations are collections of true counterfactual statements of the form “had W rather than Z been the case, Y rather than X would have been the case,” that is, a

20 The formula for such *explanantia* is the following: Take any condition C that would have cancelled the effect of Z on X. Then “C was not the case” is explanatory for X.

historiographical explanation consists of a set of minimal explanations. For example,

- (1) Scientists would not have come to believe in atoms, had Einstein not explained the Brownian motion.
- (2) Scientists would not have come to believe in atoms, had Perrin not performed his experimental work.

is a set of minimal explanation and therefore a (rather incomplete) historiographical explanation. A historiographical explanation provides information about situations where the outcome would have been different. In this way, a historiographical explanation answers many what-if-things-were-different questions.

A set of minimal explanations (i.e. a historiographical explanation), may describe (i) the effect of different factors on *explanandum* (“had Einstein not provided the explanation OR had Perrin not done the experimental work, scientist would not have come to believe in atoms”), or (ii) the effect of different alternatives (contrasts) to the actual antecedent on *explanandum* (“had Einstein provided no explanation OR had he provided a different explanation, scientists would not have come to believe in atoms”), or (iii) both.

On the basis of the counterfactual account of explanation we can say that:

A minimal explanation describes one counterfactual situation in which Y rather than X would have been the case and therefore answers only one what-if-things-were-different question; a complete explanation incorporates all minimal explanations; and a historiographical explanation incorporates some set of minimal explanations. On the basis of these explications, we are able to understand historiographical phenomena such as explanatory competition and explanatory goals.

Before proceeding to those notions, one more thing must be noted:

The account of explanation suggested in this paper does not make a distinction between causal and reason-based explanations.<sup>21</sup> The account of historiographical explanations developed here is based on counterfactuals of the form “Had W rather than Z, Y rather than X would have happened.” I take it that such counterfactual claims can also cite reasons or beliefs, and therefore my account can incorporate explanations that cite reasons and beliefs, as our example including Einstein and Perrin did. Moreover, given relevant sources,

21 See Collingwood, *The Idea of History*, and Dray, *Laws and explanations in History*, arguing for what we call here “reason-based explanations”; see Donald Davidson, “Actions, Reasons, and Causes”, *Journal of Philosophy* 60:23 (1963), 685–700, for famous criticism; and see Giuseppina D’Oro, “Historical Understanding” in *A Companion to the Philosophy of History and Historiography*, ed. Tucker (Wiley-Blackwell, 2009), discussing the issue.



we can follow the rational, as well as non-rational, reasoning (no matter how the notion of rationality is understood) of some historical actor step-by-step and tell how the agent's conclusion would have been different had a different step been taken. Therefore, my account incorporates reason-based explanations. Davidson famously argued that

[a] person can have a reason for an action, and perform the action, and yet this reason not be the reason why he did it. Central to the relation between a reason and an action it explains is the idea that the agent performed the action *because* he had the reason. Of course, we can include this idea too in justification; but then the notion of justification becomes as dark as the notion of reason until we can account for the force of that 'because.'<sup>22</sup>

Within the counterfactual account, one can have a reason for an action and perform the action without the reason being explanatory *if the action did not depend counterfactually on that reason* (in a relevant manner, see section 5 for discussion). This case is identical to one where an event E and an outcome O occur but E does not explain O despite the fact that E always leads to O *because O does not counterfactually depend on E* (in a relevant manner). The "force of because" lies in the counterfactual dependency.

It seems that a distinction between reason-based and causal explanations would make it impossible to understand how historiographical studies are able to provide unified explanatory information despite mixing these two types of explanations, and therefore positing identical deep-structure (counterfactual dependency in our case) behind both types is advantageous. To illustrate this, let's consider the following passage from John Krige. It describes the final steps before it was claimed that a new fundamental particle (the W boson) had been discovered at CERN:

According to Schopper's account, then, two factors left the CERN directorate little option but to place their faith in Rubbia's results and claim that the W had been discovered. One was intrinsic to the internal politics and institutional logic of a contemporary high-energy physics laboratory like CERN. The Stanford physicist Stanley Wojcicki has described it well: "UA1 decided to go public and publish – that certainly was their decision alone – and once that was done the news was in the public domain. In light of its importance it would have been very unusual for CERN to

<sup>22</sup> Davidson, "Actions, Reasons, and Causes", 693.



officially ignore this fact and not have a press conference.” The second factor was Schopper’s promise to Thatcher and the need to avoid doing anything that might offend one of the major contributors to the CERN budget. When the laboratory was established in the early 1950s the British government had had grave doubts about its participation in the venture; in the late 1960s it had initially refused to participate in the construction of the SPS. Now the Conservative government was hesitating about its engagement in CERN’s next big machine, the LEP, plans for which had been adopted in principle by the member states a few months before. The discovery of the W helped dispel any remaining doubts about the quality of the physics being done at CERN and hastened the final agreement to finance the electron-positron collider.<sup>23</sup>

Krige’s study, taken as a whole, discusses many types of factors that led to the claim that the W boson was found. Reasons were an essential part of the unfolding of the process as well as the technological environment, the effect of which can hardly be described without causal language. There are no distinctions between causes and reasons within the study; both are parts of the same narrative. Moreover, it is quite natural to interpret the claims in the citation above as having implicit counterfactual form: “Had the subject matter been less important, CERN could have ignored the maneuver of the UA1 team”; “Had there not been a need to be cautious with major contributors, different decision could have been made.” There is nothing mysterious in causes and reasons being united within a single explanatory narrative if we take them to carry similar dependency-relations. These remarks mean that the counterfactual account is *monistic* in the sense that it claims that all explanations in historiography share the same fundamental structure. However, it is not monistic in the sense that it would deny the existence of reason-based explanations.

### 3 Competition and Goals

As there can exist many historiographical explanations for an *explanandum* of the form “why X rather than Y”, it becomes important to understand when two such explanations compete.<sup>24</sup> Even though we cannot find *the* explanation for

23 John Krige, “Distrust and Discovery: The Case of the Heavy Bosons at CERN”. *Isis* 92:3 (2001). 537.

24 Woodward, *Making Things Happen*, does not discuss competitions and goals explicitly. I discuss them as they seem especially important in historiography.

something, it does not follow that all explanations are compatible. Note that it has been argued that no competing historiographical explanations of events can be found.<sup>25</sup> This claim is suspicious. Even if it was true that no competing explanations can be found from published historiographical studies, in the process of historiographical enquiry many possible explanations are considered and one or more are established as the most credible one. Moreover, we will see (in Section 4) that there is sometimes explanatory competition within a single research paper – a phenomenon that has gone virtually unnoticed. In order to clarify these aspects of historiography, we need to understand when two historiographical explanations are competitors.

We can use the notion of minimal explanation as our starting point. We can say that two minimal explanations are competitors if and only if they disagree on how the *explanandum* depends on factor Z. To be more precise, we can say that two minimal explanations E and E\* are competitors if and only if E says that “Z led to X and W would have led to Y” and E\* says “Z led to X and W would have led to Q” (where  $Q \neq Y$  but, as a limiting case, it is possible that  $Q = X$  if the contrast between Z and W is explanatorily irrelevant according to E\*).

Moreover, two complete explanations are competitors if they incorporate competing minimal explanations or if they incorporate different minimal explanations. The idea in the latter amendment is that if a complete explanation E cites the minimal explanation M but E\* omits M, then M is explanatory according to E but not explanatory according to E\*, as complete explanations mention every explanatory detail.

Finally, two historiographical explanations are competitors if they either incorporate competing minimal explanations or if they incorporate different sets of minimal explanation. In this case, omitting a minimal explanation does not automatically mean that its explanatory relevance is denied. It can simply mean that the minimal explanation is not relevant from the perspective of an explanatory goal (see below) and we should be careful to recognize when the difference in explanations is due to a difference in explanatory goals and when it is not.

It is important to notice that two explanations cannot be competitors (in the technical sense) if they answer different *explananda*. However, this does not mean that these explanations are equally good or relevant since some choices of *explanandum* are worse than others. I will discuss (in Section 6) why the choice of *explanandum* should be critically evaluated and on what kind of considerations that choice must be based.

Let's return to our example.

<sup>25</sup> Tucker, *Our Knowledge of the Past*, 191–197.

Why did scientists come to believe that atoms exist?

This question has (at least) two readings:

- (I) Why did scientists come to *believe* that atoms exist rather than have no beliefs at all?
- (II) Why did scientists come to believe *that atoms exist* rather than believe that atoms do not exist?

The first question can be answered by citing all sorts of factors, such as nutrition and oxygen, that are necessary for there to be cognitive agents. Had there not been nutrition or oxygen, scientists would not have had any beliefs. However, in relation to the second question, which was asked in our original example, the factors that are necessary for cognitive agents to exist are irrelevant in the explanation since those factors need to be in place in order for there to be any beliefs. Those factors do not tell us why one of the beliefs came to be held rather than the other. However, here we do not have explanatory competition since the *explananda* differ.

On the other hand, the following historiographical explanations are competitors:

- (1) Scientists believe in atoms rather than believe that atoms do not exist because Einstein *formulated an explanation* of the Brownian motion and Perrin confirmed this explanation with his *experimental work*. Had there not been such explanation or experimental work, scientists would not believe in atoms.
- (2) Scientists believe in atoms rather than believe that atoms do not exist because *Einstein* formulated an explanation of Brownian motion and *Perrin* confirmed this explanation with his experimental work. Had someone else formulated the explanation or performed the experimental work, scientists would not believe in atoms.

There are two senses in which the two can be competitors. One is the 'hard sense': they disagree whether it was the personal prestige or the details of the work that was explanatorily relevant in the process. In the case of hard competition, the explanation (1) must have the following structure:

(1<sup>c</sup>) Scientists believe in atoms rather than believe that atoms do not exist because Einstein *formulated an explanation* of the Brownian motion and Perrin confirmed this explanation with his *experimental work*. Had there not been such explanation or experimental work, scientists would not believe in atoms. However, had someone else formulated the explanation or performed the experimental work, scientists would still believe in atoms.

Both (1<sup>c</sup>) and (2) cannot be true and, therefore, there exists a hard competition between them. In such cases, the competition can be resolved only by careful historiographical research and judgement.

The other one is the ‘soft sense’. Assume that it is true that both the details of the work and the personal prestige are explanatorily relevant, that is, had the details or the person who did the work been different, the outcome would have been different. Then the difference between (1) and (2) might reside in a difference in explanatory goals. We could say, for example, that (1) aims to underline the role of explanations and experiments in the history of science, whereas (2) aims to underline the role of social hierarchies in the history of science.<sup>26</sup> It is important to distinguish these two senses of explanatory competition.

When two different explanations are offered, the most charitable reading of them usually is that they attempt to achieve different explanatory goals. This means that we can accept many different explanations of the same *explanandum*. However, this does not mean that anything whatsoever can have explanatory value (nor does it mean that hard competition never exists). We already argued that only if there are counterfactual dependencies between two events, can there exist an explanatory relationship. Moreover, asserting that something is always (or never) explanatory should make us suspicious. From the fact that Z is explanatory/irrelevant with respect to some goal, it cannot be inferred that Z is explanatory/irrelevant with respect to every goal.

It is also important to notice that even if a Z is not interesting from the perspective of the explanatory goal that the historian has, Z cannot be ignored during a historiographical research. Consider the following:

A politician says that scientists’ interest in climate change (S) is the cause of an increase in measured temperatures (T): the more interested the scientists are, the more the temperature increases, as can be seen from statistics. The politician is naturally interested in the political aspects of science and thus he wants to highlight them. However, it is a mistake to ignore the way the world is. There has been a global warming (G), and G is the common cause of both S and T. If G is ignored due to the explanatory goal, an explanation is found where none exists – S is not a cause of T even if it might appear so (see

<sup>26</sup> Perhaps it is most natural to think that soft competition is not real competition at all. However, I still think the category of ‘soft competition’ is useful since soft and hard competition have the same surface properties and we can decide whether two explanations that appear to be in competition are hard (or “real”) competitors only by further investigation. If these explanations are not hard/real competitors, then it is advantageous to have the category ‘soft competition’ which reveals that the explanations appear to exist without there being a serious tension between them. This way, we can avoid many confusions.

Section 5). Such mistakes are easily made if the distinction between soft and hard competition is not explicated.

Moreover, historians with different explanatory goals, participating in soft competition, can always attempt to find common ground in an explanation that is more complete than the ones the historians initially provide. For example, if the difference between explanations (1) and (2) above is soft and reflects different explanatory goals, the beliefs in atoms can be explained as follows:

Had there been no explanation or experimental work OR had these works been produced by someone else than Einstein and Perrin, scientists would not have come to believe in atoms.

Two historians with different explanatory goals could both accept this claim about the patterns of counterfactual dependencies. However, if they did not, we would have a hard competition between two explanations. This, again, indicates the importance of the distinction between hard and soft competition: the latter allows us to achieve deeper explanations by combining different historiographical results, the former does not.

But how should the notion of explanatory goals be understood? No neat catalogue of such goals can be given. The notion is based on the idea that historians want to highlight certain aspects of the process that led to the *explanandum*. The explanatory goal is what the explanation seeks to establish in addition to providing information about conditions on which the *explanandum* depends. An explanation may seek to pinpoint the role of some types of factors (F) in history and this attempt is centered around the (meta)explanatory goal of understanding the role of F in historical processes. In this sense, the choice of the explanatory goal aims at second-order understanding (whereas an answer to a particular *explanandum* gives us first-order understanding).

The choice of explanatory goals cannot be discussed in detail here. However, it must be noted that there are two possible approaches to the analysis of such goals.<sup>27</sup> First one is axiological. In the axiological approach, the choice of explanatory goals reflects our values and we want to understand the role of an F in history because F is considered important with respect to our values. The second approach is epistemic. In this approach, we attempt to find out a set of factors we can use to effectively explain much of history. For example, we might make the hypothesis “experiments play an important role in

---

27 Although it must be admitted that the choice of explanatory goals in actual historiography is probably based on the trends in the field rather than detailed reflection of values or epistemology.

scientific developments” and attempt to confirm or falsify this hypothesis by taking a number of historical cases and evaluating the explanatory power of experiments. Notice that sometimes the axiological and epistemic approach coincide. For example, we might attempt to give a marginalized group the recognition it deserves by pointing out how the actions of that group have shaped history.

We have seen that as long as we are able to distinguish hard competition from soft competition, we can accept that historical events and processes can be explained in various ways. There is no need for a forced choice between different explanations in the cases of soft competition. However, it would still be an advantage if something could be said about how to assess the relative merits of different historiographical explanations. I now turn to the issue of explanatory depth.

#### 4 Explanatory Depth

The notion of explanatory depth has gone mostly unnoticed in the philosophy of historiography and, unfortunately, the lack of understanding about the notion has sometimes led to the ridiculing of important historiographical issues. For example, it has been argued that

The best explanation for the independent diaries of soldiers of the same unit that state that on a certain date their unit came under heavy bombardment and therefore panicked and retreated is that indeed they came under heavy bombardment and therefore panicked and retreated. There is no need for further knowledge of psychology or human nature under fire. Had this explanation of the retreat depended on psychological theories, it would have been indeterminate, since under fire soldiers are known to retreat out of fear, become paralyzed with fear and stay put, or become emboldened with rage and charge forward. The mere description of the stimulus together with background conditions and contemporary psychology is insufficient for explaining the actual retreat. But psychology is redundant here.<sup>28</sup>

The passage is correct in claiming that sometimes providing a (rather minimal) explanation is a rather simple task, given the evidence. However, it seems that “knowledge of psychology and human nature” is too quickly dismissed.

<sup>28</sup> Tucker, *Our Knowledge of the Past*, 187.

For the very reason that “under fire soldiers are known to retreat out of fear, become paralyzed with fear and stay put, or become emboldened with rage and charge forward” it would be interesting to know why they did retreat in this case. In other words, we would like to know what properties of the group of soldiers would have led to a different behavior under the bombardment. For example, perhaps less experienced soldiers would not have retreated because they would not have understood the danger. It seems that if we knew that both the heavy bombardment and the experience of the soldiers made a difference in the outcome, we would have achieved a deeper explanation. If psychology can provide us with such explanatory knowledge, it is far from redundant. An account in explanation in historiography should be able to dissect such dimensions of explanatory practice.

In this section, I outline criteria for explanatory depth using the ideas introduced by Hitchcock and Woodward.<sup>29</sup> We have seen that *explanandum* of the form “Why X rather than Y?” is explained by a set of *explanantia* of the form “Because Z rather than W.” In science, these sets of *explanantia* can often be read from *generalizations* that describe the relationships between variables. Because of this, the notion of explanatory depth in Hitchcock and Woodward (2003) is defined with respect to such generalization. However, due to the nature of historiography, I do not characterize explanatory depth with reference to generalizations. Rather, I explicate criteria that enable us to compare the explanatory depth of two historiographical explanations (i.e. sets of counterfactual statements). This change from generalizations to collections of counterfactuals requires that the criteria of explanatory depth are somewhat rewritten from Hitchcock and Woodward.<sup>30</sup> However, we can still retain the core of the idea that “[one] generalization can provide a deeper explanation

29 Hitchcock and Woodward, “Explanatory generalizations, part II: Plumbing explanatory depth”; Woodward, *Making Things Happen*.

There are other ways to analyze explanatory depth/power, see Petri Ylikoski and Jaakko Kuorikoski, “Dissecting explanatory power”, *Philosophical Studies* 148:2 (2010), 201–219; Brad Weslake, “Explanatory Depth”. *Philosophy of Science* 77:2 (2010), 273–294; Cyrille Imbert, “Relevance, Not Invariance, Explanatoriness, Not Manipulability: Discussion of Woodward’s Views on Explanatory Relevance”. *Philosophy of Science* 80:5 (2013), 625–636; H. Bhogal, “Coincidences and the Grain of Explanation”, *Philosophy and Phenomenological Research* (2019). I follow Woodward and Hitchcock because their ideas are most closely connected to the ones discussed in the previous sections.

30 Instead of writing about “more invariant generalization”, I write about “deeper explanation”. In Hitchcock and Woodward, “Explanatory generalizations, part II: Plumbing explanatory depth”, the relative depth of an explanation depends on the range invariance of the generalization that is used in the explanation. As generalizations are rarely applicable in historiography, I do not use the term *invariance*.



than another if it provides the resources for answering a greater range of what-if-things-had-been-different questions, or equivalently, if it is invariant under a wider range of interventions”<sup>31</sup> by connecting explanatory depth in historiography of science with the ability to answer what-if-things-were-different questions. In what follows, I explicate some cases where one historiographical explanation has better ability to answer what-if-things-were-different questions than another explanation. Many more such cases can perhaps be explicated and my aim in this section is to challenge scholars to reflect on the issue of explanatory depth rather than to provide a complete list of the different dimensions of explanatory depth.

(I) We can say that E is a deeper explanation than E\* if E gives information about counterfactual alternatives to the *explanans* Z that E\* is silent about. For example, if E says

Had Einstein not explained the Brownian motion OR had he provided a different explanation, scientists would not have come to believe in atoms.

and E\* says

Had Einstein not explained the Brownian motion, scientists would not have come to believe in atoms.

then E is a deeper explanation than E\* since it answers a what-if question about Einstein giving a different explanation of the Brownian motion.

Notice that the number of what-if questions is not the only dimension that is relevant in such cases. Some what-if questions are more important than others. For example, in some context, it might be more important to know what would have happened if Einstein provided a different explanation than to know what would have happened if no explanation was provided. If we wanted to know the influence of personal prestige, this question would be interesting.

(II) E is deeper than E\* if E provides information about a factor F that is not provided by E\*. In other words, E is a more complete explanation than E\* as it makes explicit a factor that is only a background conditions from the perspective of E\*.<sup>32</sup>

<sup>31</sup> Hitchcock and Woodward, “Explanatory generalizations, part II: Plumbing explanatory depth”, 198.

<sup>32</sup> “G” makes explicit the dependence of the *explanandum* on variables treated as background conditions by G.” (Hitchcock and Woodward, “Explanatory generalizations, part II: Plumbing explanatory depth,” 187.)

For example, assume that the following explanations are soft competitors:

- (1) Scientists believe in atoms rather than believe that atoms do not exist because Einstein *formulated an explanation* of the Brownian motion and Perrin confirmed this explanation with his *experimental work*. Had there not been such explanation or experimental work, scientists would not believe in atoms.
- (2) Scientists believe in atoms rather than believe that atoms do not exist because *Einstein* formulated an explanation of Brownian motion and *Perrin* confirmed this explanation with his experimental work. Had someone else formulated the explanation or performed the experimental work, scientists would not believe in atoms.

Then it follows that (1) and (2) together (i.e. “Had there been no explanation or experimental work OR had these works been produced by someone else than Einstein and Perrin, scientists would not believe in atoms”), provide a deeper explanation than either of them alone.

According to Hitchcock and Woodward “[this] is, perhaps, the most fundamental way in which one [collection of counterfactuals] can provide a deeper explanation than another.”<sup>33</sup> The more factors we add, the more complete an explanation we have. Therefore, complete explanations are the deepest explanations.

Finally, notice that this dimension of explanatory depth is able to explain why we often want to know the intermediary events that bring about the outcome without making our ability to explain dependent on such knowledge. If we know how the effect of Z on X is mediated, we are able to answer more what-if questions and therefore have a deeper explanation. This means that our ability to explain does not depend on the knowledge of the intermediate steps but often the knowledge about those steps provides us with a deeper explanation.<sup>34</sup> Yet, it must be noticed that simply adding intermediate steps between Z and X does not make an explanation deeper. One must specify what changes in these intermediate steps would have led to Y rather than X. As Ereshefsky and Turner put it,

The more the narrative is filled in, that is, thickened, with information about intermediary events that bring about the outcome, the stronger

<sup>33</sup> “Explanatory generalizations, part II: Plumbing explanatory depth”, 188.

<sup>34</sup> It is clear that, in many cases, we are able to explain without knowing the details of a historical process. For example, when the attacking force is much more powerful than the defensive force, the victory of the attacking force can be explained by citing the relative strengths of the forces and the details of the battle do not add much to our understanding of the outcome.

the historical explanation. [...] The mere introduction of any factor does not make a historical narrative stronger. Only citing factors that are *causally relevant* to the outcome explained make a narrative stronger.<sup>35</sup>

The point here is that a fully developed account of historiographical explanation should be able to explain (i) why certain historiographical practices, such as thickening of a narrative, are fruitful and (ii) on the basis of (i) tell when such practice has made real progress. The counterfactual account of explanation explains when and why thickening of narratives provides explanatory understanding.

(III) We can now turn to dimensions of explanatory depth that are less straightforward than (I) and (II) but still of a great importance in historiography. First, we may notice that adding more and more factors into an explanation might not add to the depth of the explanation in historiography. In historiography, the historical context (whatever this means) where the *explanandum* is embedded needs to be somehow taken into an account when the *explanans* is formulated. We cannot build every feature of a historical situation into the *explanans* and treat them as equally changeable. This would, in a sense, *tear down the historicity of these situations*. For example, it might not seem right to assume that the ethos of scientists in some era could have been different. The fact that scientists had that ethos might seem to be constitutive of that era. If we wanted to understand why something happened *in the context of that era*, it would be a distortion to provide a scenario where the ethos would have been different. For example, if we added to our explanation above that

Had the scientific community lived in an Aristotelian framework, Einstein's and Perrin's work would not have had any effect.

we arguably would not have achieved a deeper explanation than the ones we already have despite our being able to answer one more what-if question.

The third dimension of explanatory depth is important if we follow the advice that historiographical counterfactuals should not “unduly disturb what we otherwise know about the original actors and their beliefs and goals.”<sup>36</sup> The third dimension prevents us from adding explanatory depth by giving answers

35 Marc Ereshefsky and Derek Turner, “Historicity and explanation”, *Studies in History and Philosophy of Science Part A* (2019) (Emphasis added).

36 P. Tetlock and A. Belkin, *Counterfactual Thought Experiments in World Politics: Logical, Methodological and Psychological Perspectives*, (Princeton University Press, 1996), 23.

to what-if questions that seem irrelevant, given the nature of the historical era under discussion. (See also the discussion of the minimal-rewrite rule below.)

(IV) Another interesting dimension is the following. We can argue that E is deeper than E\* if it is the case that

E would still be true while E\* would no longer be true, if there was a change in background conditions, due to the conceptualization of *explanantia*.<sup>37</sup>

Assume that 1905 was a somewhat boring year, and therefore Einstein's explanation of the Brownian motion was the most noteworthy event of the year.

Now, if E says:

Had Einstein not explained the Brownian motion, scientists would not have believed that atoms exist.

and E\* says:

Had the most noteworthy event of the year not happened, scientists would not have believed that atoms exist.

then E is deeper than E\*. This is due to the fact that E would still be true if the iPhone had been introduced in 1905, while E\* would no longer be true. The introduction of the iPhone would have been the most noteworthy event in 1905. However, if the iPhone had been introduced, it would no longer be true that had the most noteworthy event not happened (i.e. had the iPhone not been introduced) scientists would not have believed that atoms exist. E\* is therefore fragile with respect to changes in the background conditions. We should prefer E as it provides (or, at least, building materials for) a better *possible-cause hypothesis*<sup>38</sup> than E\*.<sup>39</sup>

37 This is not from Hitchcock and Woodward "Explanatory generalizations, part II: Plumbing explanatory depth". Instead, it is based on considerations in Woodward, *Making Things Happen*, 217.

38 Woodward uses the term, *generalization*. However, these generalizations that tell "[Z] is among the possible causes of [X]" (Woodward, *Making Things Happen*, 214), should not be confused with explanatory generalizations that give somewhat detailed information about the relationship between the exact values of Z and X. Therefore, I choose to use the term *hypothesis*.

39 See Woodward, *Making Things Happen*, 217.

This is an interesting dimension of explanatory depth since it is possible to find competing claims within a single historiographical study that can be distinguished only with respect to their explanatory depth. This means that the concept of explanatory depth turns out to be a rather fruitful historiographical tool.

Consider an illustration: In the paper “Inspiration in the Harness of Daily Labor: Darwin, Botany, and the Triumph of Evolution, 1859–1868”, Richard Bellon describes the process that led to the acceptance of Darwin’s evolutionary ideas in the British scientific community.<sup>40</sup> Bellon claims that the publication of *On the Origin of the Species* in 1859 was not a decisive event in this process. On the contrary, the book was judged to be speculative in character and against the scientific and moral standards of the community. Only after publishing a study on orchid fertilization, in which he applied his evolutionary ideas, was Darwin’s framework meet with approval by the community.

Bellon is providing an answer to the question:

Why were Darwin’s ideas accepted by the scientific community rather than not?

Bellon gives the following answer:

The *Origin* was packed with evidence, but it communicated few entirely original scientific observations, and this allowed its critics to dismiss it as vainglorious speculation untethered from the manly discipline of original discovery. [...] If Darwin had not tied the theory of the *Origin* to productive, technical, and specialized research – in the event, his floral biology, but it could have been any number of things – the ultimate reception and received meaning of the *Origin* would have been significantly different. [...] Science’s social, political, and religious respectability depended on the governance of imagination by consistently patient and humble behavior and [...] Darwin’s adversaries frequently used this ideology to bludgeon the *Origin*. Ultimately, Darwin vanquished his foes by reversing the weapon and claiming the mantle of heroic conduct for himself and his theory.<sup>41</sup>

40 Richard Bellon, “Inspiration in the Harness of Daily Labor: Darwin, Botany, and the Triumph of Evolution, 1859–1868”, *Isis* 102 (2011), 393–420.

41 *Ibid.* 395–396.

It may seem that the explanation is a rather straightforward one, at least for a philosopher of science.<sup>42</sup> The explanation of the acceptance of the evolutionary framework can be formulated as follows:

(E1) Had Darwin not shown in *Orchids* that his framework was successful and progressive (i.e. *fruitful*) in empirical research, the evolutionary framework would not have been accepted.

Yet, there is a complication in the issue. Bellon also points out that before the *Orchids* was published, Darwin's work was dismissed on *moral standards*. Darwin did not seem to work on a patient, humble and honest manner and, therefore, was not taken seriously.<sup>43</sup> Only after the *Orchids*, was Darwin accepted as a morally righteous scientist, and therefore the fulfillment of Victorian moral standards was an important factor in the acceptance of his theory. As Bellon points out, the requirements for novel discoveries and progressive research program (i.e. the requirement of fruitfulness) were an embodiment of the Victorian value system and moral standards.<sup>44</sup>

This explanation for the acceptance of Darwin's theory can be framed as follows:

(E2) Had Darwin not lived in accordance with the values of Victorian society, his evolutionary framework would not have been accepted.

Even though "living in accordance with the values of Victorian society" and "showing that a framework is successful and progressive" refer to the same thing in Darwin's case, we can still choose one explanation over the other due to considerations of explanatory depth. We could reason as follows:<sup>45</sup>

There is an asymmetry between E1 and E2. There are important changes in background conditions that would have made E2 false. For example, nothing in Bellon's paper indicates that "had Darwin done something morally wrong

42 This simplified form of analysis can be made, for example, by following Imre Lakatos, *The Methodology of Scientific Research Programmes* (Cambridge University Press, 1978); Larry Laudan, *Progress and its Problems: Toward a Theory of Scientific Growth*, (University of California Press, 1977); or Thomas Kuhn "Objectivity, value judgment, and theory choice" in *The Essential Tension*, (University of Chicago Press, 1977), 322. Needless to say, the discussion about the details of confirmation and the acceptance of theories and research programs has been enormous.

43 Bellon, "Inspiration in the Harness of Daily Labor," 403–407.

44 *Ibid.*, 395.

45 For the sake of the argument at least and without commitment to the truth of the conclusion.

that was not related to his scientific work, his theory would not have been accepted." In other words, it seems possible<sup>46</sup> that had Darwin done something morally wrong (perhaps some minor misbehavior) and produced his theory, his theory would still have been accepted. Moreover, it seems possible that had Darwin lived in another society, his theory would still have been accepted even if he did not follow the Victorian value system.

However, it is difficult to figure out any change in background conditions that would have made E1 false. What should we add to the claim "Had Darwin not proved his theory fruitful and progressive, it would still have been accepted" to make that claim true? Of course, we can say that had Darwin lived in a society that considered fruitfulness as a reason to reject a theory, his theory would not have been accepted. However, it seems difficult to tell what kind of change this would have required. The claim seems to have the form "Z would not have been the cause of X if Z occurred in the context where Z does not cause X," and thus lacks informative content.

One way to look at the difference between E1 and E2 is that one can abandon Victorian moral standards and still perform fruitful scientific work. However, it is not possible to abandon the standard of fruitfulness and still live in accordance with Victorian moral standards, as the standard of fruitfulness is an embodiment of the moral standards. The antecedent of E2 is never true when the antecedent of E1 is not true. However, the antecedent of E1 can be true even if the antecedent of E2 is not true. Therefore, following our argument-for-the-sake-of-illustration, E1 is deeper than E2 and is therefore to be preferred to E2.

In this section, we have seen how explanatory depth can be approached and how that notion can be used to assess the relative merits of different explanations and to clarify the role of certain explanatory activities (such as the thickening of narratives). The notion makes it possible to set the threshold of explanatoriness rather low (on minimal explanations) without blurring the fact that not every explanation is equally good. Explaining something is easy; explaining something in depth is difficult.

## 5 Intervention, Counterfactual History and Contingency

It is a well known fact that not just any counterfactual dependency is explanatorily relevant. For example, a storm would have occurred, had the reading of a barometer been low; however, the reading of the barometer is not explanatorily relevant.

---

<sup>46</sup> On the basis of Bellon's paper.



Woodward's account of explanation is often called 'interventionist' because it requires that explanatorily relevant counterfactual dependencies must remain invariant under an *intervention*.

The intuitive idea is that an intervention on X with respect to Y changes the value of X in such a way that if any change occurs in Y, it occurs only as a result of the change in the value of X and not from some other source.<sup>47</sup>

Consider that we ask, "Why did a storm occur?" We could answer: "Because the barometer reading fell. Had it not fallen, there would not have been a storm." This answer does not provide us with an explanation even though it answers a what-if-things-had-been-different question. The relevant answer would be: "Because the atmospheric pressure fell. Had it not fallen, there would not have been a storm." The difference between these answers is that the second one is true if an intervention on the atmospheric pressure was performed, while the first one is false if an intervention on the barometer was performed: if we hold the needle of the barometer fixed, the storm would have occurred anyway (as the atmospheric pressure fell).

Suppose someone argues that Einstein's work does not explain the belief in atoms: had Einstein not provided an explanation, because someone else already had, scientists would have come to believe in atoms. This does not mean that Einstein's work does not explain the belief, since *someone else explaining the Brownian motion* would be a cause of the belief and therefore does not count as an intervention on Einstein's work with respect to belief in atoms. A more tricky case is the following.

Suppose we remove Einstein's work from history by intervention. For example, he could have had an injury causing brain damage. However, if it was the case that someone else was working to explain the Brownian motion and succeeded in explaining it, then scientists would still have believed in atoms. It seems that the belief is not dependent on Einstein's work and therefore the work does not explain the belief. However, the interventionist account is able to deal with such cases. In the account, a factor is explanatorily relevant if a change in that factor changes the outcome *once we fix other factors in a certain way*. For example, Einstein's work explains the belief because it is true that had Einstein not provided an explanation AND had we fixed other things so that no one else provided an explanation, scientists would not have come to believe in atoms.

In the case above, we were lucky in the sense that we needed to fix factors to their actual states (*no one else provided the explanation at the time* is a part of the actual course of history). However, sometimes a factor is explanatory

<sup>47</sup> Woodward, *Making Things Happen*, 14. See the details in *ibid.*, 98.

because changing it changes the outcome, given that we fix other factors to their non-actual states. For example, someone does not attempt to make an observation O because she knows that her instrument is not accurate enough. It is not true that had the person attempted to make the observation, she would have made the observation. However, it is true that had the person attempted to make the observation AND had her instrument been more accurate, she would have made the observation. This is an explanation for “why did she not observe P rather than did?”

In cases where some factor is fixed, a minimal explanation must have the form “had W rather than Z been the case AND had F (the factor we fixed) been the case, Y rather than X would have been the case”. W alone would not have led to Y, but W and F would have.

Notice the following:

(1) Interventionism allows us to be precise in the way in which the antecedent of a counterfactual conditional is changed but does not require us to be too precise. The only requirements for the change are that (i) the change satisfies the definition of intervention, (ii) the change in the antecedent is the one that is supposed to be explanatory according to the one who formulates the explanation. This also means that interventionism provides tools to clarify obscure explanatory claims. For example, Lebow and Stein argue that the truth-value of “had the United States attacked the missile bases, the Soviet Union would have responded to an attack on Cuba with military action of its own” can be assessed only after we specify what kind of attack is in question.<sup>48</sup> Interventionism gives us the tools to clarify the claim: the one who thinks that the counterfactual is true and explanatory must, in the interventionist account, provide a scenario where intervention I is performed and tell what kind of attack was launched as the consequence of I. It should not be too difficult to provide such specification, given that interventions need to be only conceptually possible.<sup>49</sup> For example, someone could say that “had the United States decided to launch a nuclear attack ...” is the intervention she thinks makes the counterfactual true. If this is the case, then one is committed to a very specific explanatory claim “had there been a *nuclear attack* ... then the Soviet Union ...”, not the claim that any attack whatsoever would have led to a Soviet response.

It is very important to notice that one does not need to care whether a nuclear strike was really considered as an option by those in charge. As long as it is true that an intervention causing a nuclear strike would have led to an attack

48 R. N. Lebow and J. G. Stein, “Back to the Past: Counterfactuals and the Cuban Missile Crisis,” in *Counterfactual Thought Experiments in World Politics*, eds. Tetlock and Belbin (Princeton University Press, 1996), 119–48.

49 See Woodward, *Making Things Happen*, 132.

by the Soviet Union, we have an explanation for why the Soviet Union did not attack: The Soviet Union did not attack rather than attacked because there was no nuclear strike rather than there being a nuclear strike.

This brings us to an important issue, to the distinction between explanatory counterfactuals and counterfactuals that are close to the actual course of events.

(2) The so-called “minimal-rewrite-of-history” rule says we should avoid far-fetched counterfactuals in historiography and consider only those counterfactual situations that could have happened, given the actual history.<sup>50</sup> In the interventionist account, this rule is rejected when explanatory relevance is the concern. “X rather than Y because Z rather than W” is explanatory if W would have led to Y no matter how far-fetched W is. If a (conceptually possible) intervention makes W to be the case and if Y follows, we have an explanation. It might seem that this shows that interventionism does not capture the nature of historiographical explanation – we are supposed to respect the principles of historiography, such as the minimal-rewrite rule. However, I would like to turn the tables and argue that the minimal-rewrite rule is not connected with explanatory relevance *per se* but rather serves a methodological and (quasi-) metaphysical role. The minimal-rewrite rule can be diagnosed within a fully developed account of explanation.

First of all, the minimal-rewrite rule is a useful methodological rule. It might be too difficult to evaluate a counterfactual citing a far-fetched change in history because we have no cases to which we could compare the counterfactual. “Had Jill been able to punch twice as hard as George Foreman, she would have been a successful boxer (despite her bad skills)” is difficult to assess because we have never seen anyone punch like that. Would pure power be enough to compensate for lack of skills? Similarly, it is difficult to assess what would have happened if there had been a nuclear strike on Cuba. Would it have led to a nuclear war? Actual history does not provide us with cases to which we could compare our counterfactual situation (perhaps there was a plan for such a situation but plans do not always materialize, especially in surprising situations).

Secondly, the minimal-rewrite rule is directly connected to the issue of historical contingency. If a historiographical explanation is requested to show the contingency of events, as is sometimes argued, then the minimal-rewrite rule guarantees that we are able to show exactly that. For example, Beatty argues that

Narrative explanations are more demanding than counterfactual, difference-making explanations more generally, in that they go beyond

50 Tetlock and Belkin, *Counterfactual Thought Experiments in World Politics*, 23–25.

“what-if-things-had-been-different questions” to indicating *that things could very well have been different*.<sup>51</sup>

Consider the following famous definition by Ben-Menahem:

Contingency: Similar causes lead to different types of effects. High sensitivity to initial conditions.<sup>52</sup>

Or a more recent suggestion by Virmajoki (in the context of contingency of science):

It could have been the case that science has the feature  $F^*$  rather than the actual feature  $F$ , where the difference between  $F$  and  $F^*$  is considered interesting in the given context of discussion.

We can accept that  $F^*$  could have been the case if and only if (a) someone points out a counterfactual past event  $Z$  that would have led to  $F^*$ , and (b) it is shown that the occurrence of  $Z$  in the past is not an impossible (or extremely far-fetched) scenario.

Feature  $F$  can be judged to be (a) inevitable if and only if the occurrence of any  $Z_i$  is seen as impossible and (b) a truly chancy feature if  $Z$  is a part of the actual history. Moreover, the more far-fetched the occurrence of  $Z$  is judged to be, the more inevitable feature  $F$  is.<sup>53</sup>

Given these definitions of contingency, an explanation showing that

- (i)  $Y$  rather than  $X$  would have been the case, had  $W$  rather than  $Z$  been the case, and
- (ii)  $W$  was similar to  $Z$  (following Ben-Menahem) or that  $W$  was not far-fetched (following Virmajoki)

implies that  $X$  was a contingent event. This means that if we only consider historical situations that are close to the actual history and still find when  $Y$  rather than  $X$  would have been the case, we have shown the contingency of  $X$ . The minimal-rewrite rule therefore makes the phenomena of historical contingency and historiographical explanation inseparable. It seems that the alleged

51 John Beatty, “Narrative possibility and narrative explanation”, *Studies in History and Philosophy of Science Part A* 62 (2017), 31–41 (emphasis added).

52 Yemima Ben-Menahem, “Historical Contingency”, *Ratio* 10:2 (1997), 101.

53 Veli Virmajoki, “Could Science Be Interestingly Different?”, *Journal of the Philosophy of History* 12:2 (2018), 311.

plausibility minimal-rewrite rule is based on its connection to historical contingency. If we want our explanations to reveal the contingency of an historical event, then the rule is justified within an account of explanation.

However, it seems that these two phenomena, historical contingency and historiographical explanation, should be kept separated. First, it seems arbitrary to limit the *explananda* of historiography to contingent events and processes. Some interesting events and processes are probably rather inevitable but we would nevertheless like to have explanations for them.<sup>54</sup> Secondly, it is difficult to tell whether similar or plausible initial conditions would have led to a different outcome if we do not already know on what conditions the outcome depended. We need to have explanations *per se* before we are able to find a contingency-indicating explanation. The minimal-rewrite rule therefore serves important historiographical functions,<sup>55</sup> but we should not make it a defining feature of historiographical explanations. The (interventionist) counterfactual account of explanation is able to incorporate both these facets of the minimal-rewrite rule: it defines the notion of *explanation* without a reference to the rule but still recognizes its methodological and contingency-indicating roles.

Finally, it is important to make note of how the contrastive explanations based on counterfactuals differ from what is known as *study of counterfactual histories*. In these studies, the event C is assumed to be the cause of the event E and the question one tries to answer is: What would have happened had C not occurred,<sup>56</sup> or more generally, what could have happened had certain things been different in the past? On the other hand, the formulation of a contrastive explanation begins by specifying the relevant alternatives to the event E and proceeds to find out which alternatives (C\*) of C would have led to some relevant alternative (E\*) to E. One important difference is also that counterfactual histories do not need to respect “the logic of fixing other conditions” (see above) and therefore they are not automatically explanatorily relevant.

54 Sterelny calls explanations that show the resilience of an outcome “robust process explanations” (Kim Sterelny, “Contingency and History,” *Philosophy of Science* 83:4 (2016), 521–539). This important class of explanations would be impossible if we thought that an explanation must reveal the contingency of the *explanandum*. See also Virmajoki, “Could Science Be Interestingly Different?,” discussing historical contingency along the lines of the counterfactual account of explanation in more detail than we are able to do in this paper. The paper describes what considerations need to be added to the counterfactual account in order to get judgements of contingency off the ground.

55 See also discussion in Section 5 on the dimension of explanatory depth (111).

56 See Alexander Maar, “Possible uses of counterfactual thought experiments in history,” *Principia: An International Journal of Epistemology* 18:1 (2014), 87–113.

For example, we might say that “had Einstein not explained the Brownian motion, scientists would still have believed in atoms as someone else would have provided the explanation.” This might be a respectable claim in the study of counterfactual history. However, it does not imply, as we saw, that Einstein’s work is not explanatorily relevant. It is therefore extremely important to keep in mind that the counterfactual account of explanation does not imply that the study of counterfactual histories, as that practice is actually conducted, is at the heart of explaining.

## 6 The Choice of *Explanandum*

Jo Guldi has recently pointed out that the choice of *explanandum* in historiography has not received the attention it deserves.<sup>57</sup> Guldi points out that these choices are often made on the basis of canons of the given historiographical field but adds, correctly it seems, that “questions of significance [of *explananda*] are far from being settled.”<sup>58</sup> In this section, I discuss which considerations are relevant in the choice of *explananda*.

First, it might be argued that we must explain those historical events and processes that are the most relevant in the historical developments. For example, one could argue that we want to know why scientists came to believe in atoms because *that they believe in atoms* is an important explanatory factor of later developments. This way of choosing the *explananda* is quite natural and carries a promise of tightly connected historiographical narratives. However, the choice of *explananda* on the basis of their causal significance does not get off the ground before we decide which of the later developments require an explanation. Every historical event belongs to some causal chain and is therefore relevant to some historical development. If every historical event is part of some causal chain, then *being a part of a causal chain* cannot be the criterion for significance. We need to ask further which of those causal chains are significant. It seems that this choice must, ultimately, be based on our values. A historical event may receive its significance from a causal chain it belongs to only if at least some link in that chain is judged to be significant.

How such judgements based on values can be meaningfully made cannot be discussed here.<sup>59</sup> However, it must be noted that, when evaluating the

57 Guldi focuses on historiography of science, but her insight applies to all historiography. Jo Guldi, “The Modern Paradigms of Explanation: Significance, Agency, and Writing History in the Era of Climate Change”, *Isis* 110:2 (2019), 346–353.

58 *Ibid.*, 351.

59 I have discussed the issue in *Cementing Science*, Chapter 4.

significance of “why X rather than Y”, both the choice of X and the choice of the contrast must be carefully examined. It is not enough that X is significant but Y must also be a somehow interesting alternative to X. For example, of the two questions

- (I) Why did scientists come to *believe* that atoms exist rather than have no beliefs at all?
- (II) Why did scientists come to believe *that atoms exist* rather than believe that atoms do not exist?

the latter seems more significant. One could argue, for example, that we are more interested in knowing how we could have had different science than we are in knowing when there would not have been science at all. In such cases, the difference in significance lies in the contrasts. Notice that if we already knew which causal chains are significant, then the choice between (I) and (II) depends on their explanatory relevance with respect to some further *explanandum* E. To take a toy-example, if we wanted to know (E) why scientists claim that atoms exist rather than claim that they do not really exist, then (II) is a significant *explanandum* from the causal point of view because it explains this E. But again, E must be judged as significant independently of causal considerations if we are to judge (II) as significant because it explains E.

Finally, one could perhaps argue that the whole question about significance is mistaken since it has been claimed that historiography mainly explains evidence, not events.<sup>60</sup> Even though evidence undoubtedly is the basis of all historiography and inference to the best explanation an important epistemological tool, this approach puts the cart before the horse. In most cases, the evidence is gathered in order to find an explanation for something.<sup>61</sup> The notion of evidence makes sense only if one has a particular problem in mind, given the number of texts, artifacts, natural objects and so on that could serve as evidence in some possible historiographical inquiry. Moreover, it would be an extremely sorry state of affairs if historians were to focus on explaining their evidence, (i.e. if evidence was the most significant thing to be explained). Surely, historiography can answer (and in fact answers) many questions that stem from our concerns as human beings. How did we end up burning massive amounts of fossil fuels? Of course, if it is thought that everything that indicates that we are burning massive amounts of fossils is historiographical evidence, then we are “explaining evidence” here. However, in that case the whole notion of “evidence in historiography” loses its value. “Explaining evidence” would be equal to “explaining a feature of the current world.” And, again, we should ask,

60 Tucker, *Our Knowledge of the Past*, 186.

61 Of course there are cases where some materials (for example, an old text) are found and historians attempt to understand their properties.



“Which current features of the world are so significant that they require a historiographical explanation?” There is no escape from value-laden judgements of significance in historiography.

## 7 Conclusion: Balancing Epistemology and Explicatory Power

In this paper, I have argued that a fully developed account of explanation in historiography should explicate notions such as *minimal explanation*, *complete explanation*, *historiographical explanation*, *explanatory depth*, *explanatory competition*, and *explanatory goal* and clarify the choice of *explanandum*. Moreover, such an account should be able to clarify many explanation-related issues on the basis of the explications. I have pointed out how a counterfactual account of explanation is able to explicate and clarify such notions and issues. No doubt, other accounts may also be able to do so, but the details have to be worked out.

However, my aim is not to suggest that the account should be accepted on the basis of considerations in this paper alone. Rather, my aim has been to indicate what kinds of considerations are relevant when an account of explanations is discussed in the philosophy of historiography. The considerations I have pointed out are probably not the only ones. There might very well be additional notions and explanation-related issues that an account of explanation in historiography should be able to explicate and clarify. Be that as it may, serious discussions about the nature of historiographical explanation can begin only after we recognize just how many explanation-related notions and issues an account of explanation must be able to explicate and clarify.

Of course, the more *desiderata* for an account of explanation we recognize, the more difficult it becomes to satisfy all of them. Discussions about the relative importance of *desiderata* are needed. This can already be seen in relation to the counterfactual account of explanation: The price for explicatory power is rather complicated epistemology. It is well known that there cannot be direct evidence for counterfactual situations. This means that it is possible, from the perspective of the counterfactual account, that we understand different kinds of explanation-related notions but are unable to tell on the basis of the evidence when the conditions mentioned in those notions are satisfied. However, there is no reason to fall into pessimism here. There have been many discussions about the epistemology of historiographical counterfactuals.<sup>62</sup> On

62 See, for example, Maar, “Possible uses of counterfactual thought experiments in history”; P. Tetlock, R. N. Lebow, and G. Parker, eds., *Unmaking the west. “What-if” scenarios*

the basis of these discussions, it seems to be possible to work out the epistemological underpinnings of counterfactual explanations. Moreover, I take it as a healthy sign that the counterfactual account leaves open the possibility that (at least sometimes) we cannot provide an explanation for an event due to our epistemic situation. This indicates that successful explanations track connections in the world rather than project ones onto it and explains why serious historiography is so difficult and easy answers so deceiving. There is no value in explaining if success is guaranteed. Be that as it may, only the future can tell to what extent any fully developed account of explanation is able to balance different *desiderata*. We first need to formulate those accounts.

Finally, the fact that Woodward's account is a general one (in that it aims at capturing the logic of explaining in many fields) and the applicability of the account in historiographical explanations suggests that historiographical explanations do not differ fundamentally from explanations in other fields. I do not think there is anything strange in this conclusion. The similarities between historiographical explanations and explanations in other fields do not imply that we do not need an account of historiographical explanation. As we saw throughout this paper, there are issues in historiographical explanation that can be captured only by modifying and developing further a more general account of explanation. Moreover, *a successful account of explanation in historiography simply is one that is able to capture and clarify explanation-related notions and issues that arise in historiography*. It makes little sense to say that an account of explanation that is successful in these tasks is not an account of explanation in historiography because it is applicable to other fields as well. In order to argue along such lines, one should assume that historiographical explanation is *sui generis* and it seems difficult to understand how that assumption could be justified in a situation where we have a powerful and fully developed account that tells otherwise. Philosophy of historiography does not lose its relevance if it turns out that historiography is, in certain respects, similar to other fields. Rather it gains relevance by being able to show how philosophical analysis of historiography can help to clarify issues that are present in many fields. As we saw in this paper, historiographical explanation can be understood along the lines of a general account of explanation and, at the same

---

*that rewrite world history*. (The University of Michigan Press, 2006); Tim De Mey and Erik Weber, "Explanation and Thought Experiments in History," *History and Theory* 42:1 (2003), 28–38; Johannes Bulhof, "What if? Modality and history," *History and Theory* 38:2 (1999), 145–168; M. Bunzl, "Counterfactual history: A user's guide," *The American Historical Review* 109 (2004), 845–858; R. N. Lebow, "What's so different about a counterfactual?" *World Politics* 52 (2000), 550–585; Virmajoki, "Could Science Be Interestingly Different?"; and the issue in *Journal of the Philosophy of History* 10 (3).

time, we saw how that general account can be understood from a different angle by examining historiography. Philosophy of historiography should not be seen as a passive application of more general philosophical frameworks but as a field that is equal in its contribution to the quality of those frameworks.

### Acknowledgements

I thank anonymous referees for valuable comments on the previous versions of this paper.

AUTHOR QUERY:

AQ: Please provide with Position, Department, Faculty, City, Country if possible.