

# Inference to the Best Explanation and Norton's Material Theory of Induction.

June 16, 2020

## Abstract

I argue that we should consider Norton's material theory of induction as consisting of two largely independent claims. First, there is the claim that material facts license inductions - a claim which I interpret as a type of *contextualism* about induction. Second, there is the claim that there are no universal rules of induction. While a good case can be made for the first claim, I believe that Norton's arguments for the second claim are lacking. In particular, I spell out Norton's argument against the claim that all induction may be reduced to inference to the best explanation, and argue that it is not persuasive. Rejecting this part of Norton's theory does not however require us to abandon the first claim that material facts license inductions. In this way, I distinguish the parts of the material theory of induction we should happily accept from the parts about which we should be more skeptical.

## 1 Introduction.

Norton's material theory of induction, developed in works such as [11] and [12], is a very welcome addition to the literature on induction. It prompts a return to basic questions about epistemology and scientific method that have perhaps been sidelined as more technical inquiries have stolen the limelight in the literature on induction in recent years. Although there are things to agree with in Norton's material theory of induction, there are some issues it raises that require closer scrutiny. Such scrutiny, I think, leads to conclusions a little different from those Norton endorses, and it will be the goal of this paper to explain them.

In §2 of this paper I will describe Norton's material theory of induction in broad terms, then divide it more precisely into two independent parts. The general claim of the paper will be that while the first part is sound,

the second part is not. Because the second part of Norton’s theory involves the rejection of many traditional approaches to induction by arguing against them individually, to cast doubt on the second part of Norton’s theory it will be enough to point out problems with Norton’s attacks on just one of these traditional approaches to induction. In this paper, I will do so by re-considering the idea, perhaps first introduced by Harman in [6], that we should view all induction as inference to the best explanation (henceforth, IBE). To this end, in §3 I present a brief discussion of IBE. In §4 I then both present and criticize Norton’s arguments against IBE. In §5 I make some concluding remarks, suggesting how in light of these criticisms Norton’s material theory of induction should be modified to make it more plausible.

## 2 The Material Theory of Induction.

Although in broad outline the basic idea of the material theory of induction seems clear enough, stating the theory carefully is trickier than it might first appear. In this paper, we will largely focus on the case of enumerative induction, in which a generalization is inferred from a set of instances:

$$(*) \quad \frac{\phi(a) \quad \phi(b) \quad \phi(c) \quad \dots}{(\forall x)\phi(x)}$$

While not every inference of the form (\*) is justified, some are (at least, relative to some given background knowledge.) For example, the observation of many grue emeralds does not justify the belief that all emeralds are grue, and thus not every inference of the form (\*) is justified. However, the observation of sufficiently many green emeralds generally *is* taken to justify the belief that all emeralds are green (given our ordinary background knowledge), and thus some inferences of the form (\*) *are* justified.

The question of what makes the justified instances of (\*) justified is of course a notoriously difficult question. Norton’s idea is that we have not been able to find a fully general justification of those inductions that are justified because no fully general justification exists. Instead, Norton tells us that whether an inference of the form (\*) is justified simply depends on the context and subject matter. This context dependence is indicated by something Norton calls a *material fact* that *licenses* the inference (\*). So for example, it is in virtue of the fact that elements have stable melting points that we may infer from just a few careful observations of bismuth samples melting at 271°C that all bismuth samples melt at 271°C. The fact that in ordinary conditions elements have stable melting points is here the *material fact* that licenses the inductive inference in question.

Norton’s claim is that for each justified induction we perform, there is some material fact that ‘licences’ the inference. Without such a ‘licensing

fact' obtaining, we are not justified in passing from a set of instances to a universal generalization. However, Norton thinks that there is no 'universal' material fact that licenses all justified inductive inferences of the form  $(*)$ <sup>1</sup> - rather, the material fact varies from subject matter to subject matter. In this way, the explanation of why a given inference of the form  $(*)$  is justified is *context dependent*. This is perhaps the core idea of Norton's material theory of induction.

The question of whether Norton requires the reasoner to have justified belief in this material fact is exegetically tricky. One can, after all, distinguish an internalist version of Norton's theory in which the reasoner is required to have justified belief in any material facts on which her inductions depend, from an externalist version of Norton's theory in which the mere truth of the material fact (whether known or not) is sufficient for an inductive inference to be justified. In his more recent work Norton has been clearer that the material theory is to be understood externalistically. For instance, in the most recent draft of [12] Norton tells us that inductive inferences are warranted '*by matters of fact that obtain independently of any human beliefs, knowledge or awareness.*' Of course, externalist theories in epistemology bring with them a standard set of worries, and whether Norton has given a good set of replies to these is not clear to me. Nevertheless, such matters will not be the subject of this paper, and so although we will read Norton's theory externalistically we will largely put to the side worries that apply to externalist epistemologies quite generally.

In [12], Norton tries to capture his theory with many slogans. He tells us that '*all induction is local*' that '*there are no universal rules for inductive inference*', that inductive inferences are '*warranted by facts, not formal schema*' and that induction is '*contextual*'. But each of these slogans says something slightly different, and so we must proceed with caution here.

To this end, let us begin by acknowledging that Norton's theory begins on fairly safe ground. It is surely right that any justified instance of  $(*)$  will be justified only in virtue of something that goes beyond the premises. Even in the best cases, it is not *merely* in virtue of having justified belief in the premises  $\phi(a), \phi(b), \phi(c), \dots$  of  $(*)$  that a reasoner is justified in inferring  $(\forall x)\phi(x)$ , even if only because that would seem to open up the possibility of inferring the grueness of all emeralds from the grueness of all observed emeralds. Something additional therefore seems to be required for inferences of the form  $(*)$  to be justified. Norton thinks of this additional thing as the obtaining of a 'material fact'. Although questions can be raised even about all this, thus far Norton seems to be on relatively firm ground. The further idea that such material facts might vary from case to case is not implausible,

---

<sup>1</sup>Norton rightly rejects, for example, Mill's appeal in [9] to an 'axiom of the uniformity of the course of nature' that might serve this purpose.

and so slogans such as ‘induction is contextual’ are not too hard to swallow. Perhaps even the externalist idea that one might not have to believe such material facts for one’s inductions to be justified involves only a small step from this relatively firm ground.

However, at least some of Norton’s slogans say something significantly stronger than all of this, as we shall see by the end of this section. Note that in effect what Norton has done is replaced the relation of *justified inferability* that was supposed to hold between the premises  $\phi(a), \phi(b), \phi(c), \dots$  and conclusion  $(\forall x)\phi(x)$  of a justified inductive argument with a more complex relation of *licensed justified inferability* that holds between some material fact  $M$ , the premises  $\phi(a), \phi(b), \phi(c), \dots$  and the conclusion  $(\forall x)\phi(x)$  of the inductive argument. To put the point slightly differently, for Norton the relation of justified inferability is now *contextual*, and the material fact provides a sort of contextual parameter. But the natural thing to wonder now is what sort of account might be given of this licensed justified inferability relation. What does it mean to say that some fact  $M$  *licenses* the inference from the premises  $\phi(a), \phi(b), \phi(c), \dots$  to the conclusion  $(\forall x)\phi(x)$  of an argument, rendering it justified?

One view strenuously rejected by Norton is that this relation of licensed justified inferability holds just in case  $M, \phi(a), \phi(b), \phi(c), \dots$  deductively entails  $(\forall x)\phi(x)$ . Norton also rejects Bayesian analyses of this relation. Other approaches that involve analysing the relation of licensed justified inferability in terms of explanations or simplicity are also rejected by Norton. (See §3 of [11] and [13] for individual discussions of each of these approaches, as well as discussions of further approaches still.)

Norton’s rejection of these accounts of the relation of licensed justified inferability is connected with his rejection of ‘universal rules’, ‘schema’, or ‘templates’ for induction. For we can easily transform any account of the licensed justified inferability relation into a universal schema for inductive inference by saying (\*\*):

(\*\*) Infer  $(\forall x)\phi(x)$  from  $\phi(a), \phi(b), \phi(c), \dots$  iff for some material fact  $M$  the relation of licensed justified inferability holds between  $M$  and  $\phi(a), \phi(b), \phi(c), \dots$  and  $(\forall x)\phi(x)$ .

where the phrase ‘relation of licensed justified inferability’ in (\*\*) is to be replaced with the account in question. In this way, any single, unified analysis of the licensed justified inferability relation leads to a universal schema of induction, contrary to Norton’s slogan that that ‘*there are no universal rules for inductive inference*’. Thus, Norton’s rejection of universal schema requires him to reject any sort of reductive analysis of the relation of licensed justified inferability on which his theory is based.

To flesh out this point a little further, let us say that an *internalistic*

*schema* for inductive inference is a rule that tells us that if (i) we are justified in believing the premises  $\phi(a), \phi(b), \phi(c), \dots$  of (\*), and if (ii) we are justified in believing particular further facts  $F_1, \dots, F_i$  (which are determined in some way from  $\phi(a), \phi(b), \phi(c), \dots$ ), then we are justified in believing the conclusion  $(\forall x)\phi(x)$  of (\*). We also say that an *externalistic schema* for inductive inference is a rule that tells us that if (i) we are justified in believing the premises  $\phi(a), \phi(b), \phi(c), \dots$  of (\*), and if (ii) we are justified in believing some further facts  $F_1, \dots, F_i$  (determined in some way from  $\phi(a), \phi(b), \phi(c), \dots$ ), and if (iii) some further facts  $G_1, \dots, G_i$  obtain (not necessarily believed with justification by us, but still determined in some way from  $\phi(a), \phi(b), \phi(c), \dots$ ), then we are justified in believing the conclusion  $(\forall x)\phi(x)$  of (\*). Thus, an externalistic schema of induction is simply an internalistic schema of induction with the extra requirement that certain ‘facts’ must (externally) obtain for the schema to hold.

So for example, consider the (loosely stated) idea that in order for (\*) to be justified, its conclusion must provide the best explanation of its premises. If we only require that the reasoner have justified belief that the conclusion provides the best explanation of the premises, then this gives an internalistic schema of induction, with the only  $F$  just being the proposition that the conclusion of (\*) provides the best explanation of its premises. If we only require that it be true that the conclusion provides the best explanation of the premises (and we do not require that this be believed with justification), then this gives an externalistic schema of induction, with the only  $G$  just being the proposition that the conclusion of (\*) provides the best explanation of its premises, and no  $F$ s. Alternatively, consider the Bayesian theory according to which the argument (\*) is justified just in case the conditional probability of its conclusion relative to its premises and our background knowledge is sufficiently high. If we have the view that our degrees of belief and background knowledge are known to us, then suitably formulated this gives an internalistic schema of induction. But if, for example, degrees of belief are viewed ‘objectively’ and we are not required to necessarily know all these degrees of belief to be justified in making the relevant inferences, then suitably formulated this gives a merely externalistic schema of induction.

If Norton’s externalism about induction is right, we should not expect any internalistic schema of induction to capture induction in general, as internalistic schema of induction do not require the external truth of any ‘material facts’ in order to be justified. And so if Norton’s opposition to universal schema of induction is just an opposition to internalistic schema of induction, then Norton’s rejection of universal schema of induction is just a consequence of his externalism. The question I wish to consider then is whether Norton is also opposed to externalistic schema of induction.

As Norton does does not tell us in full detail what exactly a schema for in-

duction is, we must guess here. One might at first think that Norton should have no reason to reject externalistic schema of induction. And yet, Norton goes out of his way to argue against things like Bayesianism and IBE very generally. There is nothing in Norton's arguments or rhetoric that suggests that he would be fine with Bayesianism or inference to the best explanation so long as they were viewed as merely giving externalistic schema for induction - rather, his rejection of them as universal schema seems unconditional and absolute. Norton's general rhetoric in [12] is that the material theory 'relocates' the warrant for induction from 'rules' to 'facts', and there is never any indication that a certain sort of rule might be exempt from this.

But if this is right, then on second pass we must consider Norton's theory as really consisting of two quite different ideas. First, there is the idea that what makes an inductive inference of the form (\*) justified is the (external) obtaining of a material fact, which differs from situation to situation. But second, and quite separate from this, is Norton's idea that there are no universal rules for inductive inference - internalistic or otherwise. That these ideas are distinct should hopefully be clear. Accepting the first idea of Norton's theory does not commit us to accepting the second idea.

To argue for his second idea, Norton in [11], [12], [13] presents us with a series of arguments against various universal schema for induction that have appeared over the ages. My view however is that these arguments fall short, and that we are yet to see good arguments for the second part of Norton's material theory of induction, in which all universal schema for induction are unequivocally rejected. Indeed, criticizing this second part of Norton's theory will be the main goal of this paper. Rather than focusing on Bayesianism as a universal schema for inductive inference, in what follows I shall focus on Norton's criticism of the idea that IBE can provide us with a universal schema of inductive inference. I will argue that Norton's arguments against this idea fail on his own terms, and that IBE in fact remains a plausible candidate for a universal rule of induction. In this way, I create space for the view that while in accordance with the first part of Norton's theory we should view what justifies any given instance of (\*) as an external, contextually determined fact, all induction can nevertheless be subsumed under a universal schema (understood externalistically.) I shall conclude the paper by arguing that a modification of Norton's material theory in this direction only strengthens it.

### **3 Understanding IBE.**

In this part of the paper we put Norton's material theory of induction to the side temporarily in order to say a few things about IBE. IBE is a type

of inference according to which we infer the best explanation of the data we are presented with. It has been explored in detail (from very different points of view) in classic texts such as Lipton [7] and van Fraassen [18], and I will assume the reader is at least somewhat familiar with the debates that surround it.

One common criticism of IBE found not only in Norton but elsewhere (e.g., Fumerton [3]) is that IBE is so vaguely defined as to lack real substantive content. This strikes me as unfair, as at least nowadays there are many reasonably precise formulations of IBE - see for example Shupbach [15], Tesic et al. [17], and Douven [1]. Many of these formulations are developed in a relatively precise formal or probabilistic framework.<sup>2</sup> Of course, these formulations are not utterly precise in the way in which something like first order logic is utterly precise, but that would be an unreasonable standard to impose, as non-IBE based conceptions of induction do not meet such a standard either. For example, at the end of the day Norton does not tell us anything positive about what the ‘licensing’ relation is that holds between material facts and the inductions that they license, and even Bayesians tend to be silent on the question of where our priors come from in anything but the most idealized cases. It is fair to demand of a formulation of IBE that it be precise enough that we can see how it has a chance of tackling well-known or reasonable objections, but beyond this, it would be wrong to impose higher standards of clarity on it than we do its rivals.

The idea that there is an intimate link between induction and explanation is not a new one - it goes back at least to Harman [5], [6], and has been explored by White [19] and Weintraub [20] among others. The idea that all induction can be grounded in IBE, and thus that IBE can be viewed as a ‘universal schema’ for induction, was perhaps made most famous by Harman in [5], [6]. On p. 531-2 of [6], for example, he argues:

... enumerative induction is warranted only if one may infer the following conclusions: ‘That A is normally followed by B explains why A has been observed to accompany B. That A is normally followed by B will explain A’s being followed by B in the next instance.’

Here, it must be stressed that the idea is *not* that in any inductive inference, the conclusion ‘All As are Bs’ explains, of every object *o* that appears in the

---

<sup>2</sup>Having said that, I actually do not think that developing IBE within the context of a probabilistic framework is the right approach. The problem is that IBE seems to be applicable in cases in which probabilistic machinery is not. Williamson [21], for example, has persuasively argued that IBE is applicable in mathematics. I also think that IBE is applicable even in the absence of knowledge of prior probabilities, and that in fact that in many cases it is only with the help of IBE that we can generate justified prior probabilities in the first place.

premises, why it is a B given that it is an A. (This has been mistakenly supposed by Ennis [2], for example). Rather, the generalization ‘All As are Bs’ is supposed to explain the fact that some sort of data-collecting procedure has only thus far yielded As that are also Bs. (A very similar observation is made by White [19] in his discussion of enumerative induction.) It is in this sense that the conclusion of an inductive argument explains its premises. So for example, consider our inferring on the basis of a sufficiently wide variety of observations that all bismuth melts at 271°C or that the speed of light is constant. In each case we will have conducted a wide variety of experiments in which as many relevant conditions as possible were varied, and in each case, we assume that no counterexample has arisen. The best explanation for this pattern in the results of our data-collecting procedure is simply that bismuth *always* melts at 271°C, or that the speed of light *really is* constant. This of course is not to explain, of any piece of bismuth, why it melts at 271°C (as opposed to some other temperature) - such an explanation would require a detailed detour through chemistry and thermodynamics. Nor is it to explain in any deep sense why the speed of light is what it is. It *is* however to explain why a certain pattern has been noticed in a certain set of observations. It explains this pattern in the observations by positing a deeper pattern that holds quite generally.

This way of re-interpreting an enumerative induction as an IBE works quite generally. Moreover, it offers some insight as to why we perform the enumerative inductions we do while refraining from others. For example, from an observation of many green emeralds we infer the general greenness of emeralds and not their grueness because presumably the hypothesis of the greenness of emeralds is a better explanation of our observations than the hypothesis of their grueness.<sup>3</sup> The papers of Harman [5], [6] and White [19] also bring out the explanatory character of enumerative inductive inference in many other cases, so the reader in want of further examples should consult those papers.

Further misunderstandings of IBE must also be warned against. Note that in some cases, IBE might involve explicitly considering other possible explanations, and judging them inferior before inferring the ‘best’ explanation. In cases like this, the terminology ‘inference to the best explanation’ is especially descriptive of the psychological process that occurs when performing the inference - through a process of comparison we find the best explanation, and then we infer it.

In other cases, however, it is less clear that any sort of explicit comparison with rival explanations is involved. Consider the following example from

---

<sup>3</sup>The question of *why* the hypothesis of greenness provides us with a better explanation of our observations than the hypothesis of grueness is not a question I am claiming to have addressed here, though see White [19] for some highly illuminating comments on this matter.

Peirce [14]. Suppose that a scrap of paper that looks like it was torn from something larger is found, and then a book with a torn corner in one of its pages is discovered such that the scrap fits inside this torn corner perfectly, matching it in size, color, texture, etc., down to the tiniest detail. That this is to be explained by the claim that the scrap was torn from this particular page jumps out to us with such immediacy that it is hard to take any other explanation seriously. We thus find ourselves inferring that the scrap was once torn from this very page. In this case, although we infer what is the best explanation, it does not feel (to me, at least) that we do so by an explicit process of comparison of the plausibility of the explanation in question with other explanations. Indeed, insofar as other possible explanations of the ‘fit’ between the scrap and the book are inferior to the obvious one, they are inferior at least in part because they are simply not the explanation that we are most naturally led to infer. In cases like this, the terminology ‘inference to the best explanation’ needs to be understood carefully - while we *do* infer the best explanation, we do not first decide which explanation is best, and then infer it. Rather, we infer the natural explanation, and in light of this judge that it is the best and that other explanations are inferior.

Throughout this paper, I place no requirements on the psychological process by which we arrive at the conclusion in an IBE. Perhaps in some cases an explicit process of comparison is involved, and perhaps in other cases not. I therefore use IBE as an umbrella term for all inferences to best explanations, regardless of the underlying psychological mechanism.<sup>4</sup>

As shown in papers such as [5], [6], the idea that IBE forms the basis of inductive inference is corroborated by a large range of examples, and counterexamples of the sort offered by [2] can easily be shown to rest on a misunderstanding of IBE. The idea that IBE gives us a universal schema for induction must therefore be taken seriously. Of course, Norton is strongly opposed to any universal schema of induction, and so he must oppose the idea that IBE could give us such a schema. We thus turn back to his arguments now.

## 4 Norton on IBE

### 4.1 Norton’s critique of IBE.

Having discussed IBE and its general role in inductive inference, in this section I present Norton’s criticism of IBE as found in chapters 8 and 9 of

---

<sup>4</sup>Loosely following Peirce, one could perhaps use the term *abduction* rather than IBE to describe the inferences in which I am interested. I have no particular problem with this, but choose to stick with the terminology of IBE, as that is the term Norton uses.

[12]. Norton's main claim there is that putative instances of IBE in science are really just a complex amalgam of inductive and deductive inferences, and that once we flesh out the full structure of these arguments, we see that the notion of explanation plays no substantive role in them at all. Thus, Norton maintains that putative examples of IBE turn out to not be instances of IBE at all.

Norton interprets arguments from real science that have been traditionally viewed as IBEs as instead having a two step structure. In the first of these two steps, a theory (the 'preferred theory') is compared against multiple opponents (or 'foils'), and it is shown that in some important way the preferred theory is indeed better than the foils. This process of comparison does not, however, involve any appeal to explanation. Rather, Norton claims that it typically involves two sorts of argumentation: (i) pointing out that the evidence at hand (perhaps supplemented with background assumptions) contradicts one of the foils (Norton calls this 'contradiction'), and (ii) pointing out that one of the foils '*requires us to accept further assumptions for which we have no evidence*' (Norton calls this 'evidential debt'). Neither of these types of argumentation requires reference to any substantive notion of explanation, and thus Norton thinks that in this first step in which '*the favored hypothesis or theory is shown to do better than one or more foils*', no notion of explanation appears at all.

In the second step, seeing the superiority of the preferred theory to its rivals, we then infer the preferred theory. As Norton puts it, '*... [w]e are to suppose that better is best; and that best is good enough to warrant commitment. Preference becomes commitment.*' Norton sees no special use of explanation here either - all we are doing in this step is committing to the stronger theory. The net result is then that '*[a]bductions or inferences to the best explanation in actual science are carried out in two steps with some distinctive notion of explanation playing no role in either.*'

In chapter 8 of [12], Norton also describes some different, independent worries about IBE. In particular, he spends much time arguing that IBE is not precisely formulated, and also asserts that we have no sufficiently unambiguous notion of explanation on which to base IBE. Although I do not think that these criticisms are well founded, I will not dwell on them here. Instead, I would like to focus on Norton's 'explanation-free' reading of putative IBEs, and try to assess whether Norton has really purged these examples of their IBE-like character in the way he claims.

In chapter 9 of [12], Norton goes through a number of arguments from actual science generally taken to be IBEs, and claims to provide 'explanation-free' readings of them, fitting them to the two step structure just described. I will summarize three of these cases here.

Norton begins with Darwin's argument for the theory of natural selec-

tion. This is often read as an IBE in which natural selection is taken to be the best explanation for the diversity of species, and thus inferred. Norton, however, sees Darwin's argument differently. He views Darwin's argument as establishing two facts: (i) that '*[i]t is possible that the variety of species arose from descent with modification through natural selection*', and (ii) that '*[i]t is unlikely that any other admissible account can accommodate the origin of species*'. Fact (i) is established '*by the massive weight of Darwin's many examples*', and fact (ii) is established by casting doubt on the theory of independent creation, pointing out that the theory of independent creation '*... is contradicted by many arbitrary facts in natural history for which a creator would have no evident purpose.*' Presumably, step one of Norton's two step process, in which the theory of natural selection is shown to be superior to its rivals, is constituted by arguing for (i) and (ii). In step two of the argument, we then move from this comparative judgment to commitment to the theory of natural selection. According to Norton, neither of these steps rest on any notion of explanation of any sort.

Norton also considers Lyell's argument for a uniformitarian theory of geology, according to which geological phenomena are the result of causal processes still in operation occurring repeatedly over massive periods of time, rather than 'cataclysms' which punctuate geological history and radically alter the development of the earth. This argument is also often seen as an IBE. Norton, however, sees Lyell's scientific arguments in [8] as establishing two facts: (i) that '*[i]t is possible that present geological features arose over long time periods from causes now operating*', and (ii) that '*[i]t is unlikely that any other admissible account can accommodate their origin.*' Norton tells us that fact (i) '*is established by the wealth of examples in Lyell's account*', and that any argument against fact (ii) '*must speculate on presently unknown causes or known causes but of presently unknown intensity*', and thus '*tak[e] on an undischarged evidential debt.*' In the same way as Darwin's argument, establishing these two facts shows the superiority of Lyell's theory, and constitutes step one of Norton's two step process. In the second step, we then commit to Lyell's theory. Again, Norton thinks that neither of these steps rest on any notion of explanation of any sort.

Norton also considers Thomson's argument that the rays produced in a cathode tube are constituted by particles (electrons), rather than being waves in some sort of ether. This is often taken to be an IBE. However, Norton points out that Thomson's argument crucially revolves around the observation that cathode rays are deflected by magnetic fields. The particle theory accounts for this nicely, but the wave theory, Thomson argues, does not seem to be able to account for this in any clear way. Step one of Norton's reading of Thomson's argument notes the superiority of the particle theory based on these (and perhaps other) considerations, and step two then

involves committing to the particle theory, as before. Again, Norton claims that no notion of explanation is needed for this argument.

Norton considers several other examples too, but their characters are sufficiently similar to the examples just given that we do not consider them separately. Norton's claim is that these examples show that many famous putative examples of IBE in science are not really IBEs at all, but rather instances of the two-step, explanation-free pattern of inference just described. In this way, Norton perhaps even casts doubt on the idea that IBE *ever* occurs in science at all, let alone that it can function as some sort of universal schema for induction. It is this argument that we will need to assess.

## 4.2 Critiquing the Critique - The Argument.

Has Norton really given us an unproblematic, explanation-free reading of the arguments in question? Let us consider Norton's two step process more carefully. In step one, we examine some finite set of candidate theories  $\{T_1, T_2, \dots, T_n\}$ , showing that all but  $T_1$  either contradicts the evidence or requires strong assumptions for which we have no evidence. In this way, in step one we show that the theory  $T_1$  is significantly more favored by the evidence than are  $T_2, \dots, T_n$ . But having done so, in virtue of what does Norton think we are justified in inferring  $T_1$  in step two?

We have seen how Norton answers this question. Norton thinks that if the preferred theory is strongly supported by the evidence in a way that the rivals we have examined are not, we are then justified in believing the preferred theory. Norton captures this idea with the slogan that '*better is best, and that best is good enough to warrant commitment*'. There are two parts to this slogan (first that 'better is best', and second that 'best is good enough to warrant commitment'), but let us nevertheless group them together as Norton does and call the result the 'better-is-best' principle. It is armed with this principle that in step two we move from the claim that  $T_1$  is significantly more favored by the evidence than are  $T_2, \dots, T_n$ , to justified belief in  $T_1$ .

In step one of this process, we will of course only have considered a finite number of a presumably infinite set of possible rivals to  $T_1$ . (A caveat: Norton thinks there are cases in which we are actually dealing only with a small finite number of possibilities and thus may consider them exhaustively in step one. I do not think Norton is right about this, and will come back to this point in §4.4.) Given this, what allows us to infer  $T_1$  from the fact that  $T_1$  is significantly more favored by the evidence than the few rivals we have considered? Why does Norton think that his better-is-best principle is justified?

The better-is-best principle involves endorsing something like the follow-

ing inference ( $\dagger$ ):

**Premise:**  $T_1$  is significantly more favored by the evidence than are any of  $T_2, \dots, T_n$ .

**Conclusion:**  $T_1$  is significantly more favored by the evidence than any rival at all.

From the conclusion shown,<sup>5</sup> we then get justified belief in  $T_1$ , though this step will not concern us here.

Why then is the inference ( $\dagger$ ) justified? If Norton cannot give an answer to this question that is acceptable on his own terms, then he has no right to apply his better-is-best principle, and his two-step analysis must be rejected. The inference ( $\dagger$ ) is an inductive inference, and as such, Norton must think that it is warranted by some material fact. This material fact may well be different in different cases. In any particular case then, what could that material fact be? We consider four proposals, and show that they all fail in various ways. It will be the third and fourth of these proposals that perhaps warrant the closest attention.

(i) One possibility is to say that the mere truth of  $T_1$  (known or otherwise) provides a warrant for the inference ( $\dagger$ ). In this way,  $T_1$  itself then becomes Norton's material fact. Thus, the truth of Darwin's theory or Lyell's theory is what warrants our inferring that these theories are significantly more favored by the evidence than any rival from the fact that they are significantly more favored by the evidence than any rival we have considered.

Now, even in cases in which  $T_1$  is true, it is not clear to me in what sense it is the mere *truth* of  $T_1$  that warrants this inference. In addition, it seems clear that as a general way of licensing ( $\dagger$ ), this strategy will not work. For there will be cases in which the theory  $T_1$ , although it is the theory more favored by the evidence at hand than any rival, is nevertheless false. Newtonian physics relative to the evidence obtained before the 20th century is presumably an example of this. Perhaps Lyell or Darwin's theories also end up being examples of this. In such cases, the truth of  $T_1$  cannot play the role of a material fact, because  $T_1$  is not true.

(ii) Perhaps then it is the fact that  $T_1$  *just is* the theory more significantly favored by the evidence at hand than any rival that can serve as

---

<sup>5</sup>Whether the conclusion of ( $\dagger$ ) is too strong is unclear. To be justified in believing  $T_1$ , perhaps it is enough that  $T_1$  is significantly more favored by the evidence than a very large class of rivals *of a certain sort* (certainly larger than the set of rivals that have been explicitly considered.) Or perhaps it is enough that  $T_1$  is one of a *small set* of theories significantly more favored by the evidence than any rivals are. In either case, it would suffice for ( $\dagger$ ) to have a weaker conclusion than the one shown. Whether such weakenings of ( $\dagger$ ) are required is a difficult question requiring further investigation. As the arguments that follow are insensitive to such minor modifications of ( $\dagger$ ), for the sake of simplicity I shall just proceed with the version of ( $\dagger$ ) just given.

Norton's material fact (where whether this material fact is itself known is not important.)

But this too is highly problematic. The material fact under consideration is just the conclusion of ( $\dagger$ ). On this view, is it therefore simply the fact that the conclusion of ( $\dagger$ ) is true that licenses the inference ( $\dagger$ ). But this is very hard to make sense of. If the material fact that licenses an inference is supposed to be something that in some way explains why the inference in question is justified, then the material fact surely cannot be the conclusion of the inference itself. One cannot explain why an inference is justified merely by pointing out that its conclusion is true, and it is difficult to see what value there could be in a theory of induction which said otherwise.

Perhaps more compelling proposals for the material fact that warrants ( $\dagger$ ) come from taking seriously the idea that ( $\dagger$ ) is an inference about our methods of theory generation. We use ( $\dagger$ ), after all, to tell us that the small set of vaguely plausible rivals to  $T_1$  that we have examined is in some sense representative of a larger set of possible rivals, not all of which we have examined. Thus, one might think that what warrants ( $\dagger$ ) is some fact about our methods of theory generation, or even some fact about the human mind and its capacity of theory generation. We explore these sorts of options now.

(iii) One might think that, when presented with a body of evidence, the human mind simply has a general capacity to come up with a small set of candidate theories that contains the theory most strongly supported by the evidence in question. Perhaps millennia of practice at coming up with hypotheses has forged such a capacity in the human mind. For our purposes, we need not suppose that such a capacity is infallible - that is, we do not need to suppose that we are *always* able to come up with a set of hypotheses that contain the theory most strongly supported by evidence - rather, it is enough for us to suppose that we are somewhat reliably able to do so. Consider then the claim that human beings - or at least good scientists - have such a decently reliable capacity. Could such a claim be the material fact that warrants ( $\dagger$ )?

Now, whether it is right to attribute such a capacity to scientists is far from clear,<sup>6</sup> but let us not worry about this, and instead just take for granted that this proposed material fact is indeed a fact. The problem then is that we are now in a position to formulate something very much like a universal schema of induction. In particular, consider the inference schema ( $\dagger\dagger$ ):

**Premise:** Given some fixed evidence, good scientists have told us that the most obvious hypotheses are  $T_1, \dots, T_n$ .

**Premise:**  $T_1$  is significantly more favored by the evidence than are  $T_2, \dots, T_n$ .

**Conclusion:**  $T_1$  is more favored by the evidence than any rival.

---

<sup>6</sup>This question arises very naturally in the context of Stanford's arguments in [16].

If the material fact under consideration licenses the inference ( $\dagger$ ), then surely the inference ( $\dagger\dagger$ ) must be justified (though of course, only defeasibly.) But the problem is that we can now license basically any scientific inductive inference using ( $\dagger\dagger$ ). So for example, from the observation of several bismuth samples melting at  $270^{\circ}\text{C}$ , the set of hypotheses any good scientist will put forth will surely contain the hypothesis that *all* bismuth melts at  $270^{\circ}\text{C}$ . Assuming that this hypothesis can be shown to be significantly more favored by the evidence than any of the rivals the scientist has come up with, we could then use ( $\dagger\dagger$ ), licensed by the material fact in question, to show that belief in the hypothesis that all bismuth melts at  $270^{\circ}\text{C}$  is justified. In this way, Norton's two-step process and the single material fact in question can be used to show why essentially *any* warranted inductive argument in science is indeed warranted. The schema ( $\dagger\dagger$ ) then becomes for all intents and purposes a universal schema for induction, contrary to Norton's insistence that no such thing exists. Moreover, supposing we know the material fact in question, the universal schema for induction here is internalistic, and so is the sort of thing that Norton most emphatically rejects. So even if there is an interesting theory of induction here, it is not Norton's.

(iv) A more fine tuned variant of the previous proposal might focus more on the specific techniques we use to generate theories in particular subdomains of specific sciences, rather than talking in overly broad terms about capacities of the human mind. So for example, given the standard set of idealizations and tricks the classical physicist has in her toolbox, the classical physicist is quickly able to formulate a plausible set of theories and models supported by the evidence when dealing with phenomena of a simple, classical nature. This is a skill that has been forged over centuries of meticulous study of specific sorts of systems. In many areas of science, there is a similar toolbox of tricks which, when appropriately mastered and used with a good sense of judgment, allows the scientist to steer her attention to a good set of hypotheses. In talking about such a toolbox of tricks, one might even choose to divide the sciences very finely, distinguishing for example the quantum statistical mechanics of gases from the quantum statistical mechanics of liquids, drawing fine distinctions between the efficacy of the techniques used in these various sub-domains of science.

Consider then the fact that, for a given sort of scientific problem in a given sub-domain of science, a certain set of techniques judiciously applied reliably leads to a set of hypotheses containing the one most supported by the evidence at hand. Could a fact of this form serve as a material fact licensing ( $\dagger$ ) in a particular case? Unlike the previous proposal, such a material fact would differ from case to case. We need not even require this fact to be known - the mere fact that the techniques in question reliably lead to a set of hypotheses containing the one most supported by the evidence

at hand would be enough to license ( $\dagger$ ). Perhaps in areas in which scientists do not really have any such reliable set of rules-of-thumb, ( $\dagger$ ) would not end up licensed at all - a result we should probably be happy with.

Let us put aside worries about whether such material facts are really true, and how precisely we are to distinguish the various sub-domains of science from each other in any particular case. I claim that we find ourselves in more or less the same situation as the previous proposal. In particular, whether or not there is an interesting theory of induction here, it is not Norton's. The problem again is that it leads to a universal scheme of induction ( $\dagger\dagger\dagger$ ):

**Premise:** In the hands of appropriate experts, the set of techniques appropriate for dealing with the problem in question has yielded the hypotheses  $T_1, \dots, T_n$ .

**Premise:**  $T_1$  is significantly more favored by the evidence than are  $T_2, \dots, T_n$ .

**Conclusion:**  $T_1$  is more favored by the evidence than any rival.

This is a little different from the universal schema offered in ( $\dagger\dagger$ ). In this case, the schema ( $\dagger\dagger\dagger$ ) is best thought of as an externalistic schema of induction, spelt out as follows (using the terminology of §2): if (*i*) we are justified in believing that  $T_1$  is significantly more favored by the evidence than are  $T_2, \dots, T_n$ , and if (*ii*) we are justified in believing that the appropriate set of experts, applying the set of techniques appropriate to the problem in question, have yielded the hypotheses  $T_1, \dots, T_n$  (this is our single  $F$ ), and if (*iii*) it is true (though not necessarily known to us) that the set of techniques in question judiciously applied reliably leads to a set of hypotheses containing the one most supported by the evidence at hand (this is our single  $G$ ), then we are justified in believing that  $T_1$  is more favored by the evidence than any rival. Assuming that from the fact that we are justified in believing that  $T_1$  is more favored by the evidence than any rival it follows that we are justified in believing  $T_1$ , this then gives us an externalistic schema of induction. And so insofar as Norton is opposed to externalistic as well as internalistic universal schemes of induction, this proposal for how to license ( $\dagger$ ) cannot work either.

The general worry then is that it is not clear how to justify an inference like ( $\dagger$ ) in a way that does not undercut the material theory of induction. Consequently, it does not seem that Norton's two-step account of these episodes of science can be acceptable on his own terms. Perhaps the way forward is for Norton to abandon his two-step analyses of the examples in question, in such a way that inferences like ( $\dagger$ ) are avoided altogether. But then Norton is left having to explain how the relevant scientific inferences can be justified without invoking any notion of explanation, and in such a way that the main ideas of his material theory are not compromised. I do

not see any suggestions in any of Norton's texts as to how this might be done.

### 4.3 Critiquing the Critique - The Conclusion.

Note by contrast that the introduction of IBE provides us with a very straightforward understanding of why (†) is justified in the cases in which it is. In such cases, our extensive efforts to come up with rivals to  $T_1$  have yielded only theories that are contradicted by the evidence or that involve hypotheses not supported by evidence that we currently have. What is the best explanation for the fact that we have been unable to come up with any adequate rival to  $T_1$  in this way? The obvious best explanation is that there *is* no such adequate rival to  $T_1$ , and that  $T_1$  really is more favored by the evidence than any rival. In this way, (†) can easily be justified when viewed as an IBE.

Norton's two-step proposed replacement for IBE cannot be justified in terms acceptable to the material theory. This two-step replacement is however easily justified by a deeper underlying IBE. Given all this, we cannot view Norton as showing on his own terms that IBE is redundant in science, as he seemed to have hoped.

### 4.4 Critiquing the Critique - A Caveat.

To complete this discussion, we must now return to a caveat earlier made. In some cases Norton thinks we only need begin his two-step process with a finite set of hypotheses. For example, in the case of Lyell's theory, Norton seems to think that Lyell only needs to consider two options. Norton tells us in Chapter 9 of [11]:

Lyell, however, has two cases that are exhaustive. Either present geological features arose from causes now in operation; or they did not. The first case is Lyell's uniformitarianism.

Norton also sees the debate about rays in cathode tubes as involving only two alternatives:

Key to the arguments of both sides is an assumption of exhaustion: that the two alternatives they considered - matter or waves - were exhaustive. ... The assumption of exhaustion then did the critical work of allowing the step from the adequacy of each sides' account to its truth.

But if this is right, then it looks like in these cases we do *not* need anything like the argument (†) to complete Norton's two-step argument. In particular, because there are only two options on the table, and in step one we are

systematically investigating both, we do *not* need any principle telling us that the cases we have not investigated would not likely yield a different conclusion, because there are no cases that we have not investigated.

I do not think, however, that this way of trying to understand Lyell's argument works. The problem is that the negation of Lyell's uniformitarian theory is not a single theory that can clearly contradict the evidence or take on an evidential burden. Uniformitarianism asserts that geologic processes occur in more or less the same manner and frequency in the past as they do in the present, and that this is sufficient to account for all geologic change. There are many ways in which such a view could be false. It could be that geologic processes have always occurred in more or less the same manner and frequency as they presently do, but that this has been punctuated by occasional 'catastrophes'. The various characters these catastrophes might have would then lead to different pictures of geological history. Alternatively, a non-uniformitarian might maintain a catastrophe-free theory in which geological events have always had the same character, but their frequency or speed has gradually changed over time. The form of these gradual changes then also lead to very different pictures of geological history.

When Lyell points out various problems with non-uniformitarian theories or the evidential debt they accrue, he is targeting specific versions of non-uniformitarian theories, and not non-uniformitarian theories in general. So for example, on p. 4 of [8] Lyell points out specific evidence against the idea that fossils are created from inorganic materials. Lyell then goes on to criticize the idea that basalt had an aqueous and not igneous origin, as well as the idea that the existence of fossils at very high altitudes should be explained by the drying up of oceans or elevation of solid land. By a sequence of such observations, Lyell builds his case for his preferred uniformitarian theory. In each case, however, Lyell is not targeting non-uniformitarian theories in general, but rather specific versions thereof. If we want to impose a Nortonian two step structure on this argument, we would therefore have to say that in step one he is comparing his own uniformitarian theory against the best non-uniformitarian theories he can come up with, and then noticing that all the non-uniformitarian theories he can come up with are inferior in various respects to his preferred theory. In step two, he then infers his preferred theory. But if this is right, then Lyell's argument really does need to use something like the inductive argument ( $\dagger$ ), and so the question of how ( $\dagger$ ) is to be justified cannot be avoided.

The same is true of Thomson. There are infinitely many forms a non-particle theory could take, and even wave theories admit many forms. Thomson's argument that wave theories cannot be right because the bending of waves requires differences in wave speed at different places makes a host of assumptions that surely do not hold of *all* wave theories, and thus insofar

as we want to impose a two-step structure, Thomson in step one surely only has the right to say that his particle theory is stronger than the best versions of the wave theory he can think of. The inference from this to the particle theory then requires the argument ( $\dagger$ ), as before.

Of course, there are various ways Norton can carve up his two-step process, but however he does so, I claim the basic problem remains. I therefore cannot agree that Norton has provided us with explanation-free readings of the case studies in question. Insofar as the possibility of doing so is the core of Norton's critique of IBE, we must conclude that Norton has not successfully refuted the hypothesis that IBE provides us with a universal rule for induction.

## 5 Conclusions and Prospects.

Norton's arguments against universal schema of induction - and in particular, against IBE - are not convincing. What reaction should we have to this? Should we, on behalf of Norton, try to come up with better arguments against the various universal schema of induction Norton considers?

This does not seem to me to be the best path forward. Instead, I think that by embracing universal schema of induction Norton can improve the material theory of induction at one of its weakest points. One of the most serious problems with the material theory of induction is a particular vagueness that can be found in every presentation of it, including Norton's most recent [12]. Specifically, Norton never tells us what it *means* for a material fact to 'license' an induction (though as we have noted earlier, he tells us various things about what it does not mean.) Nor does he even seem to try. But without a clear conception of what 'licensing' is, the material theory of induction is really not much more than an idea for a philosophical theory, rather than a fully fledged theory itself.

But this is precisely where 'universal schema' can help Norton. To see how, let us return to Norton's bismuth example. According to IBE, the best explanation for our various observations of bismuth samples melting at  $271^{\circ}\text{C}$  is that all bismuth samples melt at  $271^{\circ}\text{C}$ . But of course, it is not simply in a vacuum that this is the best explanation of our observations. It is only given the facts that bismuth is an element, and elements have stable melting points that the explanation in question functions as the best explanation of our observations. Thus we have the following structure:

**Context:** Bismuth is an element, and elements have stable melting points.

**Observations:** Various bismuth samples were seen to melt at  $271^{\circ}\text{C}$ .

**Best Explanation:** All bismuth melts at  $271^{\circ}\text{C}$ .

This structure defines a 3-place relation between a piece of background knowledge (the ‘context’), some observations, and the best explanation of those observations. Note that the context is a conjunction of two claims, and perhaps there is even other information that could be included in the context in this case. Now it seems to me that an externalist like Norton has room to argue that not every fact in the context needs to be known - perhaps the mere external truth of certain parts of the context, along with the remaining parts of the context being believed (perhaps with justification) by us, makes it reasonable to infer the ‘best explanation’ shown from the observations given.

Thus, armed with IBE, material facts license inferences in certain cases in which were they adjoined to our general background justified beliefs, the resulting set of justified beliefs would provide a context relative to which the conclusion of the induction is the best explanation of the observations contained in the premises. This statement on its own helps clarify to a great extent what Norton’s relation of ‘licensing’ might consist in.

Nevertheless, one must be cautious here. Consider, for example, a case in which we have no background knowledge about bismuth or elements having stable melting points at all, but merely have a few observations of bismuth samples melting at 271°C. In this case, it seems counterintuitive to think that we are justified in inferring that all bismuth melts at 271°C (at least, not until we have a lot more data), and thus we would *not* want to say that the inference in question is licensed by the conjunctive claim that bismuth is an element and elements have stable melting points. The lesson here is that the material fact cannot be the context in its entirety, but merely some part of the context that we are happy to treat externalistically and not require that we believe with justification, in order to be able to say that certain explanations are indeed the best explanations of our observations. How exactly such distinctions are to be drawn is really a general challenge of externalism, and whatever version of the material theory Norton adopts, this sort of issue in one shape or form must be resolved. But putting this to the side, if we adopt the point of view I have urged here, we *can* say that the licensing relation is a species of the relation that holds whenever it is the case that were the material fact adjoined to our general background justified beliefs, the resulting set of justified beliefs would provide a context relative to which the conclusion of the induction is the best explanation of the observations contained in the premises. This on its own is surely progress.

If this sort of approach were adopted, IBE would then function as an externalistic universal schema of induction, contrary to Norton’s insistence that there is no universal schema of induction. The main advantage such an externalistic universal schema of induction would bring would be to give

content to the idea of licensing. This on its own would help to clarify Norton's material theory of induction greatly. Perhaps it would not resolve all of the issues that generically arise when considering externalist theories in epistemology, but such matters are not the theme of this essay.

With all this in mind, I find it easy to imagine a different version of Norton - I will call him Norton'. Norton' is opposed to internalistic universal schemes of induction, but is happy with externalistic universal schemes of induction. Norton' in fact uses his favorite externalistic universal scheme of induction to formulate a notion of 'licensing', and develops the material theory of induction from there. What licenses an inference turns out to be something relative to the subject matter, and need not be known to us. In this way, Norton' finds himself in agreement with the first main idea of Norton's theory distinguished in §2 of the present paper. However, even though Norton' rejects internalistic universal schemes of induction, he embraces externalistic universal schemes of induction, and it is on the basis of such an externalistic universal scheme of induction that Norton' fleshes out the material theory of induction. In this way, Norton' rejects the second main idea of Norton's theory identified earlier.

I do not know if the material theory of Norton' is ultimately plausible, but it does seem to me to offer more promise than the material theory offered by Norton. I will conclude with a conjecture: if Norton had drawn the distinction between internalistic and externalistic universal schemes of induction, he might have even have recognized that his main problem was not with externalistic schemes of induction but with internalistic schemes of induction, and Norton might thus have become Norton'.

## References

- [1] Douven, I., *Inference to the Best Explanation. What Is It? And Why Should We Care?*, In McCain and Poston [10] (eds.), 'Best Explanations: New Essays on Inference to the Best Explanation.'
- [2] Ennis, R., (1968), *Enumerative Induction and Best Explanation*, The Journal of Philosophy, Vol. 65, No. 18, pp. 523-529.
- [3] Fumerton, R (1980), *Induction and Reasoning to the Best Explanation*, Philosophy of Science, Vol. 47, No. 4 (Dec., 1980), pp. 589-600
- [4] Goodman, N. (1955). 'Fact, fiction, and forecast'. Cambridge, MA, US: Harvard University Press.
- [5] Harman, G. (1965) *The Inference to the Best Explanation*, The Philosophical Review, Vol. 74, No. 1, pp. 88-95.
- [6] Harman, G. (1968) *Enumerative Induction as Inference to the Best Explanation*, The Journal of Philosophy, Vol. 65, No. 18, pp. 529-533.

- [7] Lipton, P. (2004) ‘Inference to the Best Explanation,’ 2nd edition, Routledge.
- [8] Lyell, L. (1833), *Principles of Geology*, Volume 3, First Edition, London: John-Murray.
- [9] Mill, J. S. (1916)[1872], *A System of Logic: Ratiocinative and Inductive*, 8th ed. London: Longman, Green, and Co.
- [10] McCain, K., Poston, T. (eds.) (2018), ‘Best Explanations: New Essays on Inference to the Best Explanation.’ Oxford University Press.
- [11] Norton, J. (2003) ‘A Material Theory of Induction’ *Philosophy of Science*, 70, pp. 647-70.
- [12] Norton, J. *The Material Theory of Induction*, manuscript, forthcoming.
- [13] Norton, J. (2005) ‘A Little Survey of Induction’ in P. Achinstein, ed., *Scientific Evidence: Philosophical Theories and Applications*. Johns Hopkins University Press. pp. 9-34.
- [14] Peirce, C. S. (1878) *Deduction, Induction and Hypothesis*, *Popular Science Monthly*, Volume 13, August, pp. 470-482.
- [15] Shupbach, J., (2017) *Inference to the Best Explanation, Cleaned Up and Made Respectable.*, In McCain and Poston [10] (eds.), ‘Best Explanations: New Essays on Inference to the Best Explanation.’
- [16] Stanford, K., (2006), ‘Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives.’ New York, Oxford University Press.
- [17] Tesic, M., Eva, B., Hartmann, S., (2017) *Confirmation by Explanation: A Bayesian Justification of IBE*, unpublished.
- [18] van Fraassen, B. (1989) ‘Laws and Symmetry’, Oxford University Press.
- [19] White R. (2005) *Explanation as a Guide to Induction*, *Philosopher’s Imprint*, Vol 5, No. 2.
- [20] Weintraub, R (2013), *Induction and inference to the best explanation*, *Philosophical Studies* 166 (1), 203-216.
- [21] Williamson, T (2016), *Abductive Philosophy*, *Philosophical Forum* 47 (3-4), pp. 263-280.