**THE MATERIAL THEORY OF INDUCTION AT THE FRONTIERS OF SCIENCE**

**William Peden**

This research was funded by the European Research Council Horizon 2020 initiative, within the project *Philosophy of Pharmacology: Safety, Statistical standards and Evidence Amalgamation* grant number 639276

**1. Introduction**

A theory of induction has many purposes. It should offer or permit answers to general sceptical challenges to induction, or at least explain why answering these problems is unnecessary. Metaphorically, it should help us defend the heartland of science. The theory should also help us understand and evaluate inductive inferences in everyday science. In other words, it should help elucidate reasoning in the familiar territory of science. Additionally, it should justify the intuitively plausible inductions that occur in largely unfamiliar (for the inductively inferring scientists) domains where scientists’ background knowledge is exiguous. I shall call these unfamiliar domains ‘the frontiers of science’.

Recently, a number of philosophers have challenged John D. Norton’s Material Theory of Induction (MTI) regarding this last desideratum. Norton has raised similar problems for his own theory. In this article, I shall defend the MTI from these criticisms. While the frontiers of science have some unanswered questions for the MTI, they pose no novel fundamental issues.

In Section 2, I introduce the MTI. In Section 3, I explain its alleged problems at the frontiers of science. In Section 4, I argue that these issues can be addressed. In Section 5, I focus on a similar puzzle that Norton raises for the MTI, which raises several further methodological issues regarding Occam’s Razor and the Principle of Plenitude. I finish by answering some objections in Section 6.

**2. The Material Theory of Induction**

I shall define ‘induction’[[1]](#footnote-1) as an extrapolative inference from an evidential proposition about some characteristics of a sample to a proposition about a broader target population. We may or may not know that the sample is part of our inference’s target population. The extrapolations might also, in degenerate cases, not extend beyond the sample: if we infer ‘Most rooms in this building are ornate’ from observing some of the rooms in the building, without knowing that these rooms constitute all of the rooms, we have still made an ‘inductive inference’ in my sense, because we did not know that our sample was exhaustive.  
  
 In this sense of ‘induction’, its fallibility consists in the logical possibility that our sample may not be representative of our induction’s target population. For example, suppose that we infer that between 98-100% of star systems have planets, based on combining (a) astronomical reports that star systems have planets in 99% of cases where we can detect them with (b) our background knowledge. This particular induction is fallible because the frequency of planets in observed star systems may deviate from all star systems by more than a 1% margin of error. Given that inductive reasoning is fallible and sometimes unreasonable, why are we sometimes justified in making (tentative) inductive inferences?

A traditional answer is that we can justifiably assume that nature is uniform, so that our samples are probably representative of the target populations in the induction (e.g. Russell, 1912, p. 66). However, this approach tackles the question from an excessive level of generality. For instance, the induction about star systems does not seem to presuppose anything about nature *in general*. Additionally, we know that nature is *disuniform* in many respects. That is why we would be sceptical of inductively inferring a universal generalisation about the plumage of a newly discovered species of bird, even if our evidential proposition described a large and meticulously selected sample from that species.

A more plausible answer says that in *some* cases we have background knowledge of local uniformities. I shall call such knowledge ‘local uniformity principles’. It is this background knowledge that justifies our implicit belief that a particular sample is representative of our target population. Richard Whately (1855, pp. 256-260) develops this idea in his theory of induction. According to him, in the star system case, our inference is enthymematic, and the implicit premise is something like ‘All large samples of star systems are representative of star systems in general.’ Thus, induction is really just reasoning with implicit premises that convert the inductive argument into a deductively valid argument. If our inductive reasoning is *good*, then the implicit premises are part of our background knowledge. Yet this deductive formulation of background knowledge’s role is too strong: we are not certain that even a large sample of star systems will be representative. Perhaps in special cases (“demonstrative inductions”) we have good background reasons to expect that all of the possible samples are representative, but often we only know that “almost all”, “95 to 100%”, “75%”, or even just “most” of the samples will be representative.

Norton’s salient idea is that these more modestly qualified propositions about local uniformities in populations can provide genuine, though imperfect, justifications for our inductions (Norton 2003). (The history of this idea is beyond my scope.) Like Whately, Norton thinks that reasonable inductions are justified in virtue of our background knowledge about *local* uniformities; unlike Whately, he thinks that this justification can be non-deductive. Inductions are not justified by some very general principle about nature, nor the form of the inference, nor the error probabilities of our testing procedures, and not our subjective credences either. In the case of the star-system induction, our background knowledge of star systems assures us that the very large samples of observed star systems that astronomers have investigated (with respect to having at least one planet) are typically representative of star systems in general. This justifies our expectation that our observations are representative of our target population (at least within a 1% margin of error) and therefore our extrapolation to the hypothesis that 98-100% of star systems have at least one planet.

Furthermore, according to Norton, there is no abstract and general account of the role of local uniformity principles in inductive inferences. These principles can strengthen or weaken our inductions in different ways. For instance, Norton notes (2003, pp. 649-650) that given some background knowledge, even a small sample can provide strong inductive support. Imagine that we infer:

(1) Bismuth is a chemical element.

(2) Generally, chemical elements have uniform melting points.

Therefore, defeasibly, (3) Bismuth has a uniform melting point.

(4) This small sample of bismuth melts at 271° C ± ε.

Therefore, (5) All bismuth melts at 271° C ± ε.

While the inference of (5) from (3) and (4) is deductive, the inference of (3) from (1) and (2) is non-deductive. This latter sort of inference is variously called “direct inference”, “the statistical syllogism”, “the proportional syllogism”, and so on. I shall use ‘direct inference’. It occurs when we infer from a statistical generalisation about a set to a statement about one of its subsets[[2]](#footnote-2). In this case, the set consists of the chemical elements and the subset consists of bismuth. Since the direct inference is non-deductive, the inference is defeasible: if we know that bismuth is atypical among chemical elements, then the inference can be unreasonable for us[[3]](#footnote-3). Thus, if we knew that bismuth is one of the elements with allotropic forms (which can possess differing melting points) like carbon, then we would know a defeater for the inference. The step that we would normally make explicit is from (1) to (5). This induction seems justified in spite of the tiny sample, but only because of implicit steps using direct inference from our local background knowledge.

Let us return to the distinction between (a) the territory of science, (b) the heartland of science, and (c) the frontiers of science. Even some of the MTI’s critics grant that it gives a good account of (a) (Kelly, 2010, p. 757). There is an ongoing controversy about its capacity to defend (b) (Okasha, 2005, p. 250; Kelly, 2010; Saatsi, 2010; Worrall, 2010; Norton, 2014.) My own views are ambivalent on this latter issue (Peden, 2019). My focus in this article is whether (c) troubles the MTI.

**3. The Frontiers of Science**

In various degrees of detail, critics have presented this critique of the MTI:

(1) According to the MTI, all justified inductive inferences are warranted due to our background knowledge of suitable local uniformity principles, which justify our expectation that the sample involved is representative of the target population with respect to the hypothesis.

(2) There are many inductions in which suitable local uniformity principles are not available, but these inductions nonetheless seem justified.

Therefore, (3) The MTI frequently conflicts with our intuitions about inductions.

(4) A good theory of induction should not frequently conflict with our intuitions.

Therefore, (5) The MTI is not a good theory of induction.

Several philosophers of science have raised this problem for Norton’s theory (Ducheyne, 2008; Steel 2010, p. 172; Kelly, 2010, p. 757; Achinstein, 2010, p. 737; Jackson, 2019, p. 8)[[4]](#footnote-4). Marc Lange, though his discussion antedates Norton’s presentation of the MTI, makes this criticism of Samir Okasha’s similar theory of induction (Lange, 2002). Norton himself argues that similar examples pose an “intractable” problem for the MTI, although he does not believe that they are so frequent or counterintuitive to warrant rejecting his theory (2010, pp. 169-174).

Lange’s discussion is especially helpful, because he offers some real scientific inferences that are intuitively reasonable, yet our background knowledge lacks any apparent local uniformity principles. His first example is Henrietta S. Leavitt’s observations (using sophisticated instruments) of a correlation between the brightness of Cepheid stars[[5]](#footnote-5) and the period between the maxima in their light-curves. Her observed regularity was later generalised by Harlow Shapley to all Cepheids in the Milky Way and by Edwin Hubble to the Andromeda galaxy. Lange’s second example is Johann Daniel Titius’s induction that the mean orbital radius *r* of every planet in the Solar System is approximately equal to *r* = 4 + (2n)(3), where *n* is normalized to 10 units for the Earth, with *n* taking the value -∞ for Mercury and 0, 1, 2… for the planets going outward from Mercury. This regularity became known as Bode’s Law. It predicts that the outer planets will be roughly twice as far from the Sun as the preceding planet. Although the law was later disproved, it had some stunning early predictive successes (including from the observations of Uranus by William Herschel in 1781) and astronomers once widely believed it. Lange’s final example is that, since at least the mid-19th century, macroecologists have observed a correlation between proximity to the tropics and biodiversity. They have generalised this correlation, despite the absence of any apparent background theoretical reason to expect the observed ecosystems to be representative of all terrestrial ecosystems (Lange, 2002, pp. 229-231).

Lange grants that *some* background knowledge played a part in these inductions. Shapley and Hubble knew that the boundaries between Andromeda and other constellations were arbitrary: these boundaries did not constitute some natural kind. (They also had no reasons to expect Cepheids in Andromeda to differ from observed Cepheids.) The key to Lange’s objection is the apparent absence of local uniformity principles that would warrant the expectation that the samples are representative of the inductions’ target populations.

The fundamental point in this criticism of the MTI (and similar theories) is that premise (1) of the argument above, which I shall call the *Factivity* claim in the MTI, cannot be reconciled with our intuitions about reasonable inductions at the frontiers of science. There are many other theses within the MTI that are not being disputed in this challenge, like Norton’s emphasis on *local* background knowledge, his claim that there are rational non-deductive inferences, and his rejection of Bayesian priors as an ever-present feature of induction. The problem must also be distinguished from the Humean Problem of Induction or Nelson Goodman’s New Riddle of Induction. The objection is an alleged a conflict between the MTI and our intuitions in many inductions. To answer this objection, it is sufficient to disprove this alleged conflict, without attempting a *general* defence of the MTI.

Of course, the argument at the beginning of this section could be disputed in many ways. Supporters of the MTI could argue that our intuitions are at fault, i.e. reject premise (4). This might not be disastrous for the MTI: Peter Achinstein (2011, p. 284) notes that rejecting some inductions at the frontiers of science would not entail inductive scepticism. One could also dispute whether our intuitions in the particular examples really favour the inductions, thus rejecting premise (2). However, since science apparently progresses, in part, via inductions at its frontiers, it is hard to see how Norton could reject *all* such inferences. My strategy is to reject premise (2) on different grounds: there *are* suitable local uniformities in the examples where the inductions seem reasonable.

**4. Combinatoric Local Uniformities**

My strategy will draw on part of the voluminous literature on the Humean Problem of Induction. Lange calls this part the “Combinatoric Justification of Induction” (CJI) (Lange, 2011, p. 83). The essential idea is that, for many inductions, we know that our sample is a large subset of a finite population, and combinatorics can prove that the majority of such subsets must be representative of such populations. This provides a *defeasible* to believe that our large samples are representative, and thus suitable for inductive inferences. I shall adapt this idea to the local inductions at the frontiers of science. Unlike any satisfactory answer to the Humean Problem of Induction, I shall liberally utilise inductively-acquired background knowledge when necessary, because my goal is not to answer Hume. I aim only to argue that certain inductions are allowed by the MTI.

Consider Shapley’s extrapolation of Leavitt’s observations of a periodicity-luminosity regularity to all the Cepheids in the Milky Way galaxy. He knew that the observed Cepheids were a subset of those in the Milky Way galaxy, that there were a finite number of stars in our galaxy, and the regularity held in 25 Cepheids (Leavitt, 1912). There will be a (very large) set of distinct 25-fold subsets that can be formulated from the set of Cepheids in the Milky Way at the time of Shapley’s induction. Before proceeding, I shall make the notion of ‘representative’ more precise:

**ε1-representative:** Any subset A is ε1-representative of its superset B with respect to a typical random variable[[6]](#footnote-6) *X* if and only if the mean value of *X* in A deviates by no more than a margin of error of 10% of the mean value of *X* in B.

If 100% of the Cepheids in the Milky Way conform to Leavitt’s periodicity-luminosity regularity, then obviously all of the 25-fold subsets will be ε1-representative with respect to a binary typical random variable for conforming to that regularity. Analogously, the same would be true if 0% of them conformed. However, if exactly 1% of them conformed, then the proportion of ε1-representative subsets would be lower. It would be possible that Leavitt had sampled one of the unrepresentative 25-fold subsets. Still, the proportion would still be very high. Similarly, if exactly 99% of them conformed, then the proportion of ε1-representative subsets would be lower, but still very high. The proportion would be even smaller if 25% or 75% conformed. Thus, as the proportion of conforming stars is further from 100% or 0%, the proportion of ε1-representative subsets falls. Consequently, the proportion would be minimal if just 50%, i.e. 0.5, of Cepheids conformed to Leavitt’s periodicity-luminosity law. Note here that I am talking about *proportions* of ε1-representative subsets, not their *cardinality*. The latter would vary with the number of Cepheids in the Milky Way, but not the former. Also note that I am discussing proportions of representative *subsets*, rather than the proportions of representative samples among samples we would observe in an infinite long-run of sampling[[7]](#footnote-7).

As supporters of the CJI point out, all of these intuitions about proportions can be formalised and proven mathematically within combinatorics (Williams, 1947, Chapter 4; Stove, 1986, Chapter VI; for some important connected proofs, see Craig and Hogg, 1965, p. 81ff). In particular, the Law of Large Numbers[[8]](#footnote-8) can be combined with techniques for approximating the huge numbers involved in order to determine minimum proportions of representative subsets in any finite population and for any typical random variable. While there are numerous objections to the CJI (some of them very good) the relevant combinatoric facts are uncontroversial.

Either a Cepheid satisfies Leavitt’s law or it does not. Thus, each *n*-fold subset of the set of *n*-fold subsets of stars in the Milky Way can be described by a random variable for whether it is **ε1-representative**. We can use a binomial distribution, *not* to determine a probability of a subset being **ε1-representative** (in the MTI, we do not assume that all reasoning can be modelled via probabilities; see Norton, 2011) but rather to calculate the worst-case scenario for an induction from an *n*-fold sample that all possess a particular value of a typical random variable[[9]](#footnote-9). As noted above, the proportion of **ε1-representative** subsets will reach a minimum if the population frequency is 50%. Using a statistical method called binomial proportion “confidence” intervals, we can use a binomial table (or calculate directly) to obtain a direct result[[10]](#footnote-10). We find that a very high proportion of *n-*fold subsets *must* be **ε1-representative**. In detail, at least 77% of such samples must be **ε1-representative**, i.e. match the population frequency within a margin of error of 10%[[11]](#footnote-11). Note that I am not interpreting the 77% proportion as a probability, but rather as a material fact about proportions of subsets that match a population frequency within a margin of error.

Returning to Shapley’s extrapolation of Leavitt’s periodicity-luminosity law to the entire population of Cepheids in the Milky Way galaxy, I shall use ‘M’ for the set of 25-fold subsets of Cepheids in the Milky Way galaxy. We can rationally reconstruct Shapley’s reasoning within the MTI:

(1) At least 77% of the subsets in M are ε1-representative.

(2) The 25-fold subset of Cepheids observed by Leavitt is a subset of M.

Therefore, defeasibly, (3) The 25-fold subset of Cepheids observed by Leavitt is ε1-representative of M.

(4) 100% of the 25-fold set of Cepheids observed by Leavitt conform to her periodicity-luminosity law.

Therefore, (5) 90-100% of the subsets in M conform to Leavitt’s periodicity-luminosity law.

Finally, without counterexamples to Leavitt’s periodicity-luminosity law or other defeaters in Shapley’s background knowledge, it follows that ‘All Cepheids in the Milky Way galaxy conform to Leavitt’s periodicity-luminosity law’ is confirmed by the evidence, since this universal generalisation is true or approximately true if (5) is true[[12]](#footnote-12). The structure of this reasoning is very similar to Norton’s reasoning in the bismuth case.

A pleasing feature of this CJI-inspired reasoning is that we have supposed a worst-case scenario for the composition of the Milky Way galaxy: that just 50% of its Cepheids conform to Leavitt’s law. Nonetheless, it is consistent with (5) that the Milky Way contains a very large number of exceptions to Leavitt’s law. Stronger results require larger samples: for instance, if Leavitt had observed 2,500 Cepheids rather than just 25, similar reasoning would work for the following explication of ‘representative’:

**ε2-representative:** Any subset A is ε2-representative of its superset B with respect to a typical random variable *X* if and only if the mean value of *X* in A deviates by no more than a margin of error of 1% of the mean value of *X* in B.

Given the actual 25-fold sample, my reconstruction of Shapley’s inference could be strengthened or weakened in various ways depending on his actual background knowledge. (Like Lange’s example, my reconstruction is very historically thin.) The rationality of the inference could be undermined if Shapley possessed background knowledge that made it likely that Leavitt’s observations were unrepresentative of his target population. If he knew that her observation methods were biased towards observing Cepheids that conform to her regularity[[13]](#footnote-13), then he could not reason as I described. There is nothing unusual here: Shapley’s induction was ampliative reasoning; ampliative reasoning is non-monotonic; and non-monotonic reasoning can always be strengthened or weakened by adding more information to the premises[[14]](#footnote-14). The inference could also be replaced by a stronger one if Shapley possessed what I shall call a ‘strengthener’. If we have good reasons to think that Cepheids are more relevantly uniform than one can expect from my rational reconstruction alone, then the induction has greater justification. (Plausibly, this was the case in Shapley’s actual induction.) This would also enable a tighter definition for ‘representative’ than **ε1-representative**[[15]](#footnote-15).

A strengthener would certainly be required to justify Titius’s extrapolation of Bode’s Law to all of the planets in the Solar System. Only six planets (Mercury, Venus, Earth, Mars, Jupiter, and Saturn) were identified as planets in 1766, when Titius publicised the law (Nieto, 1972, p. 1-2). Although the minimum proportion of **ε1-representative** subsets with 6 members is greater than zero, it is extremely small, and far from enough to warrant making an inductive inference about a potentially very large (though finite) domain like the Solar System. However, astronomers at the time believed that the Solar System possessed a high degree of harmony (Cunningham, 2017, Chapter 1) and thus their background knowledge arguably justified stronger inductions than combinatoric facts alone can rationalise. Additionally, for weaker definitions of ‘representative sample’ than **ε2-representative**, Titius’s induction could be rationalised for a very wide margin of error.

Famously, in domains like psychology, neurology, the social sciences, and avian plumage, we have background knowledge of local disuniformities that generally make inductions from six-fold or 25-fold samples (or sometimes even 2,500-fold samples) weak or outright irrational. However, even in these domains, large sample sizes can sometimes overwhelm our knowledge of local disuniformity. The macroecological example Lange that describes is an instance of this point: ecosystems are very diverse, but ecologists’ evidence consists of many, many thousands of ecosystems that mostly conform to a correlation between biodiversity and proximity to the tropics[[16]](#footnote-16).

I shall discuss the limitations of justifying inductions via direct inferences using combinatoric facts in more detail in Section 6. However, I shall briefly note some important restrictions. Firstly, these combinatoric facts only apply for finite populations. Random variables of infinite populations can be distributed in the form of a Cauchy distribution or a Landau distribution, for which these facts do not apply[[17]](#footnote-17). Secondly, all of the reasoning I have described involves samples and populations that are identified using genuine scientific characteristics: Cepheids, ecosystems, planets etc. This contrasts with parallel reasoning in terms of grue objects, bleen objects, and other classifications that are not (currently) part of scientists’ conceptual schemes. Thus, I assume that we have background knowledge, obtained in part by induction, in support of the scientists’ use of their categories to describe the phenomena, rather than “gruesome” classifications[[18]](#footnote-18). Thirdly, information about the sampling method can serve as a defeater for the reasoning I have suggested[[19]](#footnote-19). Fourthly, while I have appealed to ‘confirmation’, ‘strengtheners’, and ‘defeaters’, neither I nor Norton have not rigorously defined these concepts. I shall return to this issue in Section 6.6.

When will my strategy fail at the frontiers of science? If our sample size is small and we lack strengtheners, then our inductions cannot be rationalised in terms of combinatoric facts about large samples. That makes sense. If the induction’s extrapolated characteristics are not part of the conceptual scheme of our scientific paradigm, then my arguments will also be unavailable. That also makes sense. And if we have defeaters for the direct inferences involved, then rational induction is blocked. As far as I can tell, it is precisely where inductions seem intuitive that the MTI allows induction at the frontiers of science, no more and no less.

**5. Observationally Indistinguishable Spacetimes**

Norton discusses apparently reasonable inductions at the frontiers of science which seem to lack a justification from some local uniformity principle (2010, p. 169-176). First, consider the hypothesis H1 ‘All future worldlines will be extendable by one millisecond of proper time’. There are rival hypotheses in which our spacetime stops at some point, because there can be half-Minkowski spacetimes[[20]](#footnote-20) without this extendibility property. It is provable that these rivals are consistent with any observations that we might make prior to that point *and* consistent with our contemporary understanding of the laws of physics. Norton and others have described such possibilities as “observationally indistinguishable”. Nonetheless, our evidence seems to favour H1 over its rivals. We seem to be warranted in reasoning as follows:

(1) All known worldlines in the past were extendable by one millisecond of proper time.

Therefore, defeasibly, (2) All wordlines are extendable by one millisecond of proper time.

Norton suggests that people seeking an underlying principle behind this induction appeal to a Leibnizian Principle of Plenitude (Norton, 2010 p. 173). Yet the opposite sort of reasoning seems to be made in his second example of such an induction at the frontiers of science. This example starts from the fact that there are a huge variety of observationally indistinguishable spacetimes which add structure (like unobservable wormholes) to a full Minkowski spacetime. Our evidence seems to favour the hypothesis H2 ‘These exotic additional structures do not exist’. We can apparently reason:

(1) None of these additional structures have been observed, nor does our evidence provide any reason to think that they are likely to exist.

Therefore, defeasibly, (2) These exotic additional structures do not exist.

Here, Norton suggests that the putative underlying principle is a form of Occam’s Razor, and concomitantly involves an assumption that simplicity is correlated with truth. Quite aside from the methodological challenges raised by the Principle of Parsimony and Occam’s Razor (how should we adjudicate between them?) the necessity of warranting inductions via sweeping generalisations like ‘simplicity is correlated with truth’ would be inconsistent with *Factivity* - the MTI thesis that rational inductive inferences are always warranted by *local* uniformity principles.

However, the strategy that I used in Section 4 also applies (defeaters aside) to these inductions in favour of H1 and H2. It does not require a Principle of Parsimony or a version of Occam’s Razor. In both cases, we have extraordinarily large samples and standard scientific categories. The local uniformity principles can again take the form ‘At least *r*% of the *n­*-foldsubsets of the population P are representative within a margin of error ε.’ In the cases Norton describes, *n* will be truly colossal[[21]](#footnote-21). Despite lacking reasons from theoretical physics to favour H1 or H2, we have an unimaginably vast quantity of inductive evidence for them.

In Norton’s examples, a new issue is raised by the potentially infinite cardinality of the populations. One might interpret H1 and H2 as quantifying over an infinite domain. Since infinite populations can lack means, it is possible that there are no ε1-representative samples of worldlines or spatiotemporal structures. Even so, we can still make the conditional claims that *if* the means exist, then my suggested reasoning will work. The combinatoric facts I have suggested will be speculative. I do not know whether we should ask for more from induction at science’s frontiers.

How do I explain apparent importance of plenitude and simplicity in these inductions? Firstly, both of the inductions in this section have conclusions that, in part, *assert* plenitude (H1­) and simplicity (H2). There is a tendency to believe that if an inductive inference partly concludes in a proposition P, then the inference presupposes P. For instance, it might seem that the induction from ‘All the many observed ravens have been black’ to ‘All ravens are black’ presupposes that the unobserved ravens resemble the observed ravens in their colours, when actually that resemblance is part of the ampliative content that the inference adds to our evidence (Stove, 1986, p. 11). A theory of induction that allows for ampliative inference, like the MTI, must allow that inductions do not presuppose their extrapolated propositional content – they infer it. The appearance of a presupposition of plenitude in the inference to H1 can be explained in the same way: H1 asserts that worldlines are extendable, which creates an appearance of presupposing that they are extendable[[22]](#footnote-22). This is fits with Norton’s rejection of using sweeping generalisations to justify inductions. The same is true for simplicity: the induction of H2 does not presuppose the simplicity of nature, but rather H2 asserts that part of nature (the structure of spacetime) is simple in particular ways.

But surely, Occam’s Razor is an important part of the scientific method? Material inductivists deny Occam’s Razor as a requirement to accept a general *material* principle (‘You should presuppose that nature is simple!’) but that is consistent with other interpretations. In particular, we can reinterpret both the Principle of Plenitude and Occam’s Razor as corollaries of evidentialism: one’s beliefs and confidence in them should depend only on the relevant evidence[[23]](#footnote-23). Evidentialism can be stricter or more moderate insofar as exceptions to this principle are admitted[[24]](#footnote-24). It follows from this epistemological theory that we should not believe that some part of nature is complex (in some particular sense of ‘complex’) except insofar as that complexity is supported by our evidence. Thus, given evidence in favour of the simplicity of a phenomenon and the absence of evidence to the contrary, we should believe that the phenomenon is simple. However, the same applies, *mutatis mutandis*, to complexity. For example, we are not entitled to believe that 22nd century human societies will be simple in comparison to a game of chess, given the overwhelming evidence of history, anthropology, and our other social knowledge[[25]](#footnote-25).

I also partly attribute the apparent relevance of plenitude to evidentialist intuitions: given a known regular pattern of occurrence and no evidence to expect it to stop, we should expect it to continue. Of course, we very often have reasons to expect such patterns to end, but that is just another instance of the non-monotonicity of inductive support. Evidentialism also entails a negative version of the same principle: without evidence in favour of a particular pattern of occurrence, we are not justified in expecting it. Thus, the intuitive relevance of the Principle of Plenitude and Occam’s Razor can be explained via their normative roles, rather than due to non-local presuppositions of induction. There might be better explanations than mine, but the apparent relevance of these principles of plenitude and simplicity in Norton’s examples can be explained via their role as evidentialist norms, plus our tendency to confuse the ampliative content of an inductive inference with its presuppositions[[26]](#footnote-26).

**6. Objections**

In this section, I shall only consider objections that are internal to the MTI. This focus means that I shall leave aside one of the biggest controversies regarding the CJI, which is the question of whether we must know that we have a “random” selection in a rational direct inference (Franklin and Campbell, 2004; Smart, 2013, pp. 327-329). Norton’s MTI has no such requirement for direct inferences. For example, in the bismuth case, Norton does not try to argue that bismuth is a “random” selection from the set of chemical elements, nor define what “randomness” would mean here[[27]](#footnote-27). Of course, if we possess some defeaters about our sample’s selection, then we should not follow the reasoning I have suggested[[28]](#footnote-28). I shall instead discuss objections that are consistent with the MTI.

**6.1 Are These Combinatoric Principles Sufficiently Local?**

Propositions such as (a) ‘At least 77% of the 25-fold subsets of Cepheids in the Milky Way galaxy are ε1-representative’ and Norton’s examples such as (b) ‘Generally, chemical elements are homogenous in their melting points’ perform similar logical roles (acting as general premises in direct inferences of samples’ representativeness) in my arguments and Norton’s. However, one might think that my justifications in Section 4 and Section 5 are too general for the MTI. Here, I understand ‘local’ to mean something like ‘About a subset of concrete (i.e. non-abstract) phenomena, rather than *all* concrete phenomena’. Hence, my answer might seem to reject exactly that *Factivity* thesis that allegedly caused the MTI problems at the frontiers of science.

The comparatively greater generality of my examples is partly due the greater role of background scientific theory in Norton’s examples[[29]](#footnote-29). Beyond this difference, the appearance is misleading. The proposition (a) is just as local as (b), because (a) is about the subsets of Milky Way Cephids in particular. Just as we cannot reasonably substitute ‘types of wax’ for ‘chemical elements’ in (b), we cannot substitute the name of any reference class for ‘Cepheids in the Milky Way galaxy’ in (a). For instance, we know that (c) ‘At least 77% of the 25-fold subsets of planets in the Inner Solar System are ε1-representative’ is false, because we know that there are fewer than 25 planets in the Inner Solar System. Both (a) and (b) are local, because substitutions of reference classes into these propositions will not always be rational.

In addition, there is also an epistemic similarity between (a) and (b). We know (b) because of the careful and ingenious empirical research of generations of chemists. Similarly, we know that (a) is true in virtue of Leavitt’s observations and our background knowledge that the number of stars in the Milky Way galaxy is finite. This knowledge also comes from empirical research. If we have a 25-fold sample, then a set of 25-fold subsets must exist, and combinatorics enables us to determine a lower limit for the proportion of ε1-representative subsets in any finite set of *n*-fold subsets. The limited size of the Milky Way was also an empirical discovery by astronomers. As with (b), it is our local and logically contingent knowledge (background and evidential) that is providing the real epistemic grounds for the statistical generalisation, rather than abstract and general mathematical facts.

One might still wonder whether there are non-contingent principles of direct inference, which demarcate the good and bad direct inferences from each other. Arguably, these would undermine the novelty of the MTI, as they would be general and non-local[[30]](#footnote-30). Yet even if there are reasonable and unreasonable direct inferences, it does not follow that there are any general principles for demarcating them. Consider how, in moral philosophy, one can deny that there are any such general principles of moral action, but still believe that some particular actions are good or bad. Similarly, one can doubt that there are general laws of nature (except in some strange sense, like a laundry-list of facts) but still believe that natural events occur. I grant that general principles of direct inference are worth pursuing. However, even if we had simple and purely formal principles of direct inference, that would be compatible with the two principal unusual (though not completely original) claims in the MTI: (1) that there are no formal principles *of induction* and (2) that inductions are not warranted by a general *material* implicit premise(s).

**6.2 What About Prior Probabilities?**

A strong objection to the CJI comes from Patrick Maher (1996). He notes that, in Bayesian epistemology, direct inferences can be defeated by (objective or subjective) *a priori* probabilities. For example, given that there are a gigantic number of potential correlations between Cepheids’ periodicity and luminosity in the Milky Way galaxy, a Bayesian might assign a very low prior probability to each of these regularities. If Shapley had such a low prior, this could overwhelm the combinatoric facts, entailing that Leavitt’s sample was probably unrepresentative. The MTI response is simple: strictly *a priori* probabilities have no function in inductive reasoning[[31]](#footnote-31). “Prior probabilities” that reflect background probabilistic knowledge are compatible with the MTI – just not those that reflect (allegedly) objective constraints or subjective opining[[32]](#footnote-32). Therefore, assuming the MTI, priors cannot overwhelm the combinatoric reasoning that I have adapted from the CJI to the frontiers of science[[33]](#footnote-33). This is a striking difference between Norton and Bayesians like Maher[[34]](#footnote-34).

Maher might object that this response reveals a severe problem for the MTI. Doesn’t it mean that Nortonians are vulnerable to the Base Rate Fallacy? Whether this objection works depends on how we interpret “the Base Rate Fallacy”. The accusation is untrue if the Base Rate Fallacy is the fallacy of ignoring relevant background knowledge. Norton grants that if we have strong antecedent evidence that Leavitt’s sample was unrepresentative of the Milky Way in general, then we should not extrapolate as Shapley did. Alternatively, if the Base Rate Fallacy is just the “fallacy” of not having *a priori* probabilities and “failing” to use them in any inductive reasoning, then the MTI is guilty as charged. However, the existence of the offence still needs to be proven. In particular, it is not proven that this latter “fallacy” is akin to the obviously bad reasoning in textbook examples like the Harvard Medical Test. There is much more to be said here; my point is only that the MTI is not avoiding Maher’s criticism by falling into an obviously fallacious pattern of reasoning.

**6.3 Convergence to the Truth**

Benjamin Smart (2013, pp. 329-330) criticises the type of reasoning that I have imported from the CJI, on the grounds that it does not guarantee convergence to the truth, and consequently it has not been demonstrated to satisfy what he calls the “Principle of Convergence”. This principle requires that ‘A conjecture, if false, will be rejected at some stage, and if true will never be rejected’. And it is indeed possible that, even with truly colossal sample sizes, it is logically possible that not only will we be mistaken (something that most modern epistemologists, as fallibilists accept with a shrug) but that we might tentatively accept a false hypothesis at some stage or reject a true hypothesis. The justification that Smart gives for this requirement is that “when we consider what it is to be an inductive inference, we hope that as evidence accumulates our inferences converge on the truth.” (2013, p. 329).

Certainly, we have that hope. It *would* be nice to have a method of induction that satisfies the Principle of Convergence. However, it does not follow that my suggested strategy is irrational because we cannot demonstrate that it will satisfy this principle. Similarly, imagine a sceptic who thinks that Smart is low-balling his demands: they insist that we must know that our inductive reasoning *guarantees* that we will infer a true hypothesis given even a sample of one[[35]](#footnote-35). That would also be nice. We hope that our inductions never lead us to error. It does not follow that such infallibility is a necessary condition for rational inductions. Just as we can reasonably proceed in our inductions without satisfying the requirement of infallibility, so we can proceed without satisfying the Principle of Convergence. For this reason, MTI contains no such condition. The best we can do, and all that we need to do, is to accept/reject only what our evidence warrants, and reject our hypotheses if they are no longer supported by our evidence.

One might think that there is a more modest principle that any justification of some inductive reasoning R should prove to be satisfied. Perhaps we must know that R should be in accordance with a method M that leads us to accept true hypotheses and reject false hypotheses more often than not (Lange, 2011, p. 85-86). We do not know whether the reasoning I have suggested in Sections 4 and 5 will satisfy this more modest principle. It is true that the MTI does not contain such a requirement, yet perhaps this indicates a problem for Norton’s epistemological standards. However, note that what we really want is not to have a generally reliable method, but to justifiably expect our *particular* inductive inferences to be successful. If we somehow knew that M would lead us to the right answer more often than not, we would still need to infer by direct inference that any particular induction would lead us to the right answer. (“Generally, M from *n*-fold samples will lead us to truth and we have an *n­*-fold sample in this induction, therefore M will lead us to truth in this particular induction.”) In other words, even if we had the cosmic knowledge that Lange is demanding, it would only justify our actual inductions via exactly the sort of reasoning I was using in Sections 4 and 5. Why should direct inference be inadequate in the one case and not the other?[[36]](#footnote-36)

**6.4 Grue**

I have presupposed that the scientists in the examples in Section 4 and Section 5 possess justified beliefs about the appropriate scientific categories (reference classes) for the domain in question. A popular criticism of the CJI is that, in foundational inductive cases, it is not clear that we have such information, because it presumably depends on inductive considerations[[37]](#footnote-37). Thus, the CJI arguably falls prey to a version of Nelson Goodman’s (1954, Chapter III) New Riddle of Induction: if we used such reasoning to justify our high confidence in hypotheses like ‘Most grass on Earth is green’, we would face parallel reasoning for contrary hypotheses ‘Most grass on Earth is grue’ (Indurkhya 1990; Lange, 2011, p. 85).

Regardless of whether this criticism is successful in the more ambitious case of the CJI, it does not work for my more modest claims. My goal has not been to solve the Humean Problem of Induction, nor any other general problem of induction[[38]](#footnote-38). I have only argued that the MTI does not clash with our intuitions at the *frontiers* of science. The New Riddle of Induction is a major problem regarding the *heartland* of science which, in itself, the CJI does not solve[[39]](#footnote-39). Therefore, I set aside the New Riddle of Induction in this article.

**6.5 Singular Predictions**

For some inductions at the frontiers of science, we know that the target population of our hypothesis does not contain our sample. In other words, we sometimes make singular predictions from our evidence to a prediction of a singular event. Yet the local justifications that I have made are limited to extrapolations from an observed set to its finite supersets. Can we extend this reasoning to singular predictions?

Yes, but it multiplies the possible defeaters of the inference. Suppose that we have previously inductively extrapolated from ‘All the thousands of flames that we have observed are hot’ and our background knowledge to ‘All or almost all of the flames in the universe are hot.’ An additional direct inference is need to infer to ‘The next flame that we observe will be hot’:

(1) The next flame that we observe is a flame in the universe.

(2) All or almost all of the flames in the universe are hot.

Therefore, defeasibly, (3) The next flame that we observe will be hot.

Since this inference is not deductively valid, it is vulnerable to defeaters. We might discover that the flame we are considering is generated by a previously unobserved fuel, and we might have strong reasons from theoretical chemistry to expect flames generated from this fuel to be freezing cold[[40]](#footnote-40). In general, when induction involves a target population that contains the sample, the defeaters can arise only at the level of extrapolation, whereas in the case of singular predictions they can arise *both* at the level of extrapolation or at the level of direct inference from premises about our intermediary population (flames in the universe) to our target population (the next flame that we observe).

**6.6 The Problem of the Reference Class**

I have made liberal and vague use of the concept of ‘defeaters’ and ‘strengtheners’ in this article. These concepts are familiar in the epistemology of induction and the reader is entitled to expect some clarification on how they work in the MTI. However, I cannot fulfil these expectations in this article, and no Nortonian has done so elsewhere. The most popular answer in contemporary non-deductive logic is the Bayesian approach: determine the strength of the inference by determining the conditional probability of the hypothesis given your full prior probability distribution. Since the MTI forswears prior probabilities, the Bayesian answer to this problem is unavailable.

Put another way: I have employed direct inferences to reconcile the MTI with our intuitions at the frontiers of science. We know that such direct inferences are subject to what Hans Reichenbach called “The Problem of the Reference Class” (1949, p. 374) because in science we always know that our samples are members of many different reference classes and we generally have conflicting statistical information for these reference classes. Therefore, in the absence of a systematic answer to the Problem of the Reference Class, my use of direct inference has a problem of fundamental vagueness.

In this article, I can only say that this is an outstanding *general* problem for the MTI. For instance, in the bismuth case, we know that bismuth is a chemical element and that most chemical elements are homogenous in their melting points, yet we also know that bismuth is a substance and that most substances are *not* homogenous in their melting points. The resolution of the Problem of the Reference Class is a vital issue on the agenda of material inductivists, but there is no reason in principle to believe that such a resolution is impossible. It is a very important inquiry, but it is a matter for another day[[41]](#footnote-41).

**7. Conclusion**

Even Norton’s critics often agree that the MTI performs well in the familiar territory of science. There is a lively debate about whether he can also answer the Humean Problem of Induction at the heartland of science. I have focused upon the frontiers of science. Norton’s *Factivity* thesis seems to clash with our intuitions about many inductions at these frontiers. However, there are local uniformity principles that assimilate the intuitively rational inductions into the MTI. These local uniformity principles concern the minimum proportion of representative subsets of finite populations, which can be calculated using mathematical principles of combinatorics.

I have argued that the frontiers of science pose no particular problem for the MTI. Nothing that I have said entails that the MTI is the One True Theory of Induction. There are also many questions that are set aside in this article; I do not mean to downplay their importance. I am confident that the MTI is not the last word on induction, but I have defended its status as a live option.

**References**

**Achinstein, P.** 2010. ‘Norton’s “Intractable Problem of Induction”.’ In Gregory J. Morgan, (ed), *Philosophy of Science Matters: The Philosophy of Peter Achinstein*, pp. 284-285. Oxford: Oxford University Press.

**Campbell, S.** 2001. ‘Fixing a Hole in the Ground of Induction.’ *Australasian Journal of Philosophy*, 79: 553-563.

**Campbell, S.** **and** **Franklin, J.** 2004. ‘Randomness and the Justification of Induction.’ *Synthese*, 138: 79-99.

**Cunningham, C. J.** 2017. *Bode’s Law and the Discovery of Juno: Historical Studies in Asteroid Research.* Switzerland: Sprinter International Publishing.

**Dawid, R.** 2015. ‘Turning Norton’s Dome Against Material Induction.’ *Foundations of Physics*, 45: 1101-1109.

**Descartes, R. and Kemp Smith, N.** 1952. *Descartes’ Philosophical writings*. London: Macmillan.

**Ducheyne, S.** 2008. ‘Some Worries for Norton’s Theory of Induction.’ *Philosophia Naturalis* 45: 37-46.

**Goodman, N.** 1954. *Fact, Fiction, and Forecast*. Harvard: Harvard University Press.

**Indurkhya, B.** 1990. ‘Some Remarks on the Rationality of Induction.’ *Synthese*, 85: 95-114.

**Jackson, A.** 2019. ‘How to Solve Hume’s Problem of Induction.’ *Episteme*, 16: 157-174.

**Kelly, T.** 2010. Hume, Norton, and Induction without Rules. *Philosophy of*

*Science*,77: 754-764.

**Kyburg, H. E.** 1956. ‘The Justification of Induction.’ *Journal of Philosophy*, 53: 394-400.

**Lange, M.** 2002. ‘Okasha on Inductive Scepticism.’ *Philosophical Quarterly*, 52: 226-232.

**Lange, M.** 2011. ‘Hume and the Problem of Induction.’ In D. Gabbay, S. Hartmann and J. Woods (eds), *Handbook of the History of Logic, Volume 10: Inductive Logic*, pp. 43–91. Amsterdam: Elsevier.

**Leavitt, H. S.** **and Pickering, E. C.** 1912. ‘Periods of 25 Variable Stars in the Small Magellanic Cloud.’ *Harvard College Observatory Circular*, 173: 1-3.

**Maher, P.** 2001. ‘The Hole in the Ground of Induction.’ *Australasian Journal of Philosophy*, 74: 423-432.

**Nieto, M. M.** 1972. *The Titius-Bode Law of Planetary Distances: Its History and Theory*. Oxford: Pergamon Press.

**Norton, J.** **D.** 2003. ‘A Material Theory of Induction.’ *Philosophy of Science*, 70: 647–70.

**Norton, J. D.** 2006. ‘How the Formal Equivalence of Grue and Green Defeats What is New in the New Riddle of Induction.’ *Synthese*, 150: 185-207.

**Norton, J. D.** 2007. ‘Probability Disassembled.’ *British Journal for the Philosophy of Science*, 58: 141-171.

**Norton, J. D.** 2010. ‘Observationally Indistinguishable Spacetimes: A Challenge for Any Inductivist.’ In Gregory J. Morgan, (ed), *Philosophy of Science Matters: The Philosophy of Peter Achinstein*, pp. 284-285. Oxford: Oxford University Press.

**Norton, John. D.** 2011. ‘Challenges to Bayesian Confirmation Theory.’ In Prasanta S. Bandyopadhyay and Malcolm R. Forster (eds), *Philosophy of Statistics* pp. 391-440. Oxford: Elsevier B. V.

**Norton, J. D.** 2014. ‘A Material Dissolution of the Problem of Induction.’ *Synthese*,

191: 671-690.

**Okasha, S.** 2005. ‘Does Hume's Argument Against Induction Rest on a Quantifier-shift Fallacy?’ *Proceedings of the Aristotelian Society*, 105: 237-255.

**Peden, W.** 2019. ‘Direct Inference in the Material Theory of Induction.’ *Philosophy of Science*, 86: 672-695.

**Pollock, J. L.** 1983. ‘Epistemology and Probability.’ *Synthese*, 55: 231-252.

**Pollok, J. L.** 1990. *Nomic Probability and the Foundations of Induction*. New York: Oxford University Press.

**Ramachandran, K. M. and Tsokos, C. P.** 2009. *Mathematical Statistics with Applications*. London: Elsevier Academic Press.

**Reichenbach, H.** 1949. *The Theory of Probability*. Berkeley: University of California

Press.

**Ross, W. D.** 1930. *The Right and the Good*. Oxford: Clarendon Press.

**Russell, B.** 1912. *The Problems of Philosophy*. London: Williams and Norgate.

**Saatsi, J.** 2010. ‘Form-Driven vs. Content-Driven Arguments for Realism.’ In P. D. Magnus and J. Busch (eds), *New Waves in Philosophy of Science*, pp. 8-28. London: Palgrave-Macmillan.

**Smart, B. T.** 2013. ‘Is the Humean Defeated by Induction?’ *Philosophical Studies*, 162: 319-332.

**Steel, D.** 2010. ‘What If the Principle of Induction Is Normative? Formal Learning Theory and Hume’s Problem.’ *International Studies in the Philosophy of Science*, 24: 171-185.

**Stove, D. C.** 1986. *The Rationality of Induction*. Oxford: Oxford University Press.

**Worrall, J.** 2010. ‘For Universal Rules, Against Induction.’ *Philosophy of Science*, 77: 740-753.

1. Throughout, I use ‘single inverted commas’ to refer to words, phrases, and sentences; “double inverted commas” for quotation. [↑](#footnote-ref-1)
2. Direct inference can also involve multiple statements about distinct supersets. I shall not discuss more complex direct inferences. [↑](#footnote-ref-2)
3. To be precise, it is defeasible in the sense of having an “undercutting” defeater, which undermines the rationality of the inference. A defeater is a proposition which, if we justifiably belief it, makes the defeated inference (epistemically) irrational. Deductive inference is also defeasible, in the different sense that we can have “rebutting” defeaters, which make it unreasonable to continue to accept at least one crucial premise in the deduction (Pollock, 1983 p. 233). [↑](#footnote-ref-3)
4. Richard Dawid (2015) also discusses a particular case of the frontiers of science as a problem for Norton’s theory, but the relevant details are too complex to discuss here. [↑](#footnote-ref-4)
5. Cepheids are stars that pulsate radially with regularities in both their pulsation pattern and the amplitude of the pulsations’ brightness. [↑](#footnote-ref-5)
6. A “random variable” is actually a function for assigning numerical values to outcomes, e.g. the numbers 1 to 6 for the outcomes of a six-sided die throw. In this case, these numbers could be ‘1’ for satisfying Leavitt’s regularity and ‘0’ for not doing so. By “typical random variable”, I mean one with a finite mean and variance, with a unique value for every member of its domain. [↑](#footnote-ref-6)
7. The difference is important because in order for there to be a minimum proportion of representative samples, there needs to be a mean for the random variable in question. This mean might not exist in there is an infinite population. [↑](#footnote-ref-7)
8. As applied to finite frequencies, rather than epistemic probabilities, propensities etc. [↑](#footnote-ref-8)
9. For this example, I am especially grateful to an anonymous reviewer for advice on which details to cover and which to leave aside. [↑](#footnote-ref-9)
10. In some cases, we can use a normal distribution to conveniently approximate the values, but this approximation is unreliable under a variety of conditions, including when the sample sizes are small. [↑](#footnote-ref-10)
11. The exact proportion is very slightly above 77%. For the often-used 95% confidence level, we can calculate that ~95% of 25-fold binomial samples, all of them with the value ‘1’, will match the population within a margin of error of 40.89%, which is very broad, but at least somewhat informative. [↑](#footnote-ref-11)
12. I mean ‘confirmed’ in an informal sense, rather than e.g. the Bayesian definition. I assume that, if our total evidence indicates that a hypothesis is approximately true and possibly even exactly true, then we have confirmed that hypothesis. [↑](#footnote-ref-12)
13. Such as if he knew that large Cepheids were easier to observe and more likely to conform to Leavitt’s regularity. [↑](#footnote-ref-13)
14. The same is true for the bismuth example: imagine if we knew that bismuth is a member of a subset of elements that generally have non-uniform melting points. [↑](#footnote-ref-14)
15. Alternatively, it could determine a higher minimum proportion of representative subsets, for a fixed definition of ‘representative’. [↑](#footnote-ref-15)
16. If *n* = 10,000 then well over 99% of subsets are ε1-representative. [↑](#footnote-ref-16)
17. Random variables with a Cauchy or Landau distribution (or some other types) have no means, so there can be no question of subsets matching these population means within a margin of error. [↑](#footnote-ref-17)
18. This is one reason why my arguments should *not* be seen as attempts to answer Hume’s Problem of Induction. [↑](#footnote-ref-18)
19. Imagine sampling (with replacement) marbles from an urn of unknown composition and trying to predict their shape. In the MTI, knowing that a high proportion of large subsets of marbles will be representative provides some justification for expecting a large sample to be representative. However, suppose you know that the marbles will be selected via a hole that only round marbles can fit through. However, information about sampling biases can also strengthen induction: imagine that you knew that the marbles were selected by a computerised mechanism that has been engineered to *guarantee* that an ε1-representative sample of marbles will be drawn. In general, the applicability of combinatoric facts needs to be very sensitive to local background knowledge. [↑](#footnote-ref-19)
20. More generally, any spacetime that is not a full Minkowski spacetime. [↑](#footnote-ref-20)
21. For instance, consider how many milliseconds are known to have occurred in our spatiotemporal past to form our evidence for H1. There are 86,400,000 milliseconds in just one day. With *n* = 86,400,000, over 99% of subsets that match the population mean within a 0.01% margin of error. [↑](#footnote-ref-21)
22. Or even that reality in general is plentiful, or that all *possible* realities are plentiful etc. Once you are playing the game of “finding” presuppositions for an induction, it is hard to stop. [↑](#footnote-ref-22)
23. Norton’s arguments against Bayesianism (such as in 2011) apparently use evidentialism or some very similar principle of rational belief as an implicit premise. I think that evidentialism is right and I tend to be persuaded by his arguments. However, I understand why many Bayesians are not. John Worrall makes a similar point (2010, pp. 751-752). [↑](#footnote-ref-23)
24. To illustrate, I suspect that conformity to evidentialism is only a *prima facie* duty (Ross, 1930). However, evidentialism is a very broad tent. [↑](#footnote-ref-24)
25. Here, my use of ‘complexity’ and ‘simplicity’ does not presuppose that a general characterisation of these concepts is possible. They might be family resemblance concepts. [↑](#footnote-ref-25)
26. In this section, I have only discussed inductive reasoning. Many philosophers believe that simplicity or plenitude, as material rather than evidentialist principles, have a major function in abductive reasoning, but that is independent of my arguments here. [↑](#footnote-ref-26)
27. Randomness is usually defined in terms of probability, but one of Norton’s central contentions is that the role of probabilities in induction is quite limited, contrary to probabilistic accounts such as Bayesianism. [↑](#footnote-ref-27)
28. Contrary to some suggestions (e.g. Smart, 2013, p. 329) these defeaters are not the same as merely knowing that our sample was more likely to be drawn than any other sample. Of course, Leavitt was more likely to observe Cepheids that were relatively close to the Solar System, rather than on the other side of the galaxy, but this should only act as a defeater within the MTI if we know that such Cepheids are atypical in their periodicity-luminosity relation. The mere *a priori* possibility that they are atypical is not enough, because (assuming that we are fallibilists) we do not require that induction cannot lead us into error. Additionally, knowledge of bias can *strengthen* an induction, if we know that the sampling method is biased towards representative samples. [↑](#footnote-ref-28)
29. Background theory still has a non-trivial function in my examples, including to identify good scientific categories over gruesome. [↑](#footnote-ref-29)
30. I am grateful to an anonymous reviewer for raising this point. [↑](#footnote-ref-30)
31. Apart from degenerate cases, e.g. we know *a priori* that we shall not discover inductive evidence in favour of a contradiction or against a tautology. [↑](#footnote-ref-31)
32. The latter must also be distinguished from expert probabilistic judgements, which are not arbitrary. Norton has not discussed expert probabilistic judgements, though there is nothing in the MTI to rule them out. [↑](#footnote-ref-32)
33. Scott Campbell (2001) has defended the CJI from Maher’s objections, while still allowing for *a priori* probabilities. Lange (2011, p. 84) correctly notes that direct inference is not a principle of (Subjective) Bayesian epistemology, though it can hold in special cases. [↑](#footnote-ref-33)
34. Norton criticises Bayesianism in many places (2003, pp. 659-662, 2007, 2011). The controversy is beyond this article’s scope. [↑](#footnote-ref-34)
35. Or consider Descartes’s ideal: “… certain and simple rules, such that, if a man observe them accurately, he shall never assume what is false as true.” (Descartes, 1952). Few, if any, epistemologists today would endorse this standard, but then why is the Principle of Convergence a more reasonable standard? [↑](#footnote-ref-35)
36. Henry E. Kyburg (1956) and David Stove (1986, Chapter 12) make such defences of the CJI, but I am just discussing induction at science’s frontiers. [↑](#footnote-ref-36)
37. Such as extrapolating the past performance of the classificatory system’s inferential success. [↑](#footnote-ref-37)
38. Norton (2014) is the most extensive discussion of Hume’s problem in relation to the MTI. [↑](#footnote-ref-38)
39. Stove (1986, Chapter IX) attempts to answer this problem for the CJI within a Keynesian epistemology and theory of probability. Norton has also discussed the New Riddle of Induction (Norton, 2006). [↑](#footnote-ref-39)
40. No such processes are actually known: even a ‘cool flame’ in the sense of chemistry has a temperature in the hundreds of degrees Celsius. [↑](#footnote-ref-40)
41. There is no shortage of answers in the literature. For example, John L. Pollock (1990) meticulously analyses defeaters and strengtheners in direct inference. The MTI does not presuppose any particular response to the Problem of the Reference Class. [↑](#footnote-ref-41)