DIRECT INFERENCE IN THE MATERIAL THEORY OF INDUCTION

William Peden w.j.peden@durham.ac.uk

Centre for Humanities Engaging Science and Society

50/51 Old Elvet, Durham, County Durham

DH1 3HJ

United Kingdom

John D. Norton's "Material Theory of Induction" has been one of the most intriguing recent additions to the philosophy of induction. Norton's account appears to be a notably natural account of actual inductive practices, though his theory (especially his answer to the Problem of Induction) has attracted considerable criticisms. I detail several novel issues for his theory, but argue that supplementing the Material Theory with a theory of direct inference could address these problems. I argue that *if* this combination is possible, a stronger theory of inductive reasoning emerges, which has a more propitious answer to the Problem of Induction.

Acknowledgements: I thank Julian Reiss, Wendy Parker, Nancy Cartwright, Peter Vickers, Robin Hendry, Donal Khosrowi, Tamlyn Munslow, Chien-Yang Huang, Richard Williams, and the rest of the CHESS team for their assistance in the development of this article. I also thank Rune Nyrup, John Norton and an anonymous referee for their help.

DIRECT INFERENCE IN THE MATERIAL THEORY OF INDUCTION

1. Introduction

John D. Norton's Material Theory of Induction (MTI) is a notable contemporary alternative to Bayesianism. According to Norton, scientists can make an inductive inference when they justifiably believe suitable local uniformity principles (which he calls "material postulates") that "licence" the inductive inference. I shall argue that there are several significant issues in the MTI, but these could be ameliorated by combining it with a theory of direct inference, and ideally one that is more systematic than the MTI.

In section 1, I describe the MTI. In the next sections, I examine the following problems for the MTI: (a) *lacunae* regarding evidential support, (b) the Problem of Induction, and (c) the Problem of the Reference Class. In the final section, I argue that if Norton had a suitable theory of direct inference, then he could resolve these problems.

2. The Material Theory of Induction

I shall begin with some preliminary clarifications. Firstly, I shall use 'induction' to mean non-deductive inference from observational evidence. This evidence does not have to be phenomenal: it need not describe sensations. In the sense I am using the term, 'observational' evidence consists of descriptions of facts, such as 'There is a red building across the street', that we can know via observation. This definition of 'induction' fits the use of this term by both Norton and his critics. (They do not define the term explicitly). This definition does not assume the existence of theory-free observation statements. It is distinct from Rudolf Carnap's very broad use of 'induction' to mean *all* non-deductive inferences (Carnap, 1962) which has been influential in confirmation theory. My use also differs from the understanding of 'induction' as inference from the particular to the general. Secondly, in accordance with much of the literature, I shall sometimes talk in terms of inductive *arguments*, but the object of our interest is typically scientists' inductive *inferences*; the former are linguistic entities, whereas the latter are psychological phenomena.

The best means to understand the MTI is by comparison with theories of induction in which inductive arguments are licenced by a *general* uniformity principle or relatively small set of such principles. According to Bertrand Russell (1912, Chapter VI), John Maynard Keynes (1921, 260), John O. Wisdom (1952, 162), Arthur W. Burks (1953), and many other philosophers, inductive arguments are justified due to an implicit premise (or premises) that nature is uniform, in some more precise sense of this claim. This premise guarantees that inductive inferences are reliable. These principles do *not* provide deductive certainty when they are added to the premises; instead, given a justified belief in the inductive evidence and in the absence of defeaters, the addition of a justified belief in the uniformity principles can enable a justified belief in the conclusion.

For example, Russell claims that the conclusions of our inductive arguments are probable given our beliefs because we assume that:

(1) If A and B are perfectly correlated, *n* is the number of observed things that are both A and B, and *c* is an unknown individual that is B, then it is (objectively) reliable to infer A*c* insofar as *n* is large.

(2) As *n* approaches infinity under the conditions described above, the objective probability of A*c* asymptotically approaches 1. (1912, 66.)

Norton's MTI is similar: the conclusion of an inductive argument is supported given the observational premises if we justifiably believe a suitable uniformity principle. (In contrast to Russell's theory of induction, the MTI is not probabilistic.) The salient divergence is that Norton's uniformity principles have a local scope, rather than describing the *general* uniformity of nature. They operate as implicit premises, akin to 'If P, then Q' in the enthymematic argument 'P, therefore Q' (Norton 2014, 674). And unlike the broad uniformity of nature principles, the exact contents of Norton's licensing principles vary from context to context.

Norton uses the example of the argument:

(1) All tested pure samples of the element bismuth melt at 271° C ± ε.

Therefore, (C_1) All pure samples of the element bismuth melt at 271° C ± ε.

(Where ε is an appropriately small margin of error. Measurement error is left implicit in Norton's discussion.)

Norton argues that this inductive argument is rational because we know the uniformity principle that 'Generally, elements are uniform in their melting points'. This principle is about elements, not nature *in toto*. Its explicit addition does not give the argument deductive validity, but according to Norton it "licences" the inference from (1) to (C_1). In contrast, Norton notes that the following argument seems unreasonable:

(1) All observed pure samples of wax melt at 91° C ± ε.

Therefore, (C_2) All pure samples of wax melt at 91° C ± ε.

According to Norton, the difference between the two arguments is that we lack a suitable local uniformity principle for wax in our background information. Indeed, we know that waxes include a variety of chemically distinct hydrocarbon mixtures. Therefore, even assuming we did not know that different types of wax have heterogeneous melting points, we would have some reason to suspect that they would have non-uniform melting points.

These arguments are enumerative inductions (or “universal inductions”) which are a limiting case of sample-to-population reasoning, in which one infers a universal hypothesis about a population from evidence about a sample. There are other types of sample-to-population reasoning: one might infer that “at least a majority” or “almost all and maybe every” or “48-52%” of a population have some characteristic, given the evidence of a sample. Like some other philosophers (such as D. C. Williams (1947, Chapter 5) Norton believes that others forms of induction, including demonstrative induction, eliminative induction, and Bayesian conditionalisation, are licensed by background knowledge of (1) uniformity principles, (2) universal generalisations, and (3) statistical generalisations, which have been inferred by antecedent sample-to-population inductions (Norton, 2003, 659-666)[1].

[1] Norton does not explicitly rule out the possibility that “material facts” other than local uniformity principles (which might be something like degrees of naturalness or hypothetical causal mechanisms) could licence inductions, but I have not yet located example of licensing via such principles in his work. Even naturalness and causality seem to be relevant to inductive justification because they might guide us about the uniformities and disuniformities in the actual universe or across possible universes.

Norton contrasts the MTI with what he calls “formal” theories of induction. In these theories, the rationality of good inductive arguments is determined by their formal structure. For instance, on a primitive hypothetico-deductive account of inductive inference, an inductive argument supports a hypothesis if the observational evidence can be validly deduced from the hypothesis and auxiliary background knowledge.

To summarise: on Norton’s view, our inductive inferences are rational if and only if we have suitable uniformity principles to licence these inferences. These uniformity principles are local, not general, in that they describe the uniformity of particular parts of nature, rather than all of nature.

Norton’s account has many strengths, which have been recognised even by critics like Thomas Kelly (2010, 758). The MTI fits well with standard descriptive accounts of scientific and technological practice. If you are criticising my naïve extrapolations from evidence about small subsamples of some political polling data, because you know that such subsamples are often unrepresentative of the target population, then you are making a *local* disuniformity claim, not a *global* disuniformity claim. Additionally, much of good experimental practice involves developing laboratory conditions that are sufficiently representative of the target populations, because the laboratory technicians believe that experimental conditions might differ from the natural phenomena of interest in important ways, even if the scientists do not doubt that (in some sense) nature in general is uniform.

The MTI also has normative plausibility. Firstly, it is unclear why I should be committed to the uniformity of *all* of nature in order to believe that *some* parts of it (such as observed samples of bismuth and bismuth samples in general) are uniform. Elliott Sober has persuasively argued for this point (1988, Chapter 2). Analogously, I do not have to believe that all of my family have XY sex chromosomes to believe that all the males do. Secondly, Norton has provided case studies of historical cases in which the MTI, in a notably unforced manner, licences *prima facie* good scientific reasoning. (See Norton 1994, 2003, 2010, 2011b.)

Norton's theory also has an appeal for some critics of Bayesianism, because prior probabilities and conditionalisation have no fundamental function in his theory of induction[2]. However, despites these attractions, I shall argue that there are still some important areas for improvement in Norton's theory, before proposing how to ameliorate them.

[2] An anonymous referee notes that MTI might also be compatible with some versions of Bayesianism, such as versions that only include the constraints from diachronic Dutch Book Arguments. Certainly, 'Bayesianism' is a very broad term and there are presumably forms of Bayesianism that are compatible with the MTI. Nonetheless, Norton has gone very far in developing a theory of induction that does not *depend* on Bayesianism.

3. Evidential Support

In this section, I shall argue that the MTI lacks answers to two crucial questions about evidence. I do not intend these questions as objections to the entirety of the MTI, but instead as points for further development.

3.1 Licensing by Uniformity

Norton makes liberal use of inferences that can be formalised as ‘A are generally B; *c* is A; therefore, *c* is B’ (Norton, 2014, 674). Such arguments enable us to infer claims about sample representativeness from local uniformity principles: given that elements are generally uniform in their melting points and bismuth is an element, we can infer that bismuth is uniform in its melting point, and therefore observed samples of bismuth are representative (with respect to their melting points) of bismuth in general.

However, despite its essential function in the MTI, Norton does little to justify this form of inference. He briefly discusses these inferences in (2014, 674) but mainly to claim that ‘A are generally B’ function in an analogous manner to ‘If A, then B’ claims in deductive logic. Norton says that, just as we are warranted in making a deductive inference from A to B when we accept that ‘If A, then B’, so we are warranted in inferring from A to

B when we accept ‘A are generally B’.

Even when the local uniformity belief is added as a premise to an inductive argument, this form of reasoning is still deductively invalid: “… the inference is ampliative, even with explicit adoption of the warranting fact as a premise.” (Norton 2014, 674). This entails that Norton cannot utilise the various justifications of deductive logical inferences, such as model-theoretic proofs of metalogical virtues like soundness and completeness[3]. These inferences are also not a subclass of inductive inferences, in the sense of ‘induction’ that Norton uses: ‘Ravens can generally fly, the next flock of birds I see will be ravens, therefore the next flock of birds I see will be able to fly’ includes a premise about ravens as a class, rather than a statement about *observed* ravens. Thus, their rationality is distinct from the issue of the rationality of induction, and consequently Norton’s arguments about the latter are not automatically applicable to the rationality of inferences from local uniformity principles to the hypothesis that particular samples are

[3] One might think that accepting the local uniformity principle itself provides sufficient warrant for the inference from the local uniformity principle to the hypothesis that the sample is representative (within a margin of error) of the target population. However, it will be apparent after section 4 that this is not true: even if accepting an appropriate local uniformity principle is necessary to justify inferences of the representativeness of a sample, it is not sufficient.

representative of the relevant target populations.

The mirror of this point is that a rich account of induction would give guidance about when it is irrational to accept local uniformity principles that can licence an inductive inference. Imagine that you are exploring the Antarctic and discover a large underground ecosystem. Among the animals are a species of penguins. You observe that all of the dozens of penguins that you see are blind and have white feathers. Intuitively, you might reasonably infer that almost all of this species are blind, because blindness might easily have evolved for the species in their subterranean existence. It would also seem to be irrational to infer (without further evidence) that ‘All of this species of penguins are white’, because you know that species of birds generally vary in the colour of their plumage. However, suppose that you postulated that ‘Subterranean species of penguins are generally uniform in their colour.’ Are you entitled to infer that all of the species of penguins are white?

The answer seems to be negative: we cannot just invent local uniformity principles. Naturally, Norton would agree, but it would enrich his theory to have an account of when we can and cannot make such inferences. (This account need not be a *formal* theory, in his sense of ‘formal’.) It would also help Norton answer critics who have theories of induction that have more extensive accounts of this reasoning, such as Bayesians, who can just say that we should consult the relevant conditional probabilities in our distribution.

3.2 Differences of Support

Local uniformity principles seem to licence inductive inferences to varying extents. Imagine that that you know (a) that 99% of the 200 balls in Urn 1 have the same colour and (b) that 50.5% of the 200 balls in Urn 2 have the same colour. Suppose you reach into Urn 1 and draw a grey ball, and then reach into Urn 2 and draw a grey ball. Assume that you have no background reason to think that the selected balls are unrepresentative (with respect to colour) of the class of balls in their respective urns. You seem to have more licence to infer that the next ball you draw from Urn 1 is grey than that the next ball you draw from Urn 2 is grey.

However, the MTI does not contain an account of how different sorts of evidence provide greater or lesser support for a given hypothesis, relative to different local uniformity principles. In the urns example, the differences in evidential support are obvious and might even seem to be quantitatively describable, but this transparency is not generally the case. This silence in the MTI is a problem, because we are not just interested in whether our inductive conclusions are supported by our evidence, but also *how well* they are supported. Perhaps a quantitative theory of evidential support is impossible. Perhaps we can only achieve a comparative or qualitative analysis. Perhaps there are multiple dimensions of evidential support, such as confirmation by the evidence, the reliability of the evidence, and the quantity of total evidence; these dimensions might be incommensurable. We do not know except insofar as we try. Bayesians have extensive analyses of these issues, and if the MTI has no alternative, then this is a respect in which the MTI *currently* compares unfavourably to Bayesianism.

From these gaps in Norton's theory, it does not follow that any rivals of the MTI do better (their accounts might be very flawed) but it does follow that the MTI would be enriched by supplementing it with the means of answering these questions. I shall return to this issue in section 6, in which I argue for a research programme for these additions.

4. The Problem of Induction

Many earlier discussions of the MTI criticise its implications for the Problem of Induction. In this section, I shall detail the critics' objection, and explain the dialectical position of Norton and his critics.

4.1 The Problem and Norton's Answer

There are many versions of the Problem of Induction; it would be more accurate to talk of the Problems of Induction. There is an interesting historical question of which version, if any, should be called the "Humean" Problem of Induction, but that question is beside the point of this article. Instead, I shall focus on the problem that Norton and his critics are discussing, which is interesting even if it is not Hume's problem.

Philosophers have almost all assumed that the justification of induction must be inductive or deductive. However, this presents a dilemma: (1) if we appeal to induction's past successes, then our justification is circular, since our justification involves an inductive inference, while (2) if we appeal to some implicit premises that render inductive arguments into deductively valid arguments, then our justification is too strong, because inductive inference is ampliative rather than deductive[4]. We have a paradox: we believe that induction is reliable, but this belief is very hard to justify. (Norton, 2003, 666).

Norton's answer is that each (rational) inductive inference is justified by a local uniformity principle, rather than the past successes of induction in general. For our actual inductive inferences, he argues that we should take horn (1) of the dilemma: our inductions are justified by past inductions, but these antecedent inductions were inferences of *local* uniformities. Obviously, there is a suspicion of a vicious regress. Norton acknowledges that there is a regress in his justification, but he denies that it is vicious, because there is no

[4] This is how Norton puts the problem for deductive justifications of induction. A more typical criticism is the claim that the even the weakest relevant suppressed premises (what D. C. Stove calls "validators" (1986, 12) cannot be known *a priori*, nor *a posteriori* without a justification that assumes induction's rationality.

reason to suppose that its termination must be problematic. In (2003, 668) he argues that the "branching trees" of justification might end in "brute facts of experience that do not need further justification". In (2014, 683-687) he argues that the regress might produce a benign, coherent circle[5]. Briefly, according to Norton, our inductions are justified by past inductions, but there does not need to be a vicious regress.

4.2 The Regress Problem

Philosophers of science have mostly been unsympathetic to Norton's answer to the Problem of Induction. For example, Samir Okasha (2005, 250), Thomas Kelly (2010, 760-763), and John Worrall (2010, 746) have all argue that the regress must be vicious. Their arguments assume a first inductive inference, prior to any non-observational knowledge, to which all other inductive inferences have an epistemic link. (This was a fair assumption, because Norton (2003, 668) accepts the possibility of a first inductive inference.) Their implicit argument seems to be:

[5] Norton does not explain why the circularity of traditional inductive justifications of induction is a problem, whereas the potential circularity of the corpus of science in the MTI would be benign.

(P$_1$) All justified inductive inferences must be licensed by local uniformity principles.

(P$_2$) All local uniformity principles must be justified by earlier inferences.

(P$_3$) Local uniformity principles can only license inductive inferences if these principles are justified by antecedent justified inductions.

Therefore, **(C$_1$)** If there was a *first* inductive inference and this inference was justified, then it would have to be licensed by at least one local uniformity principle, which in turn would have to be justified by an inductive inference that was justified.

Therefore, **(C$_2$)** If there were a first justified inductive inference, then there would be a justified inductive inference prior to the first justified induction, but this is contradictory.

Therefore, **(C$_3$)** There was no first justified inductive inference.

But, **(P$_4$)** If there are *any* justified inductive inferences, then there must have been at least one first justified inductive inference. (There may have been more than one.)

Therefore, **(C$_4$)** There are no justified inductive inferences.

If Norton's critics are correct, then the MTI and some plausible auxiliary premises entail inductive scepticism. This entailment would not be entirely surprising: Okasha (2005, 249) notes that Sober (1988, Chapter 2) has a theory of induction that is similar to Norton's, but he comes to a *sceptical* conclusion about induction. Since Norton denies inductive scepticism, he needs some way of avoiding C_4. Norton commits to P_1 from the outset of his primary article on the MTI (2003) and it captures much of the MTI's normative core. All parties in the controversy seem to accept P_3. However, Norton does deny P_2 (Norton, 2014, 679-680) and P_4 could also be challenged[6].

Norton provides many interesting arguments against this premise, but they stray into controversial issues in the philosophy of science and general epistemology. In his fuller answer to the Problem of Induction Norton increasingly relies on theses that are beyond the MTI (2014), which raises of the question of whether the MTI really requires so

[6] An anonymous referee notes that P_4 apparently disallows the possibility that an inductive inference could be justified by its connections with simultaneous or subsequent inferences. Indeed, we might think that the chain of inductive inferences is analogous to theories of cosmology in which there is a causal chain but no first cause. Norton's critics do need *some* premise like P_4 to connect C_3 with C_4, because C_3 on its own has little, if any, sceptical weight, and bridging principles between these conclusions cannot be much weaker than P_4. Further critical discussion of P_4 and the other premises, aside from P_3, is beyond the scope of this article.

many background assumptions to be tenable. Furthermore, his arguments that the MTI does not lead to some form of inductive scepticism have not convinced many philosophers.

In section 6, I shall argue that Norton has access to an answer to the Problem of Induction that requires only the MTI, a plausible emendation to P_3, and a theory of direct inference. Before I discuss this answer, I shall turn to one final outstanding issue for the MTI, partly because addressing this issue is a requirement of the answer in section 6.

5. The Problem of the Reference Class

As we have seen, Norton makes very considerable use of inferences of the form:

(1) Generally, Φ is Ψ.

(2) *x* is Φ.

Therefore, (3) *x* is Ψ.

It is by this form of inference that local uniformity principles "licence" inductions, but Norton does not discuss this type of inference in much detail. However, in a recent paper he does say that "The general [local uniformity principle] cited above functions both as a factual statement and a warrant for an inference... The [local uniformity principle]

“Generally, such and such.” warrants the inductive inference to “such and such.”” (Norton, 2014, 674.)

Taken literally, this claim is far too permissive. Consider the case of carbon: you know that this is an element, and that generally, elements have uniform melting points. However, if you also know that carbon is a member of the subset of elements that do *not* have uniform melting points, then you are *not* warranted in inferring that its melting point is uniform, and therefore you are not warranted in inducing from (a) the approximate observed melting point of a sample of carbon to (b) the approximate melting point of all carbon samples. It is not merely that such inferences are defeasible, but rather that it would be useful to know when the defeaters apply. (Bayesians have an account of when the defeaters apply.) Similarly, you know that ‘Generally, humans cannot run 100 metres in under 15 seconds’, but you are not warranted in inferring that Usain Bolt will be unable to run 100 metres in under 15 seconds in his next race, because you know that he is a world-class sprinter and that they can generally run that fast. When does such background knowledge defeat our inferences from (1) ‘Generally, A is B; *c* is an A’ to (2) ‘*c* is B’?

Such inferences are often called “direct inferences”, but their other names include “statistical syllogisms” and “proportional syllogisms”. None of these names are perfectly apposite, but I shall use “direct inference” in the hope that it is the most felicitous. In a simple direct inference, we reason from (1) a premise about the proportion or measure[7] of a characteristic in a reference class and (2) a premise that the subject of the conclusion is a member of this reference class, to (3) a conclusion that the subject of the conclusion does or does not have this characteristic. ‘Generally’ is one possible quantifier in premise (1), but we could also use other qualitative terms like ‘Almost always’ or ‘Almost never’, as well as numerical terms like 75%, or intervals like ‘5-10%’. For instance:

Example 1

(1) 99% of the time, Φ is Ψ.

(2) *x* is Φ.

Therefore, (3) *x* is Ψ.

Example 2

[7] By “measure”, I mean the value of an aleatory probability function (whether additive or non-additive) for that event or type of event. Some statistical statements cannot be expressed in terms of proportions, because they refer to reference classes of infinite cardinality and it does not make sense to speak of proportions in infinite reference classes. However, there can be statistical statements that describe measures of these classes.

(1) 95-99% of the time, θ is ζ.

(2) x is θ.

Therefore, (3) x is ζ.

Example 3

(1) λ is almost never ω.

(2) x is λ.

Therefore, (3) x is not ω.

Critics of the MTI could object that despite his reliance on direct inference, Norton has not answered the Problem of the Reference Class, which was first developed in detail by Hans Reichenbach (1949, 374). We always know that the subject of a direct inference is a member of a variety of reference classes: carbon is an element, but also a non-allotropic element; Usain Bolt is a human, but also a world-class sprinter.

As Alan Hájek (2007) has argued, the Problem of the Reference Class challenges a great variety of theories in formal epistemology; it is not merely a flaw of frequentism. Furthermore, the Problem of the Reference Class is not just an issue for probabilistic theories of support. (The MTI is not a probabilistic theory.) Carl G. Hempel (1965, 53-62) develops versions of the problem for a variety of non-probabilistic theories of direct inference, such as Stephen Toulmin's account[8]. Hempel notes that an insufficiently restricted use of direct inferences enables us, from the same body of data, to infer that:

(1a) The proportion of Roman Catholics among Swedes is less than 2%.

(2a) Petersen is a Swede.

Therefore, (C_1) Petersen is not a Roman Catholic.

- but also -

(1a) The proportion of non-Roman Catholics who undertake a pilgrimage to Lourdes is less than 2%.

(2b) Petersen underwent a pilgrimage to Lourdes.

Therefore, (C_2) Petersen is a Roman Catholic.

Thus, different premises can support contrary hypotheses by direct inference, *even if the premises are true*. Even if we interpret the support relation as non-probabilistic, we shall have to reckon with the problem of deciding whether C_1, C_2, or neither is supported

---

[8] Hempel, following Carnap's use of the word 'inductive' to mean non-deductive reasoning, labels these problems for direct inference "inductive inconsistencies". However, in the narrow sense of inference from *observational* premises to non-observational premises, direct inferences are not generally 'inductive'.

when our data includes (1a), (2a), (1b), and (2b). More generally, we shall have to reckon with adjudicating between conflicting evidence about the reference classes to which the target individual belongs: the Problem of the Reference Class.

Norton does not explain how we should achieve this adjudication. His only detailed statement about direct inference, quoted above, is far too permissive: I am sure that he does not mean it literally, but that leaves many important points undeveloped. Firstly, it is insufficient to add ‘And all you know about bismuth is that it is an element’, because this never obtains in any epistemologically significant case. For instance, you know that bismuth is an element *and* that it is a substance, but the frequency of uniform melting points among substances is different from the frequency among elements. Additionally, you must also have some means of distinguishing bismuth from other substances in order to know that bismuth is an element, which requires background knowledge about bismuth. Secondly, it is insufficient to add ‘And you do not believe any defeaters for this inference’, because this does not explain how to identify the defeaters, nor how to proceed when there *are* defeaters. Finally, the presence or absence of defeaters is not always apparent, which is why philosophers can disagree about direct inference. For example, the edited volume Bogdan (1982) covers part of a long-running debate between Henry Kyburg and Isaac Levi, in which they disagreed about particular direct inferences.

These are not just hypothetical issues. Suppose you have good reasons to believe that bed nets are *generally* effective at preventing malaria, but not in villages where the inhabitants often use the bed nets for fishing. If you also have good reasons to believe that a particular village will use the bed nets in this way, then you cannot infer from the general effectiveness of bed nets to the conclusion that bed nets will be effective in that village. Similarly, you might have good reasons to believe that class-size reduction is a *generally* effective means of improving average K-3 grade levels, but you might know confounding facts about Californian schools that would defeat the inference that adopting this policy will improve the state's average K-3 grade levels. Nancy Cartwright and Jeremy Hardie (2012) contains more examples of the use and misuse of direct inference in evidence-based policy.

Despite these disagreements and misuses, it does seem that there are at least some norms about direct inference. For example:

(1) Suppose you know that I am about to select a ball from a glass vat. You know that almost all the balls in the vat are red. However, you also know that the red balls are larger than the purple balls, and my selection mechanism uses an automated suction pump/blower that will only allow black or green balls through. The direct inference is blocked by this additional knowledge. Intuitively, when hypothesising about a selection from a set, you should ignore the evidence about the general proportion in the set and focus on your evidence about the frequencies of selections from that set, when these statistics conflict.

(2) Imagine you are wondering whether a student from your logic course can help

regarding a problem in integral calculus. You know that almost none of the students in your university could help. However, suppose you also know that the student has won prizes in several calculus courses. The specific details about the student and her aptitude at integral calculus suggest that she might be able to help, although most of the university's students cannot. Intuitively, subset information should be favoured over information about the set, when these statistics conflict.

Something like these intuitions have been shared by a variety of philosophers writing about direct inference: for example, (2) seems to motivate Reichenbach's claim that one should use the narrowest reference class for which one has reliable statistics (1949, 374) and Wesley Salmon's claim that one should use the narrowest reference class that meets his suitability conditions for statistical inference (1967, 90-94). A similar idea can also be found in Kyburg (2006, 45-47). There are agreements as well as disagreements about direct inference, so that philosophical theories about them can be evaluated using common ground, yet also helpful in resolving disputes.

Since Norton makes use of the licensing of inductions via information about local uniformities in his MTI, it would enrich his account to have an answer to the Problem of the Reference Class, and thus leave direct inference less mysterious. Furthermore, critics of the MTI might worry about the cogency of the theory as a whole, given that it *might* preclude any answer the Problem of the Reference Class. For instance, he rules out the

Bayesian answers, in which the relevant statistics for the intersections of the reference classes are derivable from the prior distribution.

6. Direct Inference

In this section, I shall argue that *if* an advocate of the MTI had a theory of direct inference, then the issues that I raised in the previous sections could be addressed. I shall set aside the broad and difficult question of which theory to use, but I shall have motivated the search for a supporter of Norton's theory. Logically, the MTI is consistent with a wide range of theories of direct inference. Furthermore, insights from the MTI could help guide the choice. The selected theory would not have to be formal, nor would a formal theory of direct inference (rather than induction) be inconsistent with the MTI, but I shall argue that a systematic theory would be particularly attractive.

6.1 The Problem of the Reference Class

The most important task of a theory of direct inference is answering the Problem of the Reference Class. However, the name of this problem is imperfect: it suggests that the theory should identify *one* statistical statement about *one* reference class, which can be rationally included as the statistical generalisation in the direct inference. As an alternative, we might adopt a theory of direct inference in which a direct inference might involve a *set* of conflicting statistical statements as premises. To avoid begging this question, a good theory of direct inference would provide systematic answers to these questions:

(1) When should we exclude statistical statements from the premises? For example, in the bismuth example, there seems to be a rough intuition that we should ignore the statistical information that substances are not generally uniform in their melting points, in favour of the statistical information that elements are generally uniform in their melting points.

(2) If there are conflicting statistical statements and there are no grounds for ignoring some of them, then how should we combine them? Bayesians have an account: we should look at our full probability distributions and use Bayes' Theorem. Norton has vigorously argued for rejecting Bayesianism as a general theory of the scientific method (Norton, 2011a) but that still leaves non-Bayesians like Norton and myself requiring an answer.

One might wonder why such a theory must be systematic. Perhaps our intuitive sense of appropriate or inappropriate uses of reference class data are sufficient. However, people's intuitions (including philosophers' intuitions) can differ for particular cases of direct inference. Setting up clear systems of direct inference enables us to evaluate and constructively critique the systems that might embody these divergent intuitions, which shifts the dispute onto more promising ground than personal feelings about specific cases. Furthermore, we sometimes make bad direct inferences, due to prejudice, wishful thinking, fatigue, and so on. A systematic theory can help us identify these mistakes. Finally, as in deductive logic, there could be complex or difficult cases where our intuitions are unclear or absent, but a systematic theory can provide definite answers that we have deduced from plausible principles.

The mere fact that a systematic theory would be desirable does not entail that it is possible. Perhaps, as with induction, Norton might suspect that context is king: there is nothing to say about direct inference that is both very general and very detailed. Certainly, there can be more or less contextually focused theories of direct inference, but there do seem to be at least *some* general and plausible normative claims about direct inference. For example, many philosophers (like Reichenbach and Kyburg) have endorsed the principle that, if there is a conflict between two acceptable statistical statements for a target characteristic and a given individual, and one statement describes a reference class that is known to be a proper subset of another, then that statement about the proper subset should be favoured over its competitor. That is far from a general theory of direct inference, but it at least provides some systematic guidance, and implies that at least *some* degree of systematicity is achievable. Furthermore, the claim that induction is highly context-dependent is logically independent of whether direct inference is highly context-dependent, just as both are independent of whether deduction is highly context-dependent.

How we should decide between rival theories of direct inference? Most of our classic metalogical criteria will not be helpful. For example, we know that direct inferences can have true premises and false conclusions, and so no theory of them can guarantee truth-preservation. Philosophers who have developed theories of direct inference have taken different approaches to their evaluation. One promising example is the approach to theory evaluation developed by Kyburg and Choh Man Teng (2001, Chapter 10). They propose to evaluate formal theories of direct inference using similar criteria to those used for the evaluation of formal deductive logics, such as metalogical soundness. However, these metalogical standards are adapted to the specific ambitions we have for our direct inferences, in contrast to deductive inferences. For example, one test that Kyburg and Teng propose is that, if a system of direct inference says that a conclusion is highly supported given the premises, then the conclusion should hold in a high proportion of the models of the relevant premises. It seems that we can rationally assess systems of direct inference, and choose between better or worse systems. I shall now turn to how such a systematic account of direct inference could improve the MTI.

6.2 Evidential Support

I shall now return to the two gaps concerning evidential support in the MTI that I identified in section 2. Firstly, there was the issue of when we can use direct inferences. Sometimes, as in Norton’s bismuth case, our background knowledge seems to be suitable to infer a local uniformity from a claim like ‘Generally, elements have uniformities in their melting points’. Yet we do not want scientists to arbitrarily accept local uniformity data to licence just any inductive inference. A theory of direct inference can aide our understanding of illicit inductive reasoning, by both (a) providing standards of when local uniformity data is acceptable as licensing background information for an inductive inference and (b) providing guidance for when epistemically acceptable local uniformity principles can be ignored for inferential purposes.

Secondly, there was the issue of degrees of support. Take the schema:

(1) To an extent of r, samples with characteristics Φ are representative (within a margin of error) of Ψ with respect to a characteristic F.

(2) x is a sample of Φ.

Therefore, (3) x is a representative sample of Ψ with respect to F.

(Where r is either (i) a real or interval value for a proportion or measure, or (ii) some qualitative qualifier like ‘generally’ or ‘almost always’. “Representative” here refers to a variety of possible senses of the term, such as having a mean value of F that is no more than 3% from the population mean.)

A theory of direct inference should provide an account of when we have warrant to use the schema above. Perhaps the value of r provides the relevant information about the degree confirmation for the hypothesis that the observed sample in our evidence is representative, given the information in the premises. To the degree that we have warrant to believe that our samples are representative, we are warranted in making inductive inferences from evidence about the properties of those samples to our target populations. Other forms of inductive inference, such as eliminative induction and predictions about particular unobserved individuals, would derive their warrant indirectly from these generalisations from samples.

Inevitably, my proposals for addressing these gaps in the MTI are sketchy, because we do not have a generally accepted "off-the-shelf" theory of direct inference. However, if we could combine a systematic theory with the MTI, then the concomitant account of induction would be much richer and less vague.

6.3 The Problem of Induction

I noted in section 3 that Norton's claim that his MTI dissolves the Problem of Induction was controversial. However, at the end of that section, I suggested that Norton was very close to a very promising answer. As a preliminary point, note that both Norton's critics (Okasha, Kelly, Worrall, and others) have not raised any objections to his use of direct inferences in the MTI. Even Hume takes a view of direct inference that is excessively *anti-sceptical*: "If you suppose a dye to have any bias, however small, to a particular side, this bias, though it may not appear in a few throws, will certainly prevail in a great number.” (Hume, 1793, 112.) Of course, there is no guarantee that even a very large finite sample of die rolls will match the long-run actual frequency, hypothetical frequency, or propensity of that die. Nonetheless, the quote from Hume illustrates how even sceptics about induction may reject scepticism about direct inference. Thus Norton, Hume, and contemporary critics of the MTI have common ground in their belief that direct inference can be rational. This agreement will be crucial in my arguments in this section.

Let us return to the original set-up of the Problem of Induction that Norton and his critics use. There is a dilemma between a deductive justification of induction (which is too strong) and an inductive justification of induction (which is circular). Attention to direct inference suggests that this is a false dilemma, because direct inference is neither deductive nor inductive[9].

There is no generally accepted name for justifying induction via direct inference, but I shall call it the “Combinatorial Justification of Induction” (CJI for short), following Marc Lange (2011, 83). This answer to the Problem of Induction was first proposed by D.

[9] Except in the Carnapian sense of non-deductive inference. If we adopt Carnap’s usage, there is no circularity in an inductive justification of inference from the observed to the unobserved, and thus of the Problem of Induction requires considerable rephrasing for its power to be apparent. This is one reason not to adopt Carnap’s usage.

C. Williams (1947). It has subsequently been defended by a number of philosophers of science, including David Stove (1986, Chapter VI), Timothy McGrew (2001), and Scott Campbell and James Franklin (2004).

Norton, his critics, and Hume all grant that, at least under some circumstances, contingent local uniformity principles can licence inductive inferences via direct inference. This raises the possibility of whether there is a local uniformity principle, knowable *a priori*, that could licence inductive inferences. That would entail that, even in a hyper-austere evidential state where we have only observational evidence and *a priori* evidence, there could still be justified inductions.

I shall begin by defining a key term:

ε-representativeness: A subset Z is ε-representative of its superset W with respect to a characteristic Φ if and only if the mean of Φ in Z differs by no more than ± ε from the mean of Φ in W, where ε is some finite, real, non-zero percentage value, such as 3%.

From combinatorics, we know that the size of *n* puts a lower bound on the number of ε-representative *n*-fold subsets of any finite set[10]. To see the intuition in a simple

[10] In particular, proponents of the CJI refer to the (weak) Law of Large Numbers. This is correct, *if* we interpret the Law of Large Numbers as a purely combinatoric principle for

example, imagine a million-fold set of soft toys in a glass vat. If *all* the toys are blue teddy bears, then every 1,000-fold subset of soft toys from the vat will be ε-representative with respect to being blue teddy bears, because the proportion of blue teddy bears in both the subsets and the superset will be 100%. If *none* of the toys are blue teddy bears, then once again every subset will be ε-representative, because both the subset and superset will lack blue teddy bears. If 99% or 1% of the toys are blue teddy bears, then there will be a significant proportion of 1,000-fold subsets that are not ε-representative, but the number of 1,000-fold subsets that are ε-representative will still be *proportionately* high. The proportion of subsets that are ε-representative reaches at a minimum if exactly 50% of the soft toys are blue teddy bears. Put another way, if the proportion of blue teddy bears among toys is the vat in the vat is precisely 50%, then there are the highest possible number of subsets of that vat that are *not* ε-representative.

Williams, who first developed the CJI, was struck by the combinatorial fact that, even when exactly 50% of a finite set have a characteristic Φ, the proportion of large subsets of that set which are ε-representative is still extremely high. To adapt an example from Stove, if we define ε as 3%, make the anti-inductive assumption that 50% of the soft toys in the vat are blue teddy bears, and we use theorems of combinatorics to calculate the

---

finite subsets and their supersets, rather than as a principle in the probability calculus that applies only to independent and identically distributed draws of samples from populations. Naturally, it is the latter version of the Law that normally interests practicing statisticians.

approximate proportion of ε-representative 3,000-fold subsets of the toys in the vat, the answer is $10^{8867.9-0.00087} \div 10^{8867.9}$. This proportion is astonishingly high: it is over 99%. Even if the subsets have a cardinality in the hundreds, the *minimum* proportion of ε-representative subsets can still be over 50%.

It is important to note that it is not the *relative* size of the subset to the superset that drives this result. Instead, the lower bound for the proportion of subsets that are not ε-representative is determined by (1) the *absolute* size of the subsets and (2) the fact that the superset is finite. (Infinite supersets might not even have a well-defined mean – for example, if its limiting frequency distribution is a Cauchy distribution – so Williams's reasoning does not apply to all of them.) Thus, these combinatoric facts are also true of 3,000-fold subsets of a googolplex-fold superset, just as they apply to 3,000-fold subsets of a million-fold superset[11].

The CJI proceeds as follows: suppose that you are inquiring about a set Z, where Z is defined to be a finite set. (To use an extreme example, 'The nearest group of toys near to me that has a cardinality no greater than the square of Graham's Number'. The entire observable universe is far too small to digitally represent the maximum possible cardinality

[11] We can estimate a lower bound with a higher value, if we know that the subset is large relative to the superset. At the limit, if we know that the subset's cardinality is equal to its superset's cardinality, then we know (trivially) that the subset is ε-representative.

of this set, though the set is still finite.) Suppose that *all* you know about an observed sample S of Z is that it is a large (such as 500-fold or 3,000-fold) subset of Z. By direct inference:

(1) Generally, large subsets of Z are ε-representative.
(2) S is a large subset of Z.
Therefore, (3) S is ε-representative.

If you justifiably believe that S is ε-representative of Z, then you can use the observed characteristics of S to infer the characteristics of Z, with a margin of error of ε. (Obviously, background knowledge can defeat this inference.) Furthermore, the larger the cardinality of S, *ceteris paribus,* the stronger the qualifier that can be inserted in (1). Depending on the sample S's cardinality, one could replace the qualifier "generally" by 'almost all', 'over 80%', 'over 99%' and so on. Even for small samples, we have the result that as *n* increases from 1 towards infinity, the subset becomes a member of a class of subsets with higher and higher proportions of representative subsets of any finite set to which they belong. Therefore, even relative to ultra-exiguous background information, one can justify some inductive inferences, and identify some cases of inductive evidential support.

The formal similarity to Norton's bismuth argument is striking. Moreover, direct inference and combinatorics are very solid foundations for induction. However, if Norton were to adopt this approach to the Problem of Induction, he would have to give up an assumption he (apparently) makes, which is that the licensing local uniformity principles are always contingent. The combinatorial principles that Williams and his philosophical descendants utilise are mathematical and therefore non-contingent according to most philosophies of mathematics. However, they are still local uniformity principles, in the sense that they describe uniformities (the proportion of ε-representative samples) in local domains of inquiry (the target populations of the inductive inferences). Yet, since both Norton and his critics accept direct inference, under appropriate circumstances, it is hard to see why Norton would reject direct inference in this case.

Of course, there are dozens of good extant objections to the CJI. A sharp reader has probably already formulated at least three. Some of these objections have proven relatively easy for defenders of the CJI to address. For instance, it is tempting to assert (though it is very hard to argue rigorously) that samples must be random before we can make an inductive inference, but there are detailed responses to this objection (McGrew, 2001 and Campbell and Franklin, 2004). The persuasiveness of this objection seems to come from a confusion of *sets* and *samples*. This confusion is encouraged by many proponents of the CJI, who use statistical expressions like “samples” and “populations” rather than set-theoretic expressions like “sets” and “subsets”. Talking about samples and populations naturally raise concerns about sample selections in general and can even lead to unwarranted confidence regarding sampling procedures, but it is inessential to the CJI.

The most difficult objection to the CJI seems to be the Problem of the Reference Class. We are never in the position where ‘Generally, large subsets of Z are ε-representative’ is our only information about a sample S. Even to distinguish S and its members, we must know some characteristics that differentiate it from other things that we have experienced: we need a definite description for our sample to individuate it, such as ‘That subset of soft toys from the vat *that I have in front of me right now*’. Therefore, if the CJI has any epistemological use, we would also need some means of picking out a proposition of the form ‘Generally, large subsets of Z are ε-representative’ as a suitable local uniformity principle for a direct inference of a observed sample’s ε-representativeness.

This problem is more pressing, *a fortiori*, if the CJI is to justify of actual science. A defender of the CJI would have to combine it with an account of how the beautifully complex methods of science can justified by the CJI, and the gigantic corpus of scientific knowledge justified in turn by these methods. Perhaps these methodological accounts are possible (Williams provides a brief and telegraphic sketch in Chapter 5 of (1947)) but in any detailed methodological account, there will be bodies of evidence that provide huge masses of statistical data with conflicting messages about members of multiple reference classes. Therefore, a defender of the CJI must reckon with the Problem of the Reference Class.

However, *if* we could have a theory of direct inference that could answer the

Problem of the Reference Class, and *if* we could answer the other objections to the CJI, then Norton could deny P_3 in the regress argument, that ‘Local uniformity principles can only license inductive inferences if these principles are justified by antecedent justified inductions.’ I do not know any good arguments for P_3, but it seems that philosophers of science have often assumed the truth of P_3 or something close to it. In practice, almost all good inductions *would* involve antecedently induced local uniformity principles, but there could be exceptions. Thus, Norton can adopt the following consistent group of claims:

(1) All justified inductive inferences must be licensed by local uniformity principles. (P_1 of the regress argument.)

(2) All local uniformity principles must be justified by earlier inferences. (P_2 of the regress argument.)

(3) Local uniformity principles can only license inductive inferences if these principles are justified by antecedent justified inductions or by direct inferences. (The revised version of P_3 in the regress argument.)

(4) If there are *any* justified inductive inferences, then there must have been a first justified inductive inference. (P_4 of the regress argument.)

(5) There are *some* justified inductive inferences. (The denial of inductive scepticism.)

A sceptic might object that the CJI involves such artificially exiguous background knowledge that any relation to real inductive reasoning is absent. In practice, we often have reasons (beyond combinatorial considerations) to suspect that our samples are unrepresentative: for example, I know that an opinion poll on the Dodd-Frank financial regulations, using a sample consisting only of stock traders at the New York Stock Exchange, will almost certainly be unrepresentative of US public opinion in general. Yet one of the insights from Norton's studies of induction is that more realistic scenarios can be *more* favourable to inductive reasoning, because we can have background reasons (beyond combinatoric reasoning) to think that in some cases our samples are representative, as well as evidence to suspect them in other in other cases. For example, we might have an opinion poll that has been carefully weighted using pollsters' background knowledge of response rates and other relevant information, such that it is likely to be representative. The CJI would be useful for the MTI because it would prove that even if such favourable background knowledge were absent, there could still be justified inductive reasoning.

Thus, direct inference is very promising against inductive scepticism in both (1) the epistemically rich contexts Norton analyses, where the MTI provides a rationale of inductions' justifications and (2) the exiguous contexts that philosophers like Williams contemplated, where the CJI might provide the rationale. However, all this assumes that direct inference is not itself as troubled as induction. That is why a theory of direct inference has such value.

Would Norton consider my arguments to be helpful support in his battle against inductive scepticism? The CJI avoids the abstract and sweeping formalistic claims that Norton rejects (it uses defeasible direct inferences and the local properties of specific populations) while also providing a means of ending the regress of local uniformity principles *in some contexts*. This last qualification is important, because the CJI's dependence on direct inference entails that it is defeasible: additional background information can block the relevance of purely combinatoric considerations. However, in the hyper-exiguous contexts that the critics of Norton employ in their sceptical arguments, it is supposed (by design) that there is no background information available (beyond what is knowable via mathematics, logic, and other conceptual reasoning) since his critics aim to bar any appeal to contingent background information about samples' representativeness. Ironically, it is the very austerity of the sceptical scenario which makes the CJI so promising for a benign termination of the regress in the MTI. The CJI's usefulness for answering the Problem of Induction will depend on how we conceptualise that problem; my point is that even if we formulate the challenge in the manner of Norton's critics, the CJI is a promising option within the MTI.

Additionally, amending P_3 as have suggested would put the MTI in a strictly stronger dialectical position with respect to the Problem of Induction, because Norton's present responses to his critics (as well as other criticisms of P_2, P_3, and P_4) would still be available, but the CJI would also be available. The CJI might not be viable: as noted, there are many objections to it, especially regarding the finer points of direct inference. Still, there is nothing in the letter of the MTI that is inconsistent with the CJI, and much that could be gained from their concatenation.

7. Conclusion

I have argued that there are holes in the MTI, but also that a theory of direct inference can fill these holes. I have not argued for any particular theory of direct inference. I suspect that Norton and other supporters of the MTI could improve on even the best options in the literature. As Norton has argued, local uniformity principles are crucially important to understanding inductive inference. To understand the role and justification of these local uniformity principles, we need a theory of direct inference.

Norton has developed one of the most propitious non-Bayesian theories of induction in many years. If I am correct, then the study of direct inference is an area of great pertinence and promise for those of us who share his objectives for the philosophy of induction.

REFERENCES

Bogdan, Radu J. 1982. *Henry E. Kyburg, Jr. & Isaac Levi*. Dordrecht, Holland; Boston : D. Reidel; Hingham, MA: Kluwer Boston Inc.

Burks, Arthur. W. 1953. “The Presupposition Theory of Induction.” *Review of Metaphysics*, 20 (3): 177-197.

Campbell, Scott and Franklin, James. 2004. “Randomness and the Justification of Induction.” *Synthese*, 138 (1): 79-99.

Carnap, Rudolf. 1962. *The Logical Foundations of Probability*. London: Routledge & Kegan Paul.

Cartwright, Nancy and Hardie, Jeremy. 2012. *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford: Oxford University Press.

Hájek, Alan. 2007. “The Reference Class Problem Is Your Problem Too.” *Synthese*, 156 (3): 563-585.

Hempel, Carl. G. 1965. *Aspects of Scientific Explanation, and Other Essays in the Philosophy of Science*. New York: Free Press; London: Collier-Macmillan.

Hume, David. 1793. *Essays and Treatises on Several Subjects, Vol. I: Essays, Moral, Political, Literary*. Edinburgh and London: T. D. Cadell and Bell & Bradfute and T. Duncan.

Kelly, Thomas. 2010. “Hume, Norton, and Induction without Rules.” *Philosophy of Science*, 77 (5): 754-764.

Keynes, John Maynard. 1921. *A Treatise on Probability*. London: Macmillan.

Kyburg, Henry E. 2006. “Belief, Evidence, and Conditioning.” *Philosophy of Science*, 73, (1): 42-65.

Kyburg, Henry. E. and Teng, Choh. M. 2001. *Uncertain Inference*. Cambridge: Cambridge University Press.

Lange, Marc. 2011. “Hume and the Problem of Induction.” In *Handbook of the History of Logic. Volume 10: Inductive Logic*, eds. Gabbay, Dov. M., Hartmann, Stephan, and Woods, John, 43–91. North Holland: Elsevier.

McGrew, Timothy. 2001. “Direct Inference and the Problem of Induction.” *The Monist*, 84, (2): 153-178.

Norton, John. D. 1994. “Science and Certainty.” *Synthese,* 99, pp. 3-22.

Norton, John. D. 2003. “A Material Theory of Induction.” *Philosophy of Science*, 70 (4): 647-670.

Norton, John. D. 2010. “There Are No Universal Rules for Induction.” *Philosophy of Science*, 77 (5): 765-777.

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

Norton, John. D. 2011b. “History of Science and the Material Theory of Induction: Einstein's Quanta, Mercury's Perihelion.” *European Journal for Philosophy of Science*, 1 (1): 3-27.

Norton, John. D. 2014. “A Material Dissolution of the Problem of Induction.” *Synthese*, 191 (4): 671-690.

Okasha, Samir. 2005. “Does Hume's Argument against Induction Rest on a Quantifier-Shift Fallacy?” *Proceedings of the Aristotelian Society*, 105: 237-255.

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

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

Salmon, Wesley C. 1967. *The Foundations of Scientific Inference*. Pittsburgh, PA: University of Pittsburgh Press.

Sober, Elliot. 1988. *Reconstructing the Past*. Cambridge, MA: MIT Press.

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

Williams, Donald. C. 1947. *The Ground of Induction*. New York: Russell & Russel, Inc.

Wisdom, John. O. 1952. *Foundations of Inference in Natural Science*. London: Methuen and Company.

Worrall, John. 2010. “For Universal Rules, Against Induction.” *Philosophy of Science*, 77, (5): 740-753.