

DOES THE DOME DEFEAT THE MATERIAL THEORY OF INDUCTION?

Critics of John Norton's Material Theory of Induction (MTI) have mostly focused on its relation to the Humean Problem of Induction (Okasha, 2005, p. 250). However, Hume's challenge is just one of many philosophical issues about induction. Thomas Kelly (2010, pp. 757-758) plausibly argues that the natural place to challenge Norton's theory is where apparently rational inductions occur without background knowledge of local facts, because Norton claims that such knowledge is essential for reasonable inductive inference.

I shall defend the MTI against a notably well-developed criticism along these lines by Richard Dawid (2015). He argues that Norton's theory fails to allow for an intuitively good inductive inference. My defence will not be very dependent on particular features of Dawid's example, and thereby it reveals the MTI's robustness to analogous criticisms. In Section 1, I set out Norton's theory. In Section 2, I outline Dawid's objection. In Section 3, I adapt an attempted justification of induction *as such* into a justification of the particular induction that Dawid discusses. I finish by answering some objections in Section 4.

SECTION 1 – The Material Theory of Induction

Norton has presented the MTI in several places (2003, 2010, 2014, unpublished). Its core content consists of the following claims:

Ampliative: Inductions are generally ampliative¹, in the sense of being non-deductive inferences; their conclusions have logical content that extends beyond (“amplifies”) their premises². Inductions are also extrapolative inferences: their premises are about known instances of expressions (including simple predications, but also more complex expressions) and their conclusions are about unknown instances. Not all ampliative inferences are inductive: the argument ‘Most numbers satisfy Goldbach’s Conjecture, 978 is a number, therefore 978 satisfies Goldbach’s Conjecture,’³ is ampliative but not inductive, because it does not extrapolate beyond the subjects of the premises⁴.

¹ My use of “ampliative” in this article differs from Norton’s. I thank an anonymous referee for pointing this out.

² The exceptions are demonstrative inductions, which form deductively valid arguments when all our relevant assumptions are made explicit: see Norton, 2003b, pp. 665-666).

³ Throughout, I use “double inverted commas” for quotation and ‘single inverted commas’ for reference to words and sentences.

⁴ Some philosophers, e.g. Rudolf Carnap (1962, p. 580) use ‘induction’ more broadly and include all ampliative inference

Implicit Premises: In good inductions⁵, the evidential support that a statement E provides for a hypothesis H is not due to the formal relations of E and H, but rather a fact that can be adopted as an implicit premise R in the inference that connects E and H. In accordance with **Ampliative**, R generally does not create a deductive argument, in contrast to what happens if we add ‘If P, then Q’ to ‘P, therefore Q’. Instead, when we articulate R, we can see why E provides a reason to believe H more strongly (Norton, 2014, pp. 673-674). Nothing in the MTI explicitly confines the nature of R, but in Norton’s examples R always asserts a uniformity (perhaps in a hypothetical class) such that, if R is true, samples of the type described in E will generally be representative (with respect to the induction’s target characteristic) of the induction’s target population. The target characteristic could be many things: a melting point, reacting in a particular way to the introduction of a magnetic force, possessing a colour or smell etc. I shall use the term ‘uniformity principle’ for R.

Local: Unlike many theories of induction along these lines (Hume, 1894, p. 37; Russell, 1948, Chapter 9), Norton denies that there is any single uniformity principle or small set of uniformity principles in science. Instead, many (*very* many) local uniformity principles that operate as “licences” of our good inductions⁶.

Licensing: This evidential relationship consists in providing a defeasible (and undefeated) reason to believe that our observed sample is representative, within a reasonably narrow margin of error, of an inductive inference’s target population⁷. Norton gives the following example: we are justified in believing that all samples of bismuth are representative of the population of bismuth samples with respect to their melting points, because we know that bismuth is a chemical element and that most chemical elements are uniform with respect to their melting points. Therefore, we can extrapolate – *modulo* measurement error – from the melting point of one sample of bismuth to the melting point of all samples. Contrariwise, we could not make the same inference if we knew the (false) defeater that bismuth is an allotropic element and that allotropic elements like carbon and sulphur more or less always have multiple melting points among their allotropic states (Norton, 2003b, pp. 650-651)⁸. In short, in Norton’s examples of licensing, there is a background inference about the representativeness of a sample from a local uniformity principle. I shall return to the issue of licensing on several occasions, but the main point to note is that Norton has not (in print) stipulated a randomness requirement: while information about biased sampling is an obvious type of defeater for inductions, Norton never explicitly requires that we know that all such defeaters are false.

⁵ In the sense in which we know that the inductions are good, rather than those that merely happen to successfully attain true premises.

⁶ Richard Whately (1855, pp. 256-260) held a similar view, but also believed that the role of the local uniformity principles was always to reformulate the inference as a deductively valid syllogism, and thus disagreed with **Ampliative**.

⁷ I focus on the type of licensing relations that are relevant for my discussion in this article. An anonymous referee points out that there may be other types of licensing relation, e.g. for inferences to the best explanation, which might not be reducible to this sort of licensing.

⁸ It is consistent with the definition of ‘allotrope’ that such an element could have the same melting point for all its states, but it would be an extremely astonishing coincidence, and not one that is realised in any known element.

Plurality: The numerous local uniformity principles connecting evidence with hypotheses do so in multifarious ways. Consequently, there is no general and abstract account to be given of their functions. A very normatively weak account, such as Subjective Bayesianism, might be able to accommodate much of good inductive reasoning, but according to Norton even these accounts are unable to capture the full richness of induction (Norton, 2003b, p. 662).

One strength of the MTI is that it seems compatible with both internalist and externalist theories of justification (Norton, forthcoming, reply to Job de Grefte). For instance, even though Norton’s use of terms like ‘fact’ when describing his theory can suggest externalism, he does not depend on externalist premises when they could provide him with a quick escape from a problem (Norton, 2014, pp. 682-683). In general, Norton tends to avoid reliance on particular epistemological views, like foundationalism or anti-foundationalism (Norton, 2014, p. 687). This flexibility means that the MTI can appeal to a broad variety of philosophers. In considering how the MTI can respond to Dawid’s criticism, I shall try to preserve this breadth as much as possible.

SECTION 2 – The Dome

2.1 – Norton’s Model

Dawid criticises the MTI using Norton’s own “Dome” model (Norton, 2007, pp. 166-167; Norton, 2008.) Imagine a perfectly smooth and symmetrically shaped dome, whose form is given by the equation $h = (2/3g)r^{3/2}$, where h is the distance of a straight line through from the apex point of the Dome’s surface to its bottom, r is the surface radial distance coordinate⁹, and g is the acceleration due to gravity for a free unit mass on the surface.

In the Dome experiment, a point mass is placed atop the Dome at the surface apex point. Norton develops a Newtonian model for the Dome set-up and proves that, given the experiment’s conditions, the model has multiple solutions for its equations regarding the point mass. It might instantly slide down the Dome on one path, but it might take any other direction along the surface of the Dome; it might stay still and suddenly (without any additional force acting upon it, other than the force which placed it on the Dome) follow a path after any of an infinitely extendable number of periods of time; or it might stay atop the Dome indefinitely. Therefore, the model underdetermines the point mass’s motion. The model also does not specify a probability distribution for different directions, which might warrant regarding one direction as most likely (Norton, 2008, pp. 787-788). Nor does the model identify a non-trivial disjunction of possible probability distributions for this event. The Dome model only states that the point mass will follow one of the directions (or remain stationary) described earlier, and rules out the alternatives. To put it mildly, the model is uninformative about the Dome experiment. For the Dome’s historical background, see Van Strien (2014).

⁹ The Dome’s surface radial distance coordinate is the length of a straight line from its “top” (its surface apex) to another straight line running from the centre of its flat surface to its perimeter (its radius).

While Newtonianism is false, the Dome model seems logically consistent and conceivable¹⁰. We cannot test it directly, since there are no perfectly smooth and symmetrically shaped domes, nor can we place point masses atop them. Yet we can nonetheless use the Dome to explore and criticise particular the implications of theories of induction for circumstances where our background theories are similarly uninformative.

2.2 – Dawid and the Dome

Dawid does not deny that *some* (even many) inductions are characterised as Norton suggests. However, he criticises the MTI thesis that *all* inductive inferences depend on local factual statements (Dawid, 2015, pp. 1103-1104). He asks: what if we were able to experiment with the Dome repeatedly and we observed a uniform pattern of behaviour by the point mass? For example, imagine that in 20,000 controlled experimental trials of the Dome experiment we observe that the point mass remains stationary for 16.8 seconds after it was placed on the Dome's apex¹¹ and then slides down the Dome¹². This apparently evinces that, in the next experiment, the point mass will also remain stationary for 16.8 seconds and then slide down the Dome. I shall call this prediction P_1 . I interpret Dawid's argument as follows:

(1) To be justified in inferring P_1 we must have inductively inferred a pattern for the time of excitation and direction of motion of the point mass in the Dome experiment¹³.

(2) According to the MTI, all inductions are justified only by local uniformity principles in our relevant background information.

(3) In Dawid's scenario, if there is a licensing local uniformity principle in the inference of the sliding pattern for the sliding behaviour of the point mass on the Dome, then it is supplied by the Dome model.

(4) The Dome model does not predict (deterministically or probabilistically) any particular time of excitation or direction of motion for the point mass, and thus cannot supply suitable local uniformity principles.

Therefore, (5) If the MTI is a correct theory of induction, then we cannot be justified in inferring P_1 in Dawid's scenario.

However, (6) We can be justified in inferring P_1 in Dawid's scenario.

Therefore, (7) The MTI is not a correct theory of induction.

¹⁰ I leave aside questions about whether it is "really" Newtonian, e.g. whether Newton's First Law rules out solutions other than the point mass remaining stationary (but see Norton 2003a, pp. 10–11). I shall assume only that the Dome is consistent and conceivable, else it does not pose a problem for the MTI.

¹¹ Like Dawid and Norton, in this article I shall simplify the discussion by disregarding measurement error.

¹² This is a slight adaptation of Dawid's scenario. I have increased the sample size. (Unlike in real science, replications of imaginary experiments are easy to perform.) My arguments could be adapted to the 100-fold sample that Dawid considers, but the relevant intuitions would perhaps be weaker, in ways that are inessential to the core issues. If his arguments work at all, they work (and better) with very large samples.

¹³ Some philosophers might deny (1) and argue for a more horizontal approach to predictive inferences. However, even if we can infer P_1 by an alternative route, that would still leave the problem that there does not seem to be anything wrong (without begging the question in favour of the MTI) with the reasoning that Dawid suggests.

Dawid's argument is valid. It also *apparently* consist of premises to which Norton is committed. Premise (1) is plausible given the MTI, since Nortonians cannot appeal to devices such as prior probabilities based on subjective opinions or symmetries to derive a high expectation for P_1 . Premise (2) is an integral thesis of the MTI. Premise (3) is plausible because, in the Dome scenario, there are no other accepted background hypotheses mentioned apart from Newtonian mechanics and the other components of the Dome model. Premise (4) is a point that Norton has utilised in his own discussions of the Dome. Premise (5) follows from the preceding four. Premise (6) seems plausible to some philosophers – though, as we shall see below, Norton rejects it under some conditions.

A Nortonian might reject (3) in the following way: perhaps our intuitions in Dawid's scenario are conditioned (in the psychological sense) by the past reliability of induction in our past practice. Either this background knowledge is available in Dawid's scenario, in which case we are able to utilise this general reliability to licence the prediction of P_1 , or this background knowledge is unavailable, in which case the induction is illicit and it is a virtue of the MTI that it forbids the inference. In neither case does the MTI deny the possibility of a legitimate induction.

However, there are at least three problems with this response. Firstly, the notion of a *general* "reliability of induction" is opaque. If we interpret this idea in the formal mode (in terms of the reliability of a formal schema) we face the problem that there are many types of induction: statistical inductions, singular predictions, analogies, inductions involving theoretical phenomena etc. Is the "reliability of induction" the reliability of all of these schemas or a particular type? Furthermore, any given inductive inference will be an instance of indefinitely many inferential schemas. Which schema should be used for assessing a given induction's reliability? Secondly, as Norton argues, meta-inductions about schemas do not actually occur in science (Norton, 2003b, p. 667). Indeed, the methodological importance of the *general* reliability of induction is just what the MTI rejects, because one of the most interesting features of Norton's theory is that inductions are justified by *local* facts. Thirdly, if we take the material mode and interpret the "reliability of induction" as the claim that the universe is clement to induction, then we lose the distinctiveness of the MTI, since this fact about the universe is not a local fact; it would be a return to the sweeping claims of Mill, Russell, and so on. I grant that "the reliability of induction" makes sense: we can say that a *particular* natural phenomenon is more or less uniform, so that inductions about it will tend to be more or less reliable. Inductions tend to be reliable from the properties of one instance of a chemical element to all elements, but unreliable from the psychology of one person to all people. Yet induction's *general* reliability cannot provide a *Nortonian* answer to Dawid's example¹⁴.

One might say that Dawid's criticism concerns a very unusual case. Normally, there is at least one (often many) background hypothesis licencing an intuitively reasonable induction. Perhaps, in a few special cases, we have to appeal to formal schemas or contingent local uniformity claims of a great degree of generality in order to licence intuitively correct inductions. That would allow that Norton's theory is almost always adequate, even if we must

¹⁴ The same applies for a response that Norton suggests in correspondence with Dawid (2015, p. 1106) in terms of an assumption that such regularities are law-governed: this is exactly the sort of vague and sweeping "assumption of induction" that the MTI rightly rejects.

forgo the MTI in this recondite case.

However, Dawid's criticism is robust to a wide range of variations, because it is principally driven by the lack of informative background information, rather than anything particular to the Dome itself. Consider another Newtonian variation: so-called "space invaders" (Earman, 2004 pp. 25-26). Imagine a point mass particle x in a universe of finitely many particles that conform to Newton's inverse square law. It is provable that this particle could, within this Newtonian universe, accelerate (in a finite period) to spatial infinity, but also that that particles might arrive unpredictably *from* spatial infinity. Let us call particles with the latter behaviour "space evaders". While Earman's model allows for space evaders, it does not tell us under what conditions they would occur, nor provide a stochastic model or informatively small family of stochastic models for guiding such predictions. However, if we could easily detect space evaders, but we never did in hundreds of years, we might predict that we are unlikely to encounter them in the near future, on the grounds that they apparently happen to be rare.

More generally, in science, our background information often does not provide us with rich warranting facts. For instance, in economics, the Mundell-Fleming model of a small open economy implies that it is possible for governments of such economies to (i) control their domestic monetary conditions, (ii) control their currency's exchange rate with another currency, or (iii) allow capital to freely enter or exit the country. The model also says that any combination of (i-iii) can be sustainable, *except* all three together (Fleming 1962, Mundell 1963). The model does not predict which combinations will actually occur, nor obtain results for which combinations are "better" by some standard of social welfare, but it is useful because it tells policymakers what circumstances can and cannot be realised in such economies. The analogy with the Dome model is imperfect, because the Mundell-Fleming model is not intended to be a complete description of the target system. Nonetheless, since contexts where our background models only provide us with relatively exiguous background knowledge often occur science, it is important to clarify how the MTI can handle them.

2.3 – Norton's Responses

Norton replies to Dawid's criticism in Chapter 14 of his unpublished manuscript for *The Material Theory of Induction* (Norton, unpublished, pp. 22-23). He distinguishes two possible responses that we might have after the Dome experiments. Firstly, we may become uncertain that the Dome model is complete and correct; I shall call this approach the "theory change" response. Secondly, we remain certain that it is complete; I shall call this approach the "resolute" response.

If we adopt the theory change response, then we shall expand the Dome model or develop an alternative model. Given either expansion or the development of an alternative, there will be new theoretical content in our science, and the relevant parts of our new physics will be the background information that licences the inference. In this case, we reject premise (3), i.e. that only the Dome can provide the licensing principles for the induction he describes. I agree that we might expand the Dome model. (We might want to do so anyway, given that the model is so uninformative.) However, the evidence does *not* tell against the background physics, because the Dome model is both deductively and probabilistically in accordance with the sample data. A heterogeneous pattern in the experiments would be no

less consistent (in a broad epistemic sense) with the Dome model. Obviously, the model does not predict the sample data, but nor is it disconfirmed¹⁵. It is hard to see how a hypothesis can be disconfirmed by some data without the hypothesis supporting that data's negation. If the Dome model provided reasons to expect some *other* pattern or set of patterns was likely to occur, then the observed pattern would be evidence against the Dome model. However, the model does not provide such reasons in this context, because the model does not say that the regular behaviour that Dawid imagines is any less likely than any other behaviour. In general, provided that the observed behaviours of the point mass are among those that the model deems "possible", they are evidentially irrelevant to the model.

One might argue that Newtonianism contains an implicit completeness claim, such that any physical uniform pattern that occurs can be predicted by combining statements of the initial conditions with Newtonian theory. However, adding this completeness claim (or making it explicit) entails that Newtonianism, as Norton and Dawid render it, predicts physical disuniformities where it does not predict uniformities. The Dome model would predict that, in tens of thousands of iterations of the Dome experiment, we would *not* see a uniform pattern. Yet one point of agreement for Norton and Dawid is that the Dome model makes no such prediction. Norton is correct that good scientists would seek a more informative theory given the evidence that Dawid describes, but there is nothing in the data that would require this search. The scientists' motivation for their search would be present from the start: the pursuit of models with more explanatory and predictive power. (Their interest in a better model would have increased.) The experimental data would discipline the choice of alternatives, but not disconfirm the model. Therefore, in this part of his response, Norton has not incorporated the reasoning that Dawid describes into the MTI.

If we adopt the resolute response to the pattern in our observations, then we remain certain of the Dome model's completeness. In this case, Norton claims that the sample data provides no reason to expect that P_1 will occur, because the background physics provides no basis for predictions *even if* we have such uniform sample data. Thus, given the resolute response, Norton rejects reject premise (6) – that the induction is actually justified. He does not directly argue for this claim, but he provides an analogy to a gambler at a roulette wheel. Suppose that the gambler does not doubt their physical model (perhaps a folk-physical model) of the roulette wheel and their model provides them with the background knowledge that each spin of the roulette wheel is independent. The probability¹⁶ of landing 'black' is 0.5. Due to the known stochastic independence of the roulette spins, even 20,000 spins that all land on 'black' would not justify adopting the expectation that the 20,001st spin will land on 'black', unless the gambler doubted their physical model of the roulette wheel (Norton, unpublished, Chapter 15 "Indeterministic Physical Systems" p. 22)¹⁷.

¹⁵ The reader might be tempted to apply the Principle of Indifference in order to infer that the sample data is unlikely given the Dome model: if every possible time of motion is equally probable, then 16.8 seconds is maximally improbable, since there are an infinite number of possibilities, but such *a priori* symmetry principles are inconsistent with the MTI.

¹⁶ Assuming that the green pockets, 0 and (in American roulette) 00 have been removed as options.

¹⁷ We are also ignoring the reliability of the description of the initial conditions. Questioning these merely relocates the problem, since we would only question them if we have inductively learned something from the 20,000 trials.

However, this analogy does not work. In the roulette example, the gambler has highly informative background knowledge of relevant physical probabilities. The implicit model for the roulette wheel tells us that the physical probability of landing in the ‘black’ pocket is independent of its past behaviour¹⁸. In contrast, the Dome model does not provide a conditional probability distribution for events like P_1 . It just tells us that P_1 *may* occur in the 20,001st Dome experiment.

The Dome model is consistent with the inductive inference of a finite relative frequency for point mass directions, using the sample data, and this relative frequency could (defeasibly) provide warrant for expecting P_1 ¹⁹. In what follows, I shall reject premise (3) – that the Dome model offers the only candidate for licencing local uniformity principles. Unlike Norton’s response, I shall argue that this rejection is possible even if we do not adopt some non-Newtonian physics in response to the evidence that Dawid imagines. (I agree with Norton about the case where we adopt a new model or perhaps even a new physics in reaction to the experiment.) Thus, I shall argue that Dawid’s argument is valid, but unsound, and therefore the Dome does not defeat the MTI.

SECTION 3 – Local Combinatorial Induction

3.1 The Structure of the Problem

Firstly, note that Dawid’s challenge does not concern the justification of induction in general within the MTI, but only a particular induction in a scientific context with at least some background knowledge. Here, Isaac Levi’s separation of “global induction” and “local induction” (Levi, 1967, pp. 3-6) is useful. Global induction is the practice of induction *as such*. A justification of global induction would be what David Hume requested: by what reasoning can we justify *any* use of induction? In contrast, local inductions are inferences in particular contexts of scientific inquiry. Dawid’s challenge is a problem of local induction, and therefore my argument that the MTI allows for the reasoning he discusses will be independent of whether the MTI contains or permits a positive answer to Hume’s problem.

Secondly, the inductive reasoning that Dawid describes might be outlined in a number of ways. However, the following schematic reconstruction is simple and amenable to a variety of more detailed analyses:

(E) All of a 20,000-fold sample of trials of the Dome experiment resulted in the point mass sliding down the Dome after 16.8 seconds.

¹⁸ Either for a generic spin (frequentism) or the particular spin (the propensity view).

¹⁹ Some might argue that finite frequencies are insufficient for making scientific predictions and that predictions require a causal connection or knowledge of limiting relative frequencies in a (normally hypothetical) infinite series of trials. However, the MTI does not require that local uniformity principles describe an infinite series, and Norton takes a deflationary view of causation (Norton, 2003a) and my concern is what the MTI allows.

Therefore, defeasibly²⁰:

(H) All, or at least almost all, of the trials of the Dome experiment result in the point mass sliding down the Dome after 16.8 seconds.

Therefore, defeasibly:

(P₁) The 20,001st trial will result in the point mass sliding down the Dome after 16.8 seconds.

The inference from H to P₁ is ampliative and thus non-monotonic. If an inference is non-monotonic, then it is defeasible: it is conceivable that we might accept further propositions that are consistent with the premises, but which rule out the inference as rational for us. Nonetheless, there seem to be no defeaters in Dawid's scenario, partly because the most pertinent background information (the Dome model) is extremely uninformative about the point mass's behaviour. In general, Nortonians can be sanguine about this part of the reasoning, because H can operate as local uniformity principle in the way that Norton has described in other examples. In more detail, the inference from H to P₁ can be seen as a simple inference:

(H) All, or at least almost all, of the trials of the Dome experiment result in the point mass sliding down the Dome after 16.8 seconds.

(A) The 20,001st trial will be one of the trials of the Dome experiment.

Therefore, defeasibly, (3) P₁ is true.

Nothing in the MTI rules out this step. Therefore, the difficulty in Dawid's example is just the inference from E to H. In the MTI, this inference cannot be justified via some formal relationship between these statements. A local uniformity principle is needed. It must justify the belief that the sample described in E is representative of the population described in H, i.e. the induction's target population.

3.2 A Local Justification

To justify this particular induction, I shall adapt a justification of global induction, developed by philosophers such as Josiah Royce (1913, pp. 82-88) Donald C. Williams (1947) and David Stove (1986, Chapter VI). I shall call this the 'Combinatorial Justification of Induction' (CJI)²¹. However, unlike the standard presentation of the CJI, I shall not suppose that the scientists involved only possess a very austere quantity of background knowledge, because Dawid's problem is not a global problem of induction where (arguably) such evidential exiguity is appropriate. In particular, I shall suppose that they know that the categories used to describe the phenomena (seconds, point masses, the Dome itself, and so on) are *bona fide* scientific categories that can be rationally used in inductive inferences. The

²⁰ The qualifier 'defeasibly' indicates that I am using 'therefore' in the way that Norton uses the word 'likely' in "the rough and ready sense of ordinary judgement" (Norton, 2011a, p. 10). Note that the conclusions are categorical statements, rather than probability statements.

²¹ The name is from Marc Lange (2011, p. 83).

future existence of the 20,001st trial is also not in question. Thus, what I say will be logically independent of the CJI's viability as a justification of global induction.

The type of inference from H to P₁ has many names: “the statistical syllogism”, “the proportional syllogism”, and more. However, the name that seems least likely to mislead is “direct inference” (Carnap, 1962, p. 207), since it is a type of inference, whereas “syllogism” tends to suggest a form of deductive reasoning. Direct inference is an ampliative inference from a premise about a population to a conclusion about some subset of that population. It has a crucial function in the MTI, because we must infer (defeasibly) from local uniformity principles about general populations in order to know that the samples in our inductions are representative of our target populations. (See the bismuth example given earlier.) Thus, direct inference is *defeasibly* permissible according to the MTI.

For the inference from E to H, the target population is the population of Dome experiment trials. One available fact about this population in Dawid's scenario is that it has a finite but unknown cardinality of *at least* 20,001. We know it by combining E with our background knowledge that the 20,001st trial will occur and our background scientific knowledge that the Dome experiment will only occur in finitely many trials. This last claim is realistic, but not essential, as we could limit the scope of H to a finite set by definition and it would still be just as useful for prediction²². Furthermore, we know some mathematical facts about such populations from combinatorics. As I shall explain, we know from combinatorics that, for a property like ‘slides down the Dome on a particular path after 16.8 seconds’, the overwhelming majority (some proportion that *must* be well over 99%) of 20,000-fold subsets of such a *finite* population will have a proportion that is representative of the population proportion within a margin of error of 1%.

To see why, let us allow for even the least favourable population frequency p for representative (in this sense) subsets, which is $p = 0.5$, i.e. 50%²³. Given this population frequency, the most common subset among the m -fold subsets will be mp (or round to mp as the nearest integer). Let $q = (1 - p)$ and $h = mq - n$, where n is the number of instances of the property in m . Meanwhile, π and e will be the familiar mathematical constants. Then, for large subsets, the proportion of m -fold subsets with n will be approximately²⁴:

$$\frac{1}{\sqrt{2\pi mpq}} e^{-\frac{h^2}{2mpq}}$$

– which reaches a maximum if $h = 0$. Then, $0 = mq - n$. Therefore, the maximum occurs when $n = mq$. Given that maximum, but recognising that we are only interested (in this case)

²² Assuming that the set is picked out by some suitable scientific description, rather than a set picked out by ‘grue’ or similar descriptions.

²³ If the population frequency of a property is 0 or 1, then obviously all subsets will be representative. The proportion of representative subsets falls as the proportion is further from these extremes, and hence reaches a minimum if it is 0.5.

²⁴ Using Stirling's approximation (Slomson 1997, Chapter 4) though it is possible to prove the relevant mathematical facts without this particular principle (Kneale 1949, pp. 136-142). The combinatoric approximations I describe are not the closest possible, but they are relatively easy to understand.

in approximate estimates of the population frequency, how can we determine the frequency with which a subset will be within some margin of variance? We replace h by a continuous variable z . The proportion of subsets that diverge from the maximally common proportion mq by z and $z + dz$ (i.e. between z and z plus an infinitesimal of z) is:

$$\frac{1}{\sqrt{2\pi mpq}} e^{-\frac{z^2}{2mpq}} dz$$

Finally, to obtain the proportion of subsets whose value of n lies within the range $mp \pm a$, where $t = \frac{z}{\sqrt{2mpq}}$, we can use integration:

$$\frac{1}{\sqrt{2\pi mpq}} \int_{-a}^a e^{-\frac{z^2}{2mpq}} dz = \frac{2}{\sqrt{\pi}} \int_{-0}^{\frac{a}{\sqrt{2mpq}}} e^{-t^2} dt$$

If $\gamma = \frac{a}{\sqrt{2mpq}}$ and $\varepsilon = \sqrt{\frac{2pq}{m}} \gamma$, then the proportion of m -fold subsets that are representative of the population within a margin of error of $\pm \varepsilon$ (i.e. the range of n values around mq) will be:

$$\frac{2}{\sqrt{\pi}} \int_{-0}^{\gamma} e^{-t^2} dt$$

This proportion increases proportionately to the subset size m . The particular implication that advocates of the CJI have exploited is that, if m is large, then the proportion of *representative* subsets of that type will be large²⁵. An intuitive gloss is that while the absolute number of unrepresentative subsets can be high even if the subsets are large, the proportion of such subsets is bounded by combinatoric principles that are independent of either the ratio of the subset to the population or the population size (generally unknown in scientific inductions) provided that the latter is finite. Thus, even if the population is indefinitely large (but finite) most large subsets that could be formed out of its members will be representative, where “representative” becomes definable with a tighter margin of error as “large” becomes more demanding. Provided that we are talking about subsets, what I have said so far is uncontroversial even among critics of CJIs (e.g. Maher 1996; Lange 2011).

Note that I have discussed *subsets*, rather than *samples*. That a sample is a subset is

²⁵ Unfortunately, advocates of CJIs have rarely explained the role of combinatorics in much detail. My exposition draws on John Maynard Keynes’s proofs (Keynes 1921, pp. 383-385) except that he is principally interested in the relation of the combinatorics to probabilities. The approach that Keynes uses, via Stirling’s approximation, comes from Joseph Bertrand (1889, Chapter IV). The original theorems are due to Jakob Bernoulli, whose own proof differed from the Bertrand approach (1713, p. 236). For explanations aimed at philosophers, see (Williams 1947, Chapter 4) and (Stove 1986, Chapter VI). Due to the idiosyncrasies of their theories of epistemic probability, Williams and Stove think that they are able to infer epistemic probabilities from these combinatoric propositions. Since Norton rejects the idea that all direct inferences are probabilistic (Norton, 2014, p. 674) I have avoided talking about epistemic probabilities. This is consistent with my use of direct inference, which need not be probabilistic (Toulmin 1958, p. 109).

not necessarily the most pertinent fact about it with respect to its representativeness, as I describe below in my discussion of sampling bias. For instance, (a) the minima of the *population* frequency of representative subsets and (b) the frequency of obtaining representative samples with some selection procedure (either in practice or in the limit) can diverge. In other words, the frequency with which we draw representative samples can differ from the minimum proportion of representative subsets. To take an extreme case, if our sampling method is biased, then we might never obtain a representative sample, even though the overwhelming majority of large subsets of a finite population are representative.

Nonetheless, based on Norton's bismuth example and other instances of reasoning in the MTI, I take it that, according to his theory, knowledge of a high minimum population frequency of representative samples is a *defeasible* sufficient condition for direct inference. In other words, like advocates of the CJI, Norton believes:

Proportion Direct Inference: That a proportion $r\%$ of a population K has the property F (for instance, being a representative subset) and that an individual i is a member of K jointly constitute a defeasible reason to believe that i is a K , if $r\%$ is high. If $r\%$ is low, then this fact is a defeasible reason to believe that i is not a K .

What constitutes a defeater, i.e. something such that if we believe it, we would no longer have evidence? I shall return to this issue later (especially in the conclusion) as it turns out to be an important lacuna in Norton's theory, but I can give some illustrative intuitive examples. Firstly, if we know that i is a member of a proper subset of K such that this subset's proportion $s\%$ differs from $r\%$, then intuitively we should reason using $s\%$, which could be low/high when $r\%$ is high/low. Secondly, we might know that, for the sampling procedure we used, the relative frequencies for representative subsets differ from the population proportions, i.e. that i was obtained from K by a biased sampling method.

One might think that it is a necessary condition of direct inference that we know that the sampling method is unbiased²⁶. However, this does not seem to be Norton's view, because he does not impose any such condition in his bismuth example or elsewhere. Since my defence of the MTI is internal (my aim is that certain inductions are reasonable given the MTI, not that the latter is reasonable) I shall not discuss whether Norton is correct in this regard. However, perhaps some of unease about Dawid's reasoning, where biased sampling is not excluded, might be a result of recognising that such a defeater is quite plausible, e.g. a nefarious laboratory assistant seeking some dramatic result by only recording Dome experiments that conform to the regularity in question.

' X is representative of Y ' is a symmetric relation. Consequently, if least 99% of the

²⁶ If the MTI *should* require this condition, then my reasoning could be interpreted as providing a Nortonian explanation of our intuitions in the Dome induction: Dawid does not specify anything about the sampling procedure; we tacitly make an assumption of an unbiased sampling procedure, and ignore that this assumption is unwarranted in the hyper-austere background knowledge that Dawid does specify. Consequently, premise (6) in my set-up of Dawid's argument in Section 2.2 is false, and it only appears to be true because we are used to making that assumption as a default in enumerative induction problems. Therefore, regardless of whether the MTI should require antecedent knowledge that the sample selection is unbiased, my arguments offer something to the MTI Norton's theory in responding to Dawid's criticism. See (Campbell and Franklin, 2004) for a criticism of the randomness requirement.

subsets are representative of the target population, then the target population is representative of at least 99% of its subsets. CJI advocates have the high ambition of answering Hume, but I only aim to justify the inference from E to H. I begin with a direct inference:

(B₁) Over 99% of the 20,000-fold subsets of the target population (the population of trials of the Dome experiment) will be representative²⁷. To be precise, the relative frequency of sliding down the Dome after 16.8 seconds in 20,000-fold subsets of the population of trials will be representative of the population frequency within a margin of error of $\pm 1\%$.

(B₂) The sample reported in E is a 20,000-fold subset of our target population in which 100% of the point masses slid down the Dome after 16.8 seconds.

Therefore, defeasibly, (B₃) The sample reported in E is representative of our target population within a margin of error of $\pm 1\%$. In 100% of this sample, the point masses slid down the Dome after 16.8 seconds.

From B₃, it deductively follows that $100\% \pm 1\%$ of the target population will be trials in which the point mass slides down the Dome after 16.8 seconds. And, if we interpret ‘all or almost all’ as indicating a standard that is no stricter than ‘at least 99%’, then H follows. Hence, B₁ is a local uniformity principle that warrants the inductive step from E to H. As with other paradigmatic inferences in the MTI, the induction to H is licenced by a local uniformity principle that gives us a defeasible reason to expect the sample to be representative of our target population. From H, the prediction that Dawid proposes, P₁, is also justified by direct inference²⁸. Consequently, the inductive reasoning that Dawid suggests is consistent with the MTI.

We can use B₁ when we know (perhaps only roughly and intuitively) the relevant basic combinatoric principles and *if* there are no defeaters for the direct inference in argument B, nor any defeaters for the direct inference from H to P₁. For example, we must not know that our observations are a member of a subset of Dome experiments that are unrepresentative or likely to be unrepresentative. Additionally, background knowledge about bias in the sampling procedure can be a defeater. However, in the scenario as Dawid and Norton discuss it, we lack such defeaters. In the MTI, direct inferences from proportions have at least a presumption in their favour, as in the bismuth case: we might know a defeater for the direct inference that bismuth has a uniform melting point from the hypothesis that most elements do (for example, if we knew that bismuth was an allotropic element) but without such a defeater, Norton allows the inference.

²⁷ This can be obtained directly using the combinatoric principles mentioned earlier, emphasising again that they require that the subsets in question are large. In particular, since that reasoning uses the normal distribution as an approximation, it is easy to find tables and calculators for the proportions in question. Meanwhile, when the subsets are small, the proportions can be calculated, with more computational difficulty, directly using a binomial distribution.

²⁸ We should clarify terms in the MTI like ‘licence’, ‘warrant’, ‘defeaters’, and ‘reason to believe’ in refined detail. I return to this point in the conclusion.

SECTION 4 - Objections

There are many possible objections to the reasoning that I offered in the previous subsection. Many of them utilise premises that are independent of the MTI or even in outright contradiction to it. However, the issue in question is whether the reasoning that I have offered is available in the MTI. That can be done without proving that justification should be acceptable to *everyone*. Consequently, I shall only discuss criticisms that apply if we take the MTI for granted and do not require particularly controversial additional premises.

4.1 B₁ is not the Right Sort of Fact

My argument uses a parallel between the reasoning in the CJI and the direct inferences that are typical within the MTI. Yet B₁ might seem to be different from the local uniformity principles typically used by Norton to exemplify his theory. The aforementioned facts about chemical elements' uniform melting points is a contingent fact that chemists discovered empirically. In contrast, B₁ might seem to be necessarily true and acquired by *a priori* reasoning.

These appearances are misleading. In Dawid's scenario, B₁ is true, but it is not true in the real world: the Dome experiment can never actually take place, and thus the proportion of representative trials of the Dome experiment among the total number of trials would be zero divided by zero. B₁ is false for the same reason that 'At least 99% of the 20,000-fold subsets of manticores are representative' is false. No such proportion exists. Thus, B₁ can only be contingently true or false.

We know the local uniformity principle B₁ in Dawid's scenario via *a posteriori* reasoning. In detail, we can infer it from the following argument:

(C₁) If a population is finite and has at least one 20,000-fold subset, then over 99% of the 20,000-fold subsets of that population will be representative.

(C₂) The population of trials of the Dome experiment is finite.

(C₃) There exists at least one 20,000-fold subset of the population of trials of the Dome experiment.

Therefore, (C₄) The population of trials of the Dome experiment is finite and has at least one 20,000-fold subset. (From C₂ and C₃.)

Therefore, (B₁) Over 99% of the 20,000-fold subsets of the target population (the population of trials of the Dome experiment) will be representative. (From C₁ and C₄.)

We know (C₂) and (C₃) empirically: (C₂) via our background knowledge about the Dome experiment²⁹, (C₃) via our observation of a 20,000-fold sample. Hence, B₁ is known *a posteriori*. The appearance to the contrary is perhaps because B₁ is a deductive consequence of the evidence, obvious background knowledge, and purely mathematical principles of combinatorics. This modest evidential basis of B₁ from our experience is atypical of scientific

²⁹ Assuming that "the Dome" refers to a particular object in the imaginary scenario, rather than Dome objects in general, and that we know that this object will not exist into the infinite future.

reasoning in general. When Norton uses the following reasoning:

(D₁) Chemical elements are generally uniform in their melting points.

(D₂) Bismuth is a chemical element.

Therefore, defeasibly, (D₃) The melting points of samples of bismuth are representative of the melting points of bismuth in general.

- D₁ is only known due to the empirical achievements of generations of chemists. It was acquired with much more sweat than B₁. (Presumably: remember that the Dome is a thought experiment.) Nonetheless, while the reasoning in Section 3.2 differs in some ways from the typical reasoning in the MTI, there is no significant epistemic difference between B₁ and facts like D₁. Both of these uniformity principles can only be known *a posteriori*.

A Nortonian might also worry about the generality of B₁. The use of the purely formal principles of combinatorics within the inference of B₁ might suggest that we can put this reasoning into a universal schema. Yet there are at least two reasons why the inferences in Section 3.2 why this schematisation is impossible. Firstly, the requisite local uniformity principles such as B₁ will not always be epistemically available: not all populations *have* 20,000-fold subsets, so such claims about their proportions will be false. A formalist might try to abstract away this issue by taking the subset size to be a variable, but this will only raise more problems, because if our sample size is 1, then we cannot infer that at least 99% of such subsets will be representative of their populations. One might say that the subset size could be a partly bounded variable (e.g. restricted to whole numbers greater than 25). However, the not all populations will have subsets meeting the minimum bounds, so once again the reasoning is only possible if our background knowledge of the material facts is suitable.

Secondly, my reasoning assumes that sliding down the Dome after 16.8 seconds is a projectible quality. In this case, my reasoning diverges from the CJI. For example, David Armstrong (1983, pp. 57-58) points out that, in Williams' version, the CJI reasoning apparently applies for predicates like 'grue'³⁰. (Stove's version of the CJI avoids this consequence by an unexplained restriction against 'grue' and similar predicates.) I have avoided these issues only by assuming we have antecedent scientific grounds to discriminate among predicates, because Dawid is not raising the Humean problem. Thus, in conformity with the MTI, local background knowledge – i.e. beliefs that are sensitive to the particular details of the case – limit the applicability of the reasoning in Section 3.2. However, as an anonymous referee points out, what counts as a "local fact" (or a "formalist" theory of induction) is open to dispute. If a "formalist" is willing to say that inductions always depend on the sort of contingent background knowledge that I have described, then I have no issue with formalism in that sense, nor do I see why Norton needs to object to such "formalists".

³⁰ Norton has already discussed 'grue' and the associated New Riddle of Induction from an MTI perspective. (Norton, 2006).

4.2 Universal Generalisations

Even supposing Dawid accepted all my arguments thus far, there are some simple ways that he could alter his criticism that would raise analogous new problems. I have argued that Nortonians can justify the inductive inference of P_1 by adapting the CJI to the local induction that Dawid describes. Yet the reasoning in the CJI only directly enables us to infer approximate conclusions, not universal generalisations. We can infer H: ‘All or almost all trials of the Dome experiment result in the point mass sliding down the Dome after 16.8 seconds’, but the combinatoric facts involved in my justification of this induction involve an unavoidable margin of error, and therefore we cannot use the same reasoning for hypotheses like the universal generalisation U: ‘All of the trials of the Dome experiment result in the point mass sliding down the Dome after 16.8 seconds’. Dawid could argue that we should be able to say that U is evidentially supported in his scenario, but can the MTI accommodate this intuition?

It can. If H is true and we do not believe that there are any counterexamples to U, then we know that U is either true or at least approximately true, given our total evidence³¹. Furthermore, if learning some evidence makes it more likely that a hypothesis is true or at least approximately true, then it seems that the evidence has confirmed the hypothesis. (This is somewhat broader than the standard Bayesian definition of confirmation, which requires increasing the probability that the hypothesis is *true*, not just true or approximately true.) Therefore, there is at least one plausible sense of evidential support on which E provides evidence for U, as well as for P_1 . The further question of whether U is the *best* supported hypothesis is beyond my scope; it requires developing a theory of comparative evidential support within the MTI, which Norton has not yet done. That would extend my inquiry far beyond the Dome.

4.3 Alternative Sampling Distributions

One feature of the CJI is that, in standard Bayesianism, the CJI requires additional premises about the relevant prior probabilities. In particular, an orthodox Bayesian must assign some prior to the different possible distributions of sample selections, which can undermine CJI-style reasoning (Maher, 2001; Lange, 2011, pp. 83–86). For example, it is *a priori* possible that the 20,000 observed trials of the Dome experiment were a highly unrepresentative subset of the population of Dome trials. If a Bayesian assumes a uniform probability distribution of any subset’s selection, then the CJI reasoning works: it is still possible that the sample is unrepresentative, but unlikely, because almost all such samples are representative and our sample is no less likely to be unrepresentative. One might think that a similar limitation exists for my use of combinatoric reasoning in the MTI. Why should we prefer the predictions suggested by direct inference over those suggested by other *a priori* possible sampling distributions? Such a preference might be rationally permissible, but it is mandatory?

Although a similar issue might be raised for e.g. Norton’s use of direct inference in

³¹ By ‘approximately true’, I mean the relationship between a claim like ‘The relative frequency of F in G is r ’ and a claim like ‘The relative frequency of F in G is $r \pm \varepsilon\%$ ’, not Popperian verisimilitude etc.

his bismuth example, he has not yet explicitly discussed this worry. Here is a response that is similar to Henry E. Kyburg's treatment of sampling distributions in his theory of direct inference (Kyburg, 2006). It is true that *knowledge* of more informative sampling distributions can undermine those recommended by direct inference from more general combinatoric principles. For instance, imagine that you are sampling 20,000 grains of sand from a beach to assess the proportion of silicon in the sand. If you are sampling the grains using a filter (against other objects on the beach) that with a known statistical bias in favour of low-silicon sand, then you ought to favour reasoning that incorporates your knowledge of that bias over the purely combinatoric reasoning about 20,000-fold subsets. In any such case, where there is a conflict between direct inference from general combinatoric principles and direct inference *justified* beliefs about sampling distributions, then the latter should take precedence, because they incorporate more of our relevant evidence.

However, in the Dome experiment, as it has been discussed thus far, we have no such conflict, because we do not have justified beliefs about the sampling distributions. Our only useful information for direct inference is that our observed sample is a 20,000-fold subset of the Dome trials in general, and that almost all such subsets will be representative. By contrast, while there are logically possible sampling distributions where our sample is probably (or even certainly) unrepresentative, we do not know if they are true, and thus they should not factor into our reasoning. The same is true, *mutatis mutandis*, for our reasoning from our inductively inferred generalisation about Dome trials in general to the particular prediction about the 20,001st trial. In brief, the CJI-style reasoning I have described should have precedence, because we *know* that our observed sample is a subset of Dome trials in general, whereas direct inference from conflicting sampling distributions is merely speculative.

This point reveals an important difference between the MTI and Subjective Bayesianism. In the latter, we are very free to speculate. In the former, it is only antecedently justified beliefs about physical probabilities (or local facts in general) that can factor into our reasoning. Subjective Bayesians can use something like the reasoning I have described in this article to prove that Dawid's induction is permissible according to their theory, but MTI supporters can go one step further and argue that it is mandatory. I regard this as a positive consequence of Norton's more general rejection of *a priori* science (2003a, p. 3).

CONCLUSION

A theory of induction should be able to handle both questions of global induction and local induction. Dawid's adaptation of the Dome scenario into a problem of local induction is a very clever objection, because it challenges the capacity of the MTI to handle such problems. Norton, in his responses to Dawid, has sometimes drifted from the spirit of the MTI, which suggests that it is a tough objection to address. I have argued that the background knowledge of local uniformity principles in the Dome scenario is richer than either Dawid or Norton have appreciated. Since my answer does not particularly depend on the details of the Dome model itself, the MTI emerges as a very robust theory of induction: if it can be supplemented with an account of defeaters for direct inference, then the MTI can handle both

ordinary scientific situations where licencing local uniformity principles are plentiful and the (rarer) situations where we have only extremely austere relevant background information.

Do my arguments provide a general formula for Nortonians to handle all such cases of exiguous background information? Perhaps even an answer Hume's problem of global induction? No. I have liberally drawn on background information that is normally regarded as unavailable in Hume's problem of induction, such as the implicit assumption that directions of sliding down the Dome are not 'gruesome' properties, and can be projected in inductive inferences.

Finally, my arguments have employed direct inference and the notion of defeaters, but the details of their incorporation into the MTI is an outstanding issue for Norton's theory (Peden, 2019). In the absence of clarifications of notions like 'defeaters' and the exact details of the support provided by direct inference in the MTI, one might reasonably worry that my reasoning has been too vague. I agree that Norton's theory is underdeveloped in this respect. However, it is not the Dome in particular, nor even analogous cases of weak background theory, which create that problem: it is a general area for expansion in the MTI. If this gap could be filled, then Norton's theory would provide a default presumption in favour of inductions with large sample and appropriate reference classes, which means that it would hard to argue that it is too narrow. Humeans, Popperians, and others might think that it is too broad, but that is another story. If the MTI was developed further in this respect, then it would be a tough theory to defeat.

BIBLIOGRAPHY

- Bernoulli, J. (1713). *Ars conjectandi*, Basel: Thurneysen Brothers.
- Bertrand, J. (1889). *Calcul des Probabilités*. Paris: Gauthier-Villars.
- Campbell, S. and Franklin, J. (2004). Randomness and the Justification of Induction. *Synthese*, 138(1), 79–99.
- Carnap, R. (1962). *The Logical Foundations of Probability*. London: Routledge & Kegan Paul.
- Dawid, R. (2015). Turning Norton's Dome Against Material Induction. *Foundations of Physics*, 45(9), 1101–1109.
- Earman, J. (2004). Determinism: What We Have Learned and What We Still Don't Know. In *Freedom and Determinism*, eds. Campbell, J. K.; O'Rourke, M.; Shier, D. (pp. 21–46). MIT Press: Cambridge MA and London.
- Fleming, J. M. (1962). Domestic Financial Policies Under Fixed and Floating Exchange Rates. *IMF Staff Papers*, 9, 369–379.
- Hume, D. (1894). *An Enquiry Concerning Human Understanding*. Oxford: Oxford University Press.
- Keynes, J. M. (1921). *A Treatise on Probability*. London: Macmillan and Co.
- Kelly, T. (2010). Hume, Norton, and Induction without Rules. *Philosophy of Science*, 77(5), 754–764.
- Kneale, W. C. (1949). *Probability and Induction*. Oxford: Clarendon Press.
- Kyburg, H. E. (2006). Belief, Evidence, and Conditioning. *Philosophy of Science*, 73(1), pp. 42–65.
- Lange, M. (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

(pp. 43–91). North Holland: Elsevier.

Levi, I. (1967). *Gambling with Truth*. New York: Alfred A. Knopf.

Maher, P. (1996). The Hole in the Ground of Induction. *The Australasian Journal of Philosophy*, 74(3), 423–432.

Mundell, R. A. (1963). Capital Mobility and Stabilization Policy Under Fixed and Flexible Exchange Rates. *Canadian Journal of Economic and Political Science*, 29(4), 475–485.

Norton, J. D. (2003a). Causation as Folk Science. *Philosophers' Imprint*, 3, 1–22.

Norton, J. D. (2003b). A Material Theory of Induction. *Philosophy of Science*, 70(5), 647–670.

Norton, J. D. (2006). How the Formal Equivalence of Grue and Green Defeats What is New in the New Riddle of Induction. *Synthese*, 150(2), 185–207.

Norton, J. D. (2008). The Dome: An Unexpectedly Simple Failure of Determinism. *Philosophy of Science*, 75, 786–798.

Norton, J. D. (2010). There Are No Universal Rules of Induction. *Philosophy of Science*, 77, 765–777.

Norton, J. D. (Forthcoming). Author's Responses. *Studies in the History and Philosophy of Science*. https://www.pitt.edu/~jdnorton/papers/Material_Theory_Responses.pdf (Accessed 25th October 2020)

Norton, J. D. (Unpublished). *The Material Theory of Induction*. https://www.pitt.edu/~jdnorton/papers/material_theory/material_theory_April_23_chapters/material.html (Accessed 22 June 2020)

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.

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

Royce, J. (1913). The Principles of Logic. In *The Encyclopedia of the Philosophical Sciences*, eds. Windelband, W. and Ruge, A. (pp. 67–135). North Holland: Elsevier.

Russell, B. (1948). *Human Knowledge: Its Scope and Limits*. London: Allen and Unwin

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

Slomson, A. B. (1997). *Introduction to Combinatorics*. Boca Raton: CRC Press.

Toulmin, S. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.

Van Strien, M. (2014). The Norton Dome and the Nineteenth Century Foundations of Determinism. *Journal for General Philosophy of Science* 45(1), 167–185.

Whately, R. (1855). *Elements of Logic*. Boston and Cambridge: James Munroe and Company.

Williams, D. C. (1947). *The Ground of Induction*. New York: Russell & Russell, Inc.