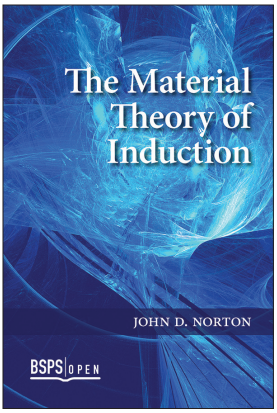


# THE MATERIAL THEORY OF INDUCTION

by John D. Norton

ISBN 978-1-77385-254-6

# 2

# What Powers Inductive Inference?[1]

## 2.1. Introduction

This chapter summarizes the case for the material theory of induction, drawing on material in other parts of the book. There are three arguments for the theory. The first two are the following:

1. *Failure of universal schemas*. Through many examples in this text, we see that no attempt to produce a universally applicable formal theory of induction has succeeded.
2. *Accommodation of standard inferences*. These same examples show that the successes of many exemplars of good inductive inferences can be explained by the material theory of induction.

These first two arguments suffice, I believe, to make a solid case for the material theory. They are developed in Sections 2.2 and 2.3. They make the case without giving an intuitive grounding for why the material approach is the right one. They establish *that* it is, not *why* it is. For the arguments succeed by showing that the alternative, formal approach fails and that the material approach works where its competitor fails. The third argument,

1 My thanks to the Fellows of and a visitor to the Center for Philosophy of Science for discussions of a draft version of this chapter on 30 November 2011 and 23 November 2014: Yuichi Amatani, Ari Duwell, Uljana Feest, Leah Henderson, Gabor Hofer-Szabo, Soazig LeBihan, Dana Tulodziecki, Adrian Wuethrich, Adele Abrahamsen, Joshua Alexander, William Bechtel, Ingo Brigandt (presenter), Sara Green, Nicholaos Jones, Maria Serban, and Raphael Scholl.

however, is grounded in the foundational question, developed in Section 2.4, of why any inductive inference should work at all—that is, "What powers inductive inference?" The question presumes that we cannot take the success of inductive inference for granted. If it works, it does so for an identifiable reason. The material theory answers:

> 3. *Inductive inference is powered by facts.* The ampliative character of inductive inference precludes universal schemas.

There are two steps in the argument for this conclusion, and they are developed more fully in Section 2.5. Briefly, the first step notes that inductive inference is, by its nature, ampliative. That is, unlike deductive inference, the conclusion asserts more than the premises. It amplifies what the premises say. For each sort of inductive inference, there will be worlds hostile to its success. Generalizing chemical properties of samples, for example, is futile in a world without stable chemical properties. Using an inductive inference presupposes that, as a factual matter, we are not in one of those hostile worlds. If the notion of these facts is construed broadly enough, commitment to them is all there is to accepting the logic. These are the facts warranting the inductive inference.

The second step specifies the character of these facts. They are not universal contingencies such as would warrant a universally applicable inductive logic. This is shown by our failure to identify a universally applicable inductive logic and our failure to exhibit such a universally warranting fact explicitly. Rather, the facts hold true only in limited domains so that there are many of them and the inductive logic each warrants has local applicability only.

The two sections following Section 2.5 illustrate these two steps. Sections 2.6 and 2.7 consider the inductive problem of extending the series 1, 3, 5, 7. It is insoluble without background facts to warrant the inference. Section 2.8 displays some more examples of warranting facts. Finally, our predisposition for treating inference formally is strong. Section 2.9 will seek to weaken the presumption that all theories of inference must be

formal by indicating limitations in the formal, non-contextual treatment of the most favorable case, deductive inference.

## 2.2. Failure of Universal Schemas

Formal approaches to inductive inference depend on supplying a universal template or schema. For example, in the last chapter, we saw the schema of enumerative induction

> Some (few) *A*s are *B*.
> Therefore, all *A*s are *B*.

Such templates are then used to generate the licit inductive inferences by substituting the content of the placeholders *A* and *B*. The enduring difficulty for formal theories is that no general account of inductive inference has provided a clearly articulated exceptionless schema. Therefore, all formal accounts fail, and by eliminating the only rivaling accounts, the material theory gains support.

That all formal schemas fail is difficult to show directly since there are many of them. What can be shown, however, is the failure of a representative sample, which is the approach taken in this book.[2] The mode of failure displayed by a given sample is sufficiently straightforward to make it likely that it will afflict all candidate schemas.

In the preceding chapter, we saw in the example of crystalline forms that the schema of enumerative induction fails. For it to be applied successfully to crystalline forms, we needed to add additional, formal conditions contrived to rule out all but the very small set of properties of crystals that support inductive generalization. The sequence of additional conditions seemed to have no discernible end. Once even a few were added, it became clear that the schema lost all semblance of generality.

In the next chapter, we will look at the requirement of the reproducibility of experiments, which is often introduced as a gold standard of evidence. On closer examination, however, it will be proven something

2 In earlier work (Norton, 2003, 2005), I sought to be more systematic. I showed how virtually all accounts of inductive inference fell into one of three families, each powered inductively by a single idea. Since the sample of failures reviewed here are spread over the three families, we have some assurance that they are adequately representative of the range of accounts.

less. We will see that it is a guide whose verdict is sometimes accepted and sometimes discarded. There is no formal rule that tells us when the principle is to be upheld and when not. It is a principle that holds except when it does not. The following chapter looks at reasoning by analogy, a form of inductive inference whose use has pervaded science from antiquity to the present. Once again, we will see that the bare schema is too impoverished to be used exceptionlessly. Efforts over the past century to augment the schema have led to supplements of monumental size while still not delivering a self-contained formal schema.

This pattern of failure continues in subsequent chapters. While considerations of simplicity are often invoked in discerning the bearing of evidence, they do not rest upon a factual principle of parsimony in nature. Notions of simplicity prove sufficiently elusive that there is no clear formulation of such a principle. Similarly the slogan "inference to the best explanation" is so familiar that one might presume that there is some hidden inductive power in explanation. The presumption fails on closer examination. Our notions of explanation are too varied and vague to harbor powers sufficient to support a universal scheme of inductive inference.

Finally, a series of chapters investigates what is, momentarily, the favored account of inductive inference in the literature in philosophy of science, Bayesian inference. Any aspirations of universal applicability fail. Several chapters develop cases in which a probabilistic logic cannot apply since such a logic would contradict symmetries inhering in the cases. There is a rich literature that seeks to establish the necessity of probabilities in representations of belief and inductive support. An examination of these arguments shows them all to be circular. This circularity is developed at length in a chapter devoted to the scoring rule approach. Finally, any Bayesian analysis is inductively incomplete in the sense that it always requires inductively potent prior probabilities to be specified externally. I report work elsewhere that shows that this incompleteness is not specific to the Bayesian system but troubles any calculus meeting certain weak requirements. It follows that no single calculus can cover all the inductive inferences of science. To repeat an earlier conclusion: all induction is local.

These examples embody modes of failure that, I believe, afflict all candidates for universal schemas of inductive inference. The schemas may simply be too vaguely specified at the outset to count as a logic of

induction, as is the case with inference to the best explanation. Or, if they are precisely specified, they prove too permissive and authorize too much, as is the case with enumerative induction. Efforts to restrict the schemas may specialize them so narrowly to one particular domain that they lose their universality. Or efforts may burden them with more conditions. And in adding them, we may need to import new notions—natural kinds, explanation, lawfulness—which in turn require further conditions for their explication, and so on without termination.

## 2.3. Accommodation of Standard Inferences

The last section offered a preview of the failure of familiar, formal schemas for inductive inference. These schemas were devised because each, to some degree, fits some collection of inductive inferences we deem licit. The second argument for the material theory is merely the flip side of this failure. Where the formal approach fails for these repositories of licit examples, the material theory succeeds.[3]

Once again, this can be read from the analyses of the previous and subsequent chapters. Curie inferred inductively from the crystalline form of mere specks of radium chloride to all samples of radium chloride. What licensed the inference was a hard-won fact from nineteenth-century work on crystals. It is what I have called the Weakened Haüy's Principle: "Generally, each crystalline substance has a single characteristic crystallographic form."

In the next chapter, we will look at the requirement of the reproducibility of experiments. This requirement proves not to be a universal inductive principle but is shown rather to arise in connection with a loosely affiliated but irregular collection of inductive inferences concerning repeated experiments. The otherwise inexplicable irregularity of such inferences becomes intelligible when we recognize that they are warranted by two classes of facts: those specifying when some process will yield the

3 In Norton (2003), I worked through the three families of accounts of inductive inference and showed briefly how the inferences of each account were materially warranted. The treatment of so many accounts there is necessarily brief. In this book, I seek to show the material warrant for standard examples of successful inductive inferences in much greater depth. As a result, fewer examples are treated.

experimental outcome of interest; and those specifying what may confound the experimental outcome. These facts specify when a replication of an experiment is evidentially significant. More importantly, they specify when the replication is not evidentially significant. The variation in the facts from case to case explains the irregularity of the whole collection.

Arguments from analogy are so varied in their form that, as we shall see in Chapter 4, they defy complete characterization even by quite elaborate formulae. The material theory resolves the problem by conceiving analogy in the same manner as scientists. For them, analogy is not an argument form but a fact that asserts the similarity of two systems. This fact warrants inductive analogical inference. The resulting inferences have as varied a form as the facts of analogy themselves. It is this broad range of variation that defeats efforts to find a universal formal characterization.

This pattern of material reconstruction persists with the analysis of inductive inferences grounded in notions of simplicity or explanation. Invocations of simplicity in specific cases are shown to be abbreviated invocations of background facts. Since the background facts vary from case to case, their summary in an inductively potent principle of parsimony is precluded. Similarly, in specific inferences to the best explanation, explanatory relations contribute nothing to the evidential import. Real examples of this sort of inference in science succeed through the mere adequacy of the favored hypotheses to the evidence and our success in eliminating its competitors by prosaic, non-explanatory means.

Finally, where the probabilistic representation of strengths of inductive support is appropriate, it is because there are specific background facts that warrant them. The examples are many, varied, and familiar. Both quantum mechanical and statistical mechanical systems in physics are governed by probabilistic physical laws. These laws provide the warrants for the probabilistic inductive inferences over them. In biology, mechanisms of inheritance in population genetics are governed by probabilistic laws. They, too, warrant probabilistic inferences. An important background probabilistic fact in many areas of the biological and social sciences is the presumption of sampling randomly from a population. This fact is important, for example, in the forensic identification of suspects through DNA analysis. It warrants the probabilistic inferences reported. A related case arises in controlled trials where subjects are randomized

into a test and control group. If the randomization is probabilistic, it introduces background probabilistic facts that can warrant probabilistic inferences about whether the effect measured could arise in case the treatment is ineffective.

These examples instantiate a familiar pattern. Whenever a cogent inductive inference appears in a science, it has proven possible to trace the warrant for the inference to background facts.

## 2.4. The Mystery of Inductive Inference

The discussion so far has been devoted to the two most visible problems associated with inductive inference:

*1. Which are the good inductive inferences?*

To answer this, we must specify how we distinguish the good from the bad inferences. The material theory of induction says we do so by identifying warranting facts; we do not seek the warrant in universal schemas. This first problem is entangled with another problem that is more fundamental but largely overlooked in the present literature. How can inductive inference work at all? That is,

*2. What powers inductive inference?*

Once we accept that inductive inference is powered by background facts, it becomes clear why the answer to the first question must lie in identifying the warranting facts.

The second question needs some elaboration. It is easy to take for granted that induction lets us do something remarkable. It lets us amplify our knowledge. We pay a small price for this amplification. Our new knowledge is not as certain as the old knowledge from which we proceeded. Sometimes the uncertainty is large. In important cases, the uncertainty is minuscule. Whether it is small or large, we still seem to get more than we should. The problem—the big mystery of induction—is to understand how this amplification can happen.

To sharpen the sense of why we need a solution to this second problem, consider an analogous problem. Imagine that we are in ancient Greece and encounter an oracle. In the darkness, we see the dim outline of the

sibyl, wailing and flailing. Her cries fall silent, and she issues several sharp proclamations that, over the course of time, turn out to be mostly accurate. And all of this for the price of a goat and few drachma in her bronze bowl. Were this to happen, we would not be satisfied merely to note that this oracle has extraordinary predictive powers. We would want to know how this were possible. What is it in the order of things that enables this sibyl to make these predictions?

The puzzle is the same with induction. It performs a similar miracle, but without the movie-quality special effects. Experience gives us a small part of space for a small span of time. Yet from knowledge of this fragment, we come to be sure that all things began some 14,000,000,000 years ago in an intense conflagration; that tiny smudges of light in the night sky are great galaxies of stars that duplicate our sun many times; and much more, down to the most minuscule structure of microbial life. We must ask, then, what is it in the order of things that allows induction to do this? What powers inductive inference?

The dominant trends in the present literature are incapable of satisfactorily answering these questions. To answer them adequately, both questions above need to be treated together. We cannot hope to know which are the good inductions without a clear and explicit idea of what powers induction. Answers to these questions in the literature have followed the model of deductive inference. This has driven us astray for millennia. It has led us to seek a non-contextual account of what powers induction and a formal answer to the problem of which are the good inductive inferences. Neither works for induction. The central claim of this chapter is that a successful account of induction is contextual and material.

## 2.5. The Foundational Argument

The most compact argument for a material theory of induction proceeds by answering the foundational question of what powers induction. It is powered by facts. As indicated in the introduction, the argument has two premises.

*Premise 1. Inductive inference is ampliative.* This means that the conclusion of an inductive inference amplifies. It asserts more than the premises. This distinguishes inductive inference from deductive inference.

For deductive inferences merely restate what we have already presumed or learned. There is no mystery in what powers deductive inference and permits its conclusions. We are just restating what we already have in the premises. The warrant lies fully within the premises. If we know all winters are snowy, it follows deductively that some winters are snowy.[4] This derives from the premise "all." If something is true of all, it is thereby true of some. The context in which we infer plays no role in powering the deductive inference. The inference succeeds no matter what either "winter" or "snowy" might mean. The meaning of "all" is enough to uphold the conclusion regardless of context. The inference is valid independently of whatever other facts may obtain about weather and climate.

It is quite different with inductive inference. From the premise that all past winters have been snowy in some location, we infer inductively that the next winter will be snowy there. Yet it is entirely possible that this prediction fails. When we conclude in favor of the prediction, we assert more than the premises warrant. Such a conclusion is viable only in certain worlds. Hospitable worlds include those where the climate is stable. An inhospitable world would be one experiencing global warming, in which the past pattern of snowy winters does not continue unaltered. We can generalize the crystallographic family of a crystalline substance from one sample to all because our world is hospitable through the background fact of Haüy's principle. But we cannot generalize the size of the one sample to all, for there are no background facts providing for restrictions on possible sample sizes. Correspondingly, we can generalize sizes of living organisms, for different types of organisms are restricted by their physical constitutions to specific scales. Insects cannot grow to human scale because their structures would be too weak to support their weight and they could no longer breathe by diffusion. Similarly, humans cannot shrink to the scale of insects. A shrunken human brain would have too few neurons for our cognition. At least this is true in our world, which is hospitable to the generalization. A science-fiction world, where the normal laws of science are suspended, however, might be another story.

4 To be clear, I follow the informal conversational presumption and tacitly assume that "All winters are snowy" is not true vacuously; that is, the truth of the proposition requires that there are some winters.

The examples above illustrate the general point: the factual assumption that our world is a hospitable one is the fact that, if true, warrants the inductive inference. But it may not always be apparent that this fact warrants the inference. It may appear that the warrant is still provided by some sort of schema. The inference to a future snowy winter, we may think, is still warranted by the schema:

All past *A*s have been *B*.
Therefore, the next *A* will be *B*.

This supposition, however, is incomplete. If used at all, the schema would have a purely intermediate role. It does not have universal applicability. We can use it in the case of a snowy winter only because the requisite background facts authorize it when we make the specific substitutions: "winter" for *A*; "snowy" for *B*. That is, a cascade of warrants may pass through a schema. The cascade terminates in facts that are the final warrant of the inference.

It is essential here to distinguish two ways that an inductive inference can fail: either by loss of an inductive bet in a hospitable world or by failure of an inductive inference in an inhospitable world. The first arises because accepting a warranted inductive inference still involves a risk. In a hospitable world with a stable climate, it is a warranted induction to infer from a past history of snowy winters that the next winter will be snowy. The next winter, however, may turn out not to be snowy. When a climate is stable, such fluctuations would be rarer but nevertheless possible. Losing an inductive bet like this must be distinguished from the second case in which it is imprudent to take the bet in the first place. If the background facts are of a warming climate in some location, then the background facts do not warrant the inference. If one persists and makes the inference, the conclusion may prove false. The failure reflects the lack of warrant of the inference, not a failure arising from traditional inductive risk.

The material theory of induction arises from the recognition that the truth of these background factual presumptions is all that is needed for the inductive inference to be warranted. One might imagine that this might not be so. The facts, we might suppose, play only a partial role in warranting the inductive inference. Might there still be a residual universal

formal schema or inductive rule that contributes to the warrant? If so, such a schema or rule would be subject to the same analysis just given. If it functions to authorize an inductive inference, then it amplifies what is already asserted in the premises and all other background facts. It cannot be universal in application, for there would be worlds inhospitable to it. And we should only use the rule or schema where it is hospitable to do so. That is, the warrant for its use is the factual supposition that the world is hospitable to it. Once again, the inductive warrant has terminated in facts that should be included with the true background facts needed to warrant the inductive inference at issue. In other words, the truth of the background factual assumptions, when construed broadly enough, is all that is needed to authorize the inductive inference. With that, we arrive at the first major tenet of a material theory of induction:

*Inductive inferences are warranted by facts.*

What remains open is the precise character of the warranting facts. There is little we can say at the general level about the nature of these facts. In particular cases, their character will be straightforward. Our inference to a future of snowy winters is warranted by the assumption that our local climate will persist pretty much as it has, so that winters without snow are possible but unlikely. If the climate warms sufficiently, however, these facts may fail and with them the inductive inference.

In some cases, the background facts may be such that the inductive inference would be deductive if we explicitly added the warranting fact as a premise. Then the inference would be an enthymeme, a deductive inference with a hidden premise. An example is this version of Curie's inference from the preceding chapter:

This sample of radium chloride is monoclinic.
*Generally*, each crystalline substance has a single characteristic crystallographic form (Weakened Haüy's Principle).

---

Unless exceptions encoded by "generally" intervene, all samples of radium chloride are monoclinic.

But it would also be entirely natural to detach the "Unless…" clause and have the inference:

> This sample of radium chloride is monoclinic.
> *Generally*, each crystalline substance has a single characteristic crystallographic form (Weakened Haüy's Principle).
>
> ———————————————————————————
>
> All samples of radium chloride are monoclinic.

This inference is inductive for we are taking the risk that the exceptions suggested by *generally* do not arise.

Corresponding complications arise if we infer inductively in the Bayesian framework. If we infer from prior probabilities to posterior probabilities by means of likelihoods using Bayes' theorem, then the inference is deductive. If we broaden the context, this ceases to be so. Propositions asserting evidence and background facts are not provided to us with probability measures. We add them. In doing so, we accept that we can represent their mutual relations of inductive support probabilistically and that their inductive consequences follow from the probability calculus. In this process, we take an inductive risk that probabilistic analysis correctly represents these relations. If we also proceed as normal people do and accept a proposition with a very high posterior probability as established, then we take a second inductive risk in detaching the qualification of high probability.

The second premise places a restriction on the character of the warranting facts:

*Premise 2. There is no universally applicable warranting fact for inductive inferences.* This premise requires support, part of which is supplied by other arguments in this book that seek to establish that there is no universally applicable logic of induction. For if there were, then there would be a universally applicable warranting fact according to Premise 1.

A more direct grounding for the second premise lies in our failure to exhibit such a universally applicable warranting fact. It has been long sought, like the philosopher's stone—and with equal success. The best-known attempt at characterizing it is Mill's principle of the uniformity of nature: "The universe, so far as known to us, is so constituted that whatever is true in any one case is true in all cases of a certain description; the

only difficulty is, to find what description" (Mill 1904, book 3, chap. 3, p. 223). To this, he added: "Whatever may be the proper mode of expressing it, the proposition that the course of nature is uniform is the fundamental principle, or general axiom of Induction" (p. 224). It is a general fact about the world that holds in all domains in which we may seek to infer inductively. It is the one, universal fact that would power all inductive inference.

The trouble with Mill's principle is that, read literally, it is false; and read charitably it is so vague as to be unusable. Take the literal reading. Our world is *not* uniform in all its aspects. Indeed, the world fails to be uniform in virtually all its aspects. Otherwise, we would live in a largely homogenous environment. At best, the world is uniform in a very few, quite special properties that end up figuring in what we take to be laws of nature. This last statement is the charitable reading. The real challenge for the principle is to specify just what its special properties are. Yet through the vague generality of its formulations, it provides no such specification. At best, the principle deflates to a weak existential claim: there are uniformly implemented properties in nature, but we do not know precisely which they are. Or, more generally, nature is regular and orderly but in a way that we cannot state or grasp compactly enough to implement as a principle that can be employed practically in a logic of induction.

That the principle needs this shield of ignorance to protect it from scrutiny suggests that there is no real content hidden behind the shield. The principle has ceased to have any practical value in our inductive investigations. Wesley Salmon (1953, p. 44) long ago wrote its obituary: "the general result seems to be that every formulation of the principle of the uniformity of nature is either too strong to be true or else too weak to be useful." This completes the argument for the premise.

If the facts warranting inductive inference are not universal truths, then they must be truths of restricted domains, and the inductive inferences they warrant will be restricted to those domains. It may well be that the inferences warranted in some restricted domain have a regular structure. Then we have an inductive logic applicable to just that domain. For example, Haüy's principle warrants an inductive logic that looks formally like enumerative induction but is restricted to generalizations concerning the crystallographic family of samples of crystalline substances. A general

statement of this restriction is the second major tenet of a material theory of induction:

*All induction is local.*

Philosophers are good at finding clever but ineffective loopholes. The following loophole is one that few can resist. If each domain has its own material facts that warrant inductive inferences in it, why not just combine them all? The resulting conjunction would be a single, *huge* fact that warrants inductive inferences in all domains.

It would be correct to assume that this huge conjunction would warrant inductive inferences in all domains. But it would not provide an escape from the necessarily local character of inductive inferences claimed by the material theory. That locality now reappears in the irreducibility of the huge conjunction to anything more compact. It remains just a single, huge conjunction of this fact and that fact and that other fact and so on, with many, many more conjuncts. To use the huge conjunction in any particular domain, we have to locate within the immensity the particular facts that applies to that specific domain, extract the particular facts while ignoring all others, and apply them. The warranting of inferences in that specific domain will still be done by facts prevailing just in that domain. The existence of the huge conjunction provides no universally applicable schema beyond the one already central to the material theory of induction: to identify the warrant of an inductive inference, seek facts that prevail in that domain.

The next two sections will illustrate the first and second premises respectively of the argument of this section.

## 2.6. The Inductive Inference on 1, 3, 5, 7, …[5]

To quickly see the importance of background warranting facts, an inductive inference problem bereft of background facts will help: Given the initial sequence of numbers 1, 3, 5, 7, how should this sequence continue? That the sequence could continue in many different ways is a trivial

5 This example and a briefer version of the argument of the previous section are given in Norton (2014).

mathematical fact. If the only restriction is that these are the first four terms of an infinite series, then there is an infinity of varying continuations. The lack of specification makes it impossible to favor any one in particular—that is, to pick among the deductively authorized possibilities. Without some specification of background facts, to infer inductively about the continuation is impossible.

The possibilities are greatly reduced if we assume naturally that the sequence is governed by some simple rule. There are still many possible continuations. The sequence may just be the odd numbers:

1, 3, 5, 7, 9, 11, 13, 15, …

Or it may be the odd primes, including one:

1, 3, 5, 7, 11, 13, 17, …

Or it may be the digits of the decimal expansion of 359/2,645:

1, 3, 5, 7, 2, 7, 7, 8, 8, 2, 8, …

While the possibilities in these cases are reduced, the inductive problem is still intractable since the notion of a "simple rule" remains underspecified. This makes finding other continuations merely a challenge to our ingenuity in writing laws that look simple in some sense we happen to find congenial.

Another approach embeds the sequence in a context for which we have more information. The numbers may be drawn from a randomizing lottery machine. The fact of randomization then authorizes a probabilistic analysis. Probabilistic inductive support is distributed uniformly over the remaining, undrawn numbers. Or perhaps the numbers appear in a question on an IQ test or in the interrogation of a psychologist we believe is intent on tricking us. These different background facts would then authorize different inferences about the continuations, although the complexity of the background would make discerning their precise character troublesome.

## 2.7. The Law of Fall

It is easy to suppose that the preceding inductive problem is merely a contrivance, unrelated to real problems of inductive inference in science, and thus one that we need not strive to accommodate in our account. This supposition is wrong. The problem is in fact one of the classic problems of inductive inference in science. This particular number sequence happens to figure in one of the great discoveries in the history of science. In his *Two New Sciences* (1638), Galileo presented his law of fall. In one form, the law asserts that the distances fallen in successive units of time stand in the ratios 1 to 3 to 5 to 7 and so on; that is, in the ratio of the odd numbers. Galileo's pathway to this law was long and convoluted. However, at least one part of it quite likely involved experimentally measuring the distances that bodies fall and the time this takes. In *Two New Sciences* ([1638] 1954, pp. 178–79), Galileo describes an experiment in which a ball is timed rolling down a grooved ramp. The ramp is a surrogate for free fall that slows the motion sufficiently to enable time measurements using Galileo's crude methods. Stillman Drake (1978, p. 89) has identified an early Galileo manuscript that, Drake argues, records the results of just such an experiment.

So let us pose a simple Galileo-like inductive problem. Given that the incremental distances fallen in successive units of time are in the ratios 1 to 3 to 5 to 7, what will be the distances in subsequent times? Using resources available to Galileo, how might this be solved?

We have a good idea of Galileo's methods. One element was that he presumed fall to be governed by a rule that was expressible simply in the mathematical techniques available to him. The idea is indicated in *Two New Sciences*. Galileo reflects on the gains in speed of falling bodies and asks of them, "why should I not believe that such increases take place in a manner which is exceedingly simple and rather obvious to everyone?" (p. 161). Galileo's inference is warranted by a fact: the simple behavior of bodies in free fall. Galileo's rhetorical question leaves the notion of simplicity at issue underspecified and thus leaves underspecified just which inference is authorized. If we read Galileo's writings more broadly, we find a stronger statement that identifies the notion of simplicity at issue. In a famous passage in *The Assayer*, he wrote:

> Philosophy is written in this grand book, the universe, which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and read the letters in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures without which it is humanly impossible to understand a single word of it; without these, one wanders about in a dark labyrinth. (1623, pp. 237–38)

This is a form of Platonism, which asserts that the world is structured as a copy of perfect mathematical forms. This factual statement about the world then warrants an inference to a simple mathematical rule as the continuation of the sequence 1, 3, 5, 7, ….

This approach may at first be appealing. The world does admit simple mathematical description. Why can we not use this fact to underwrite inductive inferences? The appeal fades rapidly under closer scrutiny. There are three problems.

First, if one is not a Platonist, one judges the warranting fact to be a falsehood and thus the inference an inductive fallacy. The success of mathematical methods in science since Galileo does not, in my view, justify the Platonic view. Rather, as I have argued elsewhere (Norton 2000, Appendix D), the success merely reflects the post hoc adaptability of mathematics to new scientific discoveries.

Second, attempts to employ the Platonic idea fall prey to the problem that the mathematical imagination can conjure up vastly more structures than are implemented in reality. Seek simple laws written in the wrong mathematical language, and our investigations will stall and fail. Einstein became a mathematical Platonist during his later-life search for a unified field theory.[6] His efforts were stymied by just this problem since he sought laws that could be simply expressed in the mathematics of tensors and the like on four-dimensional space-time manifolds. Subsequent theorizing in quantum gravity has branched out in the mathematical structures it uses

6 This is recounted in Norton (2000).

and typically does not posit a four-dimensional space-time manifold as a primitive.

Third, when Galileo investigated falling bodies, the mathematics accessible to him was limited to methods drawn from Euclid. They comprised the barest sliver of the mathematics we now employ. It would be naïve to assume that the Platonic blueprint of nature is drawn with the mathematics of this tiny sliver.

## 2.8. Invariance under the Change of the Unit of Time

In the face of these mounting difficulties, we may well wonder whether Galileo had the sufficient background facts to warrant what still appears to be a good inference. Fortunately, he did assume another background fact, which was perfectly tuned to warrant the inference and eliminate all but one of the open possibilities. This aspect of his work, however, typically receives scant attention.

Galileo's experimental methods were unable to fix a precise unit of time. At best, he could determine that, in one experiment, successive intervals of time were equal. He realized that his experimental result was stable in spite of this variability of time units. In measuring fall, he recovered the same ratios, 1 to 3 to 5 to 7 and so on, no matter what unit of time he used. This important fact is stated by Galileo in *Two New Sciences* when he presents this odd-number formulation of his law of fall. He wrote:

> Hence it is clear that *if we take any equal intervals of time whatever*, counting from the beginning of the motion, such as AD, DE, EF, FG, in which the spaces HL, LM, MN, NI are traversed, these spaces will bear to one another the same ratio as the series of odd numbers, 1, 3, 5, 7. ([1638] 1954, p. 175; emphasis added)

The invariance of the result is asserted by the text I have italicized.[7]

7 Galileo's Latin *quotcunque tempora aequalia* is literally "however so many equal times." Crew and de Salvio render it as "any equal intervals of time whatever." Their looser rendering fits with the overall context in allowing both the number and duration of the intervals to vary. An important part of the context is the earlier statement of the law of fall from which this

With a little arithmetic, we can see how this invariance under change of units of time works. In successive units of time, the body falls the following distances: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, …. Now replace the original unit of time with a new unit equal to two of the old units. The distances fallen in successive units of time with the new unit are

$$1 + 3, 5 + 7, 9 + 11, 13 + 15, 17 + 19, \ldots$$
$$= 4, 12, 20, 28, 36, \ldots$$
$$= 4 \times 1, 4 \times 3, 4 \times 5, 4 \times 7, 4 \times 9, \ldots .$$

Galileo's law requires only that these distances be in the ratios 1 to 3 to 5 to 7 and so on. Hence, we can neglect the factor of 4 and observe that the ratios conform to the law. This invariance obtains, Galileo asserts, no matter which unit of time we select.

The remarkable fact is that there are few laws of fall that respect this invariance. Using techniques in calculus and functional analysis not available to Galileo, it is possible to prove that the *only* laws are these. If $d(t)$ is the fall distance in the unit of time $(t - 1)$ to $t$, then[8]

$$d(t) \text{ is proportional to } t^p - (t - 1)^p,$$

where $p$ is any real number greater than 0 (see Norton 2014a). This means that prior to any measurements, the scope of the law admissible is already reduced to these very few possibilities.

What now gives the inference great strength is that there is just one free parameter in the law, $p$. It follows that securing just one ratio of distances experimentally fixes the law uniquely. For example, take the first ratio that Galileo would have measured, $d(2)/d(1) = 3$. It follows that $p$ must satisfy

corollary is derived. The law is first introduced as "during any equal intervals of time whatever, equal increments of speed are given to it" (p. 161). Galileo's Latin *dum temporibus quibuscunque aequalibus* is correctly rendered by Crew and de Salvio as "during any equal intervals of time whatever," where *quibuscunque* has no restriction to number or duration. These unrestricted, equal time intervals are the ones that reappear in Corollary I.

8 There is a suppressed proportionality constant in the statement. It is suppressed since Galileo's law concerns ratios of the quantities $d(t)$, and the constant will not affect those ratios.

$$3 = \frac{2^p - (2-1)^p}{1^p - (1-1)^p} = \frac{2^p - 1^p}{1^p} = 2^p - 1.$$

The unique solution is $p = 2$ so that $d(t)$ is proportional to

$$t^2 - (t-1)^2 = t^2 - (t^2 - 2t + 1) = 2t - 1.$$

Hence, for successive times $t = 1, 2, 3, 4, \ldots$, we have $d(t) = 1, 3, 5, 7, \ldots$—that is, the odd numbers.

This is a remarkable result, and it is worth restating: if invariance under change of units of time is to be respected, the only continuation of the two-membered sequence of incremental distances fallen

$$1, 3$$

is the sequence of odd numbers

$$1, 3, 5, 7, 9, 11, 13, \ldots.$$

Of course, Galileo could not have known this result in all generality. But it is quite likely that he was aware of how restrictive the invariance was. One needs only to try out a few alternatives to the odd-number sequence arithmetically to realize that all simple alternatives fail. Drake (1969, pp. 349–50) notes that a correspondent of Galileo's, Baliani, reported that Galileo had used the invariance as a "probable reason" for the odd-number rule.

While Galileo did not elaborate in *Two New Sciences* on this result, Christiaan Huygens soon did. When he was only seventeen years old, Huygens found the result independently, prior to reading Galileo's *Two New Sciences*.[9] One statement of what he found is given in a letter of 28 October 1646 to Marin Mersenne (Huygens 1888, pp. 24–28). We see there that Huygens arrived at his result by considering two possibilities: that either the incremental distances fallen in subsequent, equal intervals of time increase in an arithmetic progression, or that they increase in a geometric progression. Only one case gave non-trivial results: an arithmetic

9 I thank Monica Solomon for drawing my attention to Huygens' work and for sending me a copy of his letter and other supporting materials.

progression in the ratios of the odd numbers, 1, 3, 5, 7, …. The demonstration is creditable but less than general since it overlooks the possibility of expressions for the incremental distances $d(t)$ with values of $p$ other than 2 in the formula $t^p - (t - 1)^p$. Thus it precludes by supposition many other progressions that would give a law of fall whose ratios would remain unchanged under change of units of time. While one might imagine ways that the demonstration could be rendered more general, there seems to be no obvious way to arrive at the general proof without mathematical techniques stronger than those available to Galileo and Huygens—for instance, those used in Norton (2014a).[10] This may explain why Galileo did not elaborate on the result in *Two New Sciences*.

Our Galileo-like inductive inference problem admits a ready solution. We take as a premise that the ratios of the incremental distances fallen in successive units of time are 1 to 3 to 5 to 7. There are two warranting facts accessible to Galileo: that the rule governing the sequence is expressible simply; and that the rule is invariant under change of units of time. Only a small amount of arithmetic exploration will show that this invariance likely rules out all extensions other than the odd numbers. A fuller analysis shows that the second invariance by itself is sufficient to warrant the inference.

## 2.9. Can Bayes Help?

One might imagine that the general inductive problem of extending the initial sequence 1, 3, 5, 7 is one where a Bayesian method would excel. But would it succeed without the need for specific background facts despite everything that has been said so far? In short, the answer is that it does not provide a successful, universal treatment of the problem. There are two striking failures in the analysis. First, Bayesian analysis fails to offer any inductive learning from the evidence of the initial sequence 1, 3, 5, 7. Second, prior probabilities control the analysis, but the requirement that they normalize prevents them being set in a manner that is universally benign.

10 One way is to consider not the incremental distances $d(t)$ but the total distance $s(t)$ fallen by time $t$. Then it is easy to show that the invariance is satisfied by setting $s(t)$ proportional to $t^p$ for any $p > 0$. However, showing that these are the *only* laws satisfying the invariance is harder.

To proceed, we will see how a Bayesian analysis might help us decide between two extensions of the sequence 1, 3, 5, 7:

The odd numbers $H_{\text{odd}}$: 1, 3, 5, 7, 9, 11, 13, 15, …
The odd primes with one $H_{\text{prime}*}$: 1, 3, 5, 7, 11, 13, 17, …

using the evidence $E$: 1, 3, 5, 7.
The ratio form of Bayes' theorem asserts that

$$\frac{P(H_{\text{odd}} \mid E)}{P(H_{\text{prime}*} \mid E)} = \frac{P(E \mid H_{\text{odd}})}{P(E \mid H_{\text{prime}*})} \cdot \frac{P(H_{\text{odd}})}{P(H_{\text{prime}*})}.$$

Since each $H_{\text{odd}}$ and $H_{\text{prime}*}$ deductively entails $E$, we have $P(E \mid H_{\text{odd}}) = P(E \mid H_{\text{prime}*}) = 1$. Therefore, Bayes' theorem reduces to

$$\frac{P(H_{\text{odd}} \mid E)}{P(H_{\text{prime}*} \mid E)} = \frac{P(H_{\text{odd}})}{P(H_{\text{prime}*})}$$

According to the theorem, what have we learned from the evidence $E$? The prior probabilities $P(H_{\text{odd}})$ and $P(H_{\text{prime}*})$ represent our initial uncertainty about the two hypotheses; the posterior probabilities $P(H_{\text{odd}} \mid E)$ and $P(H_{\text{prime}*} \mid E)$ represent their new values after incorporating evidence $E$. The reduced form of Bayes' theorem just tells us that conditionalizing on the evidence makes no difference to our comparative uncertainty concerning the two hypotheses. The ratio of the prior probabilities is the same as the ratio of the posterior probabilities. This will be true for any pair of hypothesized sequence that starts with 1, 3, 5, 7. In short, we have learned nothing new from the evidence as far as our decision between the two hypotheses is concerned.

Hypotheses logically incompatible with the evidence will be eliminated. Take, for example, the natural numbers represented by $H_{\text{nat}}$: 1, 2, 3, 4, 5, 6, …. Since $H_{\text{nat}}$ is logically incompatible with $E$, we have $P(E \mid H_{\text{nat}}) = 0$, and the posterior probability will be $P(H_{\text{nat}} \mid E) = 0$. But this result is not an inductive result. We have simply eliminated all hypotheses deductively incompatible with the evidence. The deductive result is easily obtained without the probability calculus or any other inductive manipulations. Where we need help is with the inductive problem. Does the evidence $E$ favor some hypotheses among those with which it is deductively compatible?

Here, the Bayesian analysis has failed to provide anything useful. Our inductive preferences are unchanged by learning the evidence.

This is a somewhat discouraging start. Nevertheless, it will be instructive to press on and ask what our posterior probabilities may be with specific prior probabilities. The analysis bifurcates according to whether we are subjective or objective Bayesians. If we are subjective Bayesians, then our prior probabilities are merely expressions of prejudice, constrained only by compatibility with the axioms of the probability calculus. We might decide that these prejudices dictate that $H_{\text{odd}}$ has three times the probability of $H_{\text{prime}*}$. Then we conclude for our posterior probabilities that

$$P(H_{\text{odd}} \mid E) = 3\, P(H_{\text{prime}*} \mid E).$$

Looking at the equation, it may seem that we have learned something. But we have not. The threefold difference in posterior probabilities is a direct restatement of our prior prejudices.

If we are objective Bayesians, we will seek prior probabilities that objectively reflect what we know. In this case, by supposition, we know nothing initially, so we have no reason to prefer one hypothetical sequence over any other. Hence, the appropriate prior probability will assign the same, small probability $\varepsilon$ to each hypothesis. That is, we have

$$P(H_{\text{odd}}) = P(H_{\text{prime}*}) = \varepsilon.$$

The reduced form of Bayes' theorem now tells us:

$$P(H_{\text{odd}} \mid E) = P(H_{\text{prime}*} \mid E).$$

Once again, we have learned nothing. Our initial assumption was that all hypotheses are equally favored, and this remains true for any pair compatible with the evidence.

This last conclusion overlooks a complication that will gravely trouble both subjective and objective Bayesians. The prior probability distribution must normalize; that is, the prior probabilities assigned to all the possible sequences must sum to unity. There is an uncountable infinity of possible sequences.[11] This means that, in a strong sense, most sequences must be

11 To see that the set is at least continuum sized, we should note that a subset of sequences using the digits 1 and 2 only can be mapped one-to-one onto the real numbers in the interval [0, 1].

assigned zero prior probability. Once a sequence has been assigned zero prior probability, its posterior probability on any evidence whatever will also be zero. This means that no evidence, no matter how favorable, will move us to entertain the sequence in the slightest. Hence, both subjective and objective Bayesians must make unavoidably damaging decisions, prior to any evidence, as to which few sequences will be learnable.

Of course, there are ways we might try to work around the problem. We might try to retain the uniform prior probability distribution simply by dropping the requirement of normalization and using so-called "improper priors." This violation may be excused if it turns out that, after conditionalization, the posterior probability distribution is normalizable. Normalizability is not achieved in this case, however. There are infinitely many sequences beginning with 1, 3, 5, 7. After we conditionalize on this evidence, we will be assigning equal non-zero probability to each sequence in this infinity of sequences. Normalization will fail.

More drastically, we might retain a uniform prior probability distribution by the artifice of simply choosing a finite subset of sequences and casting the rest into the darkness of zero probability. If we eschew the uniformity of prior probabilities for variable probabilities, we can expand the set of sequences with non-zero prior probabilities to a countably infinite set. As long as the prior probabilities diminish fast enough as we proceed through the set, the sum of the probabilities can be unity, as normalization requires. One way of achieving this diminution is to assign these varying non-zero probabilities only to sequences that are arbitrarily long, but always of finite length. If we do this, we need some rule to decide which sequences are more probable and which are less. A popular choice is to use a prior probability distribution advocated by Solomonoff (1964). Briefly, describable sequences, like 1, 2, 1, 2, 1, 2, …, have greater prior probability than sequences without simple descriptions. A prior probability distribution is implemented by penalizing each sequence's probability by an exponential factor $(1/2)^N$, where $N$ is the length of the shortest

---

The sequence 1, 1, 2, 2, 1, 1, 2, 2, … is mapped to the fraction in binary notation 0.00110011…, etc. To see that the set is no bigger, we should note that we can map any sequence to a real number in [0, 1] by replacing the symbol "," by the symbol "0". The sequence 1, 3, 5, 7, 9 , 11, 13 … is mapped to the real 0.1030507090110130…, etc. The map is not "onto" because some real numbers, such as 0.100010001, have no corresponding sequence.

description possible for the sequence.[12] Bayesian analysis that employs this prior probability distribution is celebrated with joyous but naïve enthusiasm as providing a "complete theory of inductive inference" (Solomonoff 1964, p. 7) or "universal induction" (Rathmanner and Hutter 2011).

The difficulty is that the comparative judgments of a prior probability distribution will never go away. They determine how we might discriminate between $H_{\text{odd}}$ and $H_{\text{prime}^*}$ on learning the evidence $E = 1, 3, 5, 7$. Thus the selection of this prior probability distribution is not benign. It must be justified by something solid. Are we to suppose that, as a general proposition, our world favors sequences with short Turing machine programs? This favoring might be credible in specific contexts, such as one where we know that people are thinking up the sequences. But we are to suppose that this favoring is true prior to any restriction whatever on where these sequences may appear. It is hard to see any reason for why the world would universally prefer to present us with number sequences that are computable and in such a way that exponentially penalizes sequences with longer programs. The literature supporting the Solomonoff approach holds otherwise and matches its joy in its solution of the inductive problem with equally joyous pronouncements grounding the approach. Authors of this literature often resort to appeals to simplicity through "Occam's Razor" (Solomonoff 1964, p. 7; Rathmanner and Hutter 2011, p. 1101). This reveals an inflated reverence for the insights of a medieval scholastic who wrote six centuries before Turing conceived the notion of a universal Turing machine. For more deflation of simplicity, see Chapter 6.

In short, the challenge of accommodating the requirement of normalizability greatly complicates the analysis. More generally, the Bayesian analysis itself creates troubles that multiply and whose intractability deepens the more we try to resolve them. We could continue to wrestle with them. Or we could see that the very fact that we face lingering problems of this gravity tells us that Bayesian analysis is just the wrong instrument for this inductive problem. Compare this with the simplicity of the material analysis of the problem of extending 1, 3, 5, 7. Once we locate the appropriate context, as in Galileo's law of fall, we find that the requirement

12 $N$ is usually taken to be the length of the shortest Turing machine program that would output the sequence.

of invariance under change of units of time fixes the extension all but completely.

## 2.10. Warranting Facts

What might other warranting facts look like? Once we realize that familiar facts may serve also to warrant inference, we see that we are surrounded by such warranting facts.

Cosmology seeks to discover the structure of the universe on the largest scale. If the universe is infinite in spatial extent, then the finite portion observable by us is minuscule. What we see is infinitely outweighed by what we cannot see. The essential assumption that allows us to proceed from what we can see to what we cannot is the "cosmological principle." It asserts that the universe is roughly homogenous in its large-scale properties. While this wording may seem somewhat vague, standard applications of the principle employ it unambiguously. In our vicinity of the universe, matter is distributed roughly uniformly in galaxies in a space of constant, possibly zero, curvature. The cosmological principle authorizes us to infer that this condition obtains everywhere in the whole universe. Much of modern cosmological theory proceeds from this authorization.

Assume we have some isolated system with a given quantity of energy and entropy. The principle of the conservation of energy—the first law of thermodynamics—authorizes us to infer that, however else it changes, the same isolated system will have the same energy at any point in the future. The second law of thermodynamics authorizes us to infer a similar conclusion about the entropy of the system: it will be the same or greater. A careful statement of the second law merely allows that, with very high probability, the entropy of such systems will be the same or greater. Hence, the conclusion is warranted inductively but with very great certainty.

Assume we have some experiment performed in an isolated laboratory. The principle of relativity authorizes us to infer that a uniformly moving replica of the experiment will yield the same result. A more careful factual statement of the principle allows that it would hold only in regions of space-time that are remote from intense gravitational fields and thus unaffected by the curvature of space-time revealed through the general

theory of relativity. So the factual principle informs us that, *mostly*, the same experimental result will obtain. Thus, the inference is inductive.

These examples are designed to illustrate a progression in two aspects. First, a progression from the more general to the more specific and local. Second, a progression from examples where the mediating facts authorize the conclusion deductively to those where they authorize them inductively. The next and final example extends the progression farther to a case of greatly narrowed scope and greater inductive risk.

Assume we set up some simple chemical process whose feed includes nitrogen gas. A general fact of chemistry is that nitrogen gas is quite unreactive. Its diatomic molecules are held together by a strong triple bond that is hard to break. This general fact authorizes us to infer, at a relatively high level of inductive certainty, that the simple chemical process will leave the nitrogen gas unaltered. We are not assured of the conclusion with deductive certainty. There are extreme conditions under which nitrogen gas can be compelled to enter into reactions, as the Nobel Prize-winning work of Haber and Bosch demonstrated a century ago. Their Haber-Bosch process enables the chemical industry to synthesize ammonia from nitrogen and thereby manufacture both fertilizers and explosives.

This progression gives us factual principles of increasingly narrow scope that warrant inferences inductively. The material theory of induction places no lower limit on the size of the domain over which these factual principles operate.

## 2.11. A Non-Contextual, Formal Logic is Exceptional

The scope of successful applications of deductive logic that are non-contextual and formal is enormous. It is one of the great achievements of human thought. Its success makes it easy to think that the right way and only way to analyze inference is with non-contextual, formal theories. Correspondingly, then, one might think of a materially warranted logic as some kind of failure, perhaps the result of insufficient efforts to find that elusive, universal formal logic of induction. I will argue in this section that the success of non-contextual, formal accounts of deductive logic is exceptional. Hence, we should not use our familiarity with deductive logic to set

our expectations for inductive logic. We should not allow this to make us expect that there is a non-contextual, formal logic of induction.

### 2.11.1. The Undeserved Success

Which are the good deductive inferences? As long as the problems are kept simple, most people have a good instinctive grasp of the deductive consequences of their knowledge, and they manage without external guidance. But the limits are readily breached. If each thing has a cause, does it follow deductively that there is one ultimate cause for all things? If for every moment of time there is a later moment of time, does it follow that time endures infinitely? Novices relying on instinct can easily falter in the face of such traps. Can we find an instrument that systematically and reliably separates the good deductions from the bad? The means of discerning the good deductions is so familiar to anyone with a familiarity with modern logic that it is easy to underestimate the difficulty of the problem.

This problem was all but solved millennia ago with a simple, profound observation. To illustrate with a modern example, if you know that "All electrons have spin half," then you know that "Some electrons have spin half." The deductive inference is assured even if you have no idea of what an electron is and even less of an idea of what "spin half" is. You can make the inference merely by attending to the form of the sentences and ignoring the material. You start with "All *A*s are *B*" and know that you are then authorized to infer that "Some *A*s are *B*." You can ignore the details about electrons and spin; all that matters is the form of the sentences.

That deductive inference can proceed in such a simple and efficient manner is a marvel. It is the basis of a formal theory of inference, for we separate out the allowed inferences from the prohibited inferences merely by looking at their form. Specifying the logic then simply amounts to providing a list of schemas, such as

All *A*s are *B*.
Therefore, some *A*s are *B*.

To use the schemas, we replace *A* by anything we like and *B* by anything else we like and—bingo!—there's a valid deductive inference.

This example shows that the success of the schema depends on the non-contextuality of deductive inference. We can transport this schema to any domain, substitute anything for *A* and *B*, and still be assured that a valid inference results.

This simple schema is just the beginning. Generations of logicians have supplied us with a growing repertoire of schemas that embrace many logical operators. Sentential logic, for instance, employs the connectives "not," "or," and "and." One of De Morgan's laws is the schema

Not-(*A* and *B*).
Therefore, (not-*A* or not-*B*).

Predicate logic includes individuals and their relational properties, and it allows us to quantity over the individuals. If all things "*x*" gravitate "*G*(*x*)," then it is false that something exists that does not gravitate. This is an application of the schema

For all *x*, *G*(*x*).
Therefore, not-(there exists *x*, not-*G*(*x*)).

Modal logic introduces modal operators, like "It is possible that…" and "It is necessary that…." Tense logic introduces temporal operators, such as "It is always…" and "It is sometimes…."

### 2.11.2. Context Dependence of Connectives

In the face of the successes of deductive logic, it may seem that the scope of formal methods in logic is unlimited. However, lingering and recalcitrant anomalies limit the scope of the formal approach. Such anomalies manifest in deductive logic when the logical terms used have meanings that are context dependent. Does "some" just mean "at least one"? Or does it mean "more than one but not too many"? The answer varies with the context. Consider this mathematical assertion: "For some *x*, the quotient $1/x$ is undefined." Here, "some" can mean "one or more," and the single case of $x = 0$ is the one that makes the sentence true. But consider "some" in the following context: "Some voters disapprove of the governor's decision." This "some" requires more than one voter, but probably not a majority.

This difference matters in the formal theory, for not all schemas we may wish to use for "some" will apply everywhere. Consider

> Some *A*s are *B*.
> Therefore, more than one *A* is *B*.

The schema applies to the "some" of the voters but not to the "some" of division by *x*. The schema is context dependent; it is not universally applicable.

The humble conditional "If *A* then *B*" has proven to be a more notorious locus of this sort of trouble. A natural understanding is that this conditional is true when knowing *A* authorizes you to know *B* as well. That is, the conditional can be a premise in the argument form modus ponens:

> If *A* then *B*.
> *A*.
> Therefore, *B*

The validity of the inference is secured if the conditional "If… then…" is the "material conditional." Accordingly, "If *A* then *B*" is the same as "Either *A* is false or *B* is true." Thus, if we happen to know that *A* is true, then we know the first option ("*A* is false") fails. So that leaves the second, "*B* is true." Hence, the material conditional has done the job of allowing us to proceed from knowing *A* to knowing *B*.

All of this may seem quite fine until one realizes that, with this understanding, the conditional "If *A* then *B*" turns out to be true whenever *B* is true, no matter what *A* says. That is, both statements "If pigs have wings, then the sky is blue" and "If the grass is green, then the sky is blue" turn out to be true, material conditionals simply because the sky is blue. The natural objection is that an "If *A* then *B*" statement can only be true if there is something in the antecedent *A* that makes the consequence *B* true. The objection fails in these last examples. Whether pigs have wings or the grass is green is irrelevant to the blueness of the sky. But the statement "If the sunset is red, then the sky is blue" can be a true conditional. For the sunset is red because the blue light from a setting sun has been scattered away by the air, and the blue light comprises the blue sky. The blue of the sky is directly relevant to the red of the sunset.

Ingenious systems of relevance logic have sought to formalize the schemas into which "If… then…" properly enters, if understood relevantly. However, deciding just what is relevant to what is a delicate issue that may embroil us in significant portions of science. The blueness of the sky results from the Rayleigh scattering of blue light by the air's nitrogen and oxygen atoms, which just happen to be the right size for the job. Likewise, arcane facts in atomic theory are also relevant but perhaps not as directly relevant as the redness of the sunset. This tells us that relevance is context dependent and may vary in strength. Indeed, relevance may prove to be so diffuse that it may not be possible to separate off a small, tight formal logic of relevance as anything other than a crude gloss of a richer relation that is inextricably connected with the factual material of the science.

More generally, the success of a universally applicable formal logic of deduction depends on deductive inference being non-contextual. Whenever simple connectives fail to have a non-contextual meaning, as in the examples above, the logic in which they appear ceases to be universal.

### 2.11.3. Sellars' and Brandom's Material Inference

The anomalies for a formal theory of deductive inference above focused narrowly on logical connectives ("If…, then…") and operators ("Some…"). And I have argued that such connectives have a context-dependent meaning that is incompatible with their universal applicability—or at least they cannot have such applicability if we fix their meanings once and for all. Wilfrid Sellars and Robert Brandom developed a broader and more powerful critique of formal approaches to inference in general, not just deductive inference.

Their concerns were not limited to connectives but to all terms that appear in inferences. Their core idea is that the meaning of the terms in propositions is what makes good the inferences in which they correctly appear. Brandom (2000, p. 52) provides an example of the inference from "Pittsburgh is to the west of Princeton" to "Princeton is to the east of Pittsburgh." We recognize this as a good inference, but not for formal reasons. Rather, it is good because of the contents of the concepts of east and west. That is, the matter of the inference makes it good.

When I developed the material theory of induction, I was not aware of Sellars' and Brandom's notion of material inference and, in particular,

Brandom's use of the term "material inference." I learned of it through a lovely note written by Ingo Brigandt (2010), which usefully develops and applies the notion of material inference.

The difficulty is that my notion of material inference and that of Sellars and Brandom differ slightly, as far as I can see. This means that it would have been better at the outset if I had chosen another name. For Brandom, the above inference is material since it is made good by the concepts invoked in the premises. In my view, it is material since I locate the warrant for the inference in the background material fact: if something is east of something else, then the second is west of the first. Here, I leave open whether this difference is consequential or merely a different entry point into a collection of views that largely agree.

## REFERENCES


Brandom, Robert. 2000. *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.

Brigandt. Ingo. 2010. "Scientific Reasoning Is Material Inference: Combining Confirmation, Discovery, and Explanation." *International Studies in the Philosophy of Science* 24: pp. 31–43.

Drake, Stillman. 1969. "Galileo's 1604 Fragment on Falling Bodies (Galileo Gleanings XVIII)," *British Journal for the History of Science* 4: pp. 340–58.

———. (1978) 2003. *Galileo at Work: His Scientific Biography*. Chicago: University of Chicago Press. Reprint, Mineola, NY: Dover.

Galilei, Galileo. (1623), 1957. *The Assayer*. In *Discoveries and Opinions of Galileo,* edited and translated by Stillman Drake, pp. 231–80. New York: Doubleday & Co.

———. (1638, 1914) 1954. *Dialogues Concerning Two New Sciences*, translated by Henry Crew and Alfonso de Salvio. MacMillan. Reprint, New York: Dover.

Huygens, Christiaan. 1888. *Oeuvres Complètes de Christiaan Huygens*. Vol. 1. La Haye: Martinus Nijhoff.

Mill, John Stuart. 1904. *A System of Logic: Ratiocinative and Inductive*. New York and London: Harper & Brothers Publishers.

Norton, John D. 2000. "'Nature in the Realization of the Simplest Conceivable Mathematical Ideas': Einstein and the Canon of Mathematical Simplicity." *Studies in the History and Philosophy of Modern Physics* 31: pp. 135–70.

———. 2003. "A Material Theory of Induction." *Philosophy of Science* 70: pp. 647–70.

———. 2005. "A Little Survey of Induction." In *Scientific Evidence: Philosophical Theories and Applications*, edited by P. Achinstein, pp. 9–34. Baltimore: The Johns Hopkins University Press.

———. 2014. "A Material Defense of Inductive Inference." *Synthese* 191: pp. 671–90.

———. 2014a. "Invariance of Galileo's Law of Fall under the Change of the Unit of Time," http://philsci-archive.pitt.edu/id/eprint/10931.

Rathmanner, Samuel and Marcus Hutter. 2011. "A Philosophical Treatise of Universal Induction." *Entropy* 13: pp. 1076–1136.

Salmon, Wesley C. 1953. "The Uniformity of Nature." *Philosophy and Phenomenological Research* 14: pp. 39–48.

Solomonoff, Ray. 1964. "A Formal Theory of Inductive Inference." *Information and Control* 7: pp. 1–22; pp. 224–54.