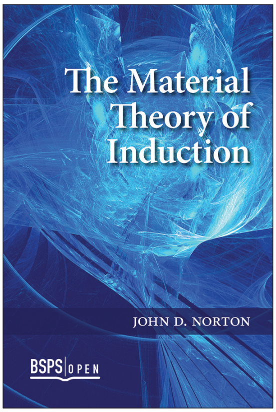

THE MATERIAL THEORY OF INDUCTION

by John D. Norton

ISBN 978-1-77385-254-6

Acknowledgement: *We acknowledge the wording around open access used by Australian publisher,* **re.press**, *and thank them for giving us permission to adapt their wording to our policy* http://www.re-press.org

8

Inference to the Best Explanation: The General Account

8.1. Introduction

This chapter and the next address the inductive inference form known as "inference to the best explanation" or "abduction." The main idea is that a theory or hypothesis must do more than merely accommodate or predict the evidence. If it is to accrue inductive support from the evidence, it must explain it. Since multiple explanations are possible, we are enjoined to infer to the best of them. This means that greater explanatory prowess confers greater inductive support. In 1964, Arno Penzias and Robert Wilson found puzzling residual noise in their radio antenna that turned out to be cosmic in origin. Subsequent investigation showed it to be thermal radiation of 2.7 degrees kelvin. The radiation was explained by Big Bang cosmology as the much diluted and cooled thermal radiation left over from the hot Big Bang over 10^{10} years ago. The competing steady-state cosmology and other now less well-known models could provide no comparably strong explanation. Cosmologists inferred to Big Bang cosmology as the best explanation.

Inference to the best explanation, however, has proven to be an especially troublesome case for my project. The difficulty does not lie with the material theory of induction. The difficulty lies with inference to the best explanation itself as an inductive inference form. Beyond the simple sketch just given, its elaboration is noticeably thin in the literature.

This thinness persists in spite of efforts to deepen our understanding of the inference form. A starting point is the notion of explanation itself in science. The general literature in philosophy of science has sought to elevate the notion beyond mere psychological satisfaction with some theory or hypothesis. It has become a core notion in philosophy of science and a subject of intense philosophical scrutiny. As far as abductive inference is concerned, the hope has been that this scrutiny will reveal something in the nature of explanation that makes it peculiarly potent in powering abductive inferences, and that this in turn will enable a more precise statement of the general rule of abduction. This expectation has set the scene for decades of frustration. Philosophical analysis of explanation has failed even to find a univocal sense of explanation at work in science. Instead, it has found multiple, competing senses of explanation. This multiplicity indicates that the notion is a loose one—an umbrella concept covering several disparate notions. They have no common core, such as might power a formal, inductive inference schema. As a result, the literature has provided no universal, formal account of abductive inference. Even the best developed accounts offer only superficial descriptions that use terms like "explains" and "loveliest" without precise, formal definitions.

Most of the analysis of this and the next chapter, then, is devoted to an attempt to do better at understanding just how the inferences designated as abductive work. These efforts draw on a series of canonical examples of abduction in science, described in the next chapter. My initial hope was that these examples would reveal the secret ingredient in good explanations that rewards explanatory prowess with inductive support. I would then seek its material underpinning. The plan has failed, and the secret ingredient has proven elusive. The inductive support proved time and again to come indirectly through weaknesses of competing explanations as opposed to from some special virtue of the preferred explanation. This has led to the curious notion developed in this chapter of "inference to the best explanation without explanation."

The upshot is that inference to the best explanation is an overrated argument form. Its strength is its visceral appeal. We apply it when we have a hypothesis or theory that fits the evidence in a strikingly satisfying manner. It just feels right, even if that feeling is created retrospectively from sanitized textbooks accounts. What remains is to move our affection for the argument form from psychology to reason. That is, we need to find a unified account of just how the inference works and what warrants it.

If explanatory prowess is what powers the argument, then there is good reason to suspect that no such unified account can be given. For we have to hope that a heterogeneous notion of explanation can somehow underwrite a homogeneous inductive power. In this case, inference to the best explanation will remain merely a label for a heterogeneous group of inferences powered more by visceral intuitions than good reasons.

In coming to these conclusions, I join a persistent, minority tradition in the philosophy of science that has deprecated the importance of explanation in inferences identified as abductive. The conclusions conform with those of Timothy Day and Harold Kincaid:

> In short, appeals to the best explanation are really implicit appeals to substantive empirical assumptions, not to some privileged from of inference. It is the substantive assumptions that do the real work. (1994, p. 282)

They associated this view with the similar approach to arguments based on simplicity advocated by Elliott Sober (1980) and also developed here in Chapters 6 and 7. Bas Van Fraassen's (1977; 1980, chap. 5) pragmatic deflation of explanation is well known. More recently, William Roche and Elliott Sober (2013) made their main claim clear in the title of their paper: "Explanatoriness is evidentially irrelevant, or inference to the best explanation meets Bayesian confirmation theory." Kareem Khalifa et al. (2017) argued that inference to the best explanation does not provide a fundamental argument form. Rather, its instances are reducible to other inferences, and these are not unifiable by a simple scheme.

Section 8.2 below recalls the identification of abduction as an argument form by scientists, most notably, Charles Darwin. Sections 8.3 to 8.6 provide a brief survey of the philosophical literature on inference to the best explanation. This literature is so large that the survey is necessarily brief and incomplete.[1] The survey yields the unhappy result that this

1 For another overview, see Igor Douven (2016). I also do not explore the literature that investigates the inference to the best explanation from a Bayesian perspective, such as Valeriano Iranzo (2008) and Leah Henderson (2014). The reason is that Bayesian analysis cannot be applied everywhere, as later chapters in this book will show. Thus, the Bayesian analysis has at best narrow applicability.

literature has done a poor job of developing inference to the best explanation as a general argument form. There are three problems.

First, the basic concepts invoked remain imprecisely defined. Worse, efforts to explicate these concepts trigger a death spiral of multiplying problems: clarifying one concept requires the introduction of several new ones that in turn require their own clarifications.

Second, the selection of illustrative examples is commonly poor. Examples in science are often just named or glossed hastily and claimed to support some favored conclusion. We shall see in the next chapter that a closer examination of canonical examples commonly returns conclusions at variance with the existing literature. Most importantly, explanation will be shown to play a minor role in them.

Third, there is a strong tendency to employ illustrative examples that involve human action. They are poor surrogates for the corresponding scientific examples. In the case of examples involving human action, it is obvious immediately that the favored explanation is correct and that the exploration of alternatives is, at best, a perfunctory exercise. There really are no credible, competing explanations for the origin of bootprints in freshly fallen snow. We might try to suppose that the snow just happened to settle into the shape of boot, complete with a boot's characteristic tread pattern. But the thought is too strained to bear serious consideration. Scientific examples are quite unlike this. It is far from obvious that the Big Bang is the unique, credible explanation of the cosmic background radiation. As we shall see in the next chapter, the real work in the examples involves establishing with some effort that no other explanation can likely succeed.

This unsatisfactory situation is resolved, I will argue in this chapter and the next, if we abandon the search for a single, unified formal account of these inferences. Instead, if we approach the examples materially, on a case-by-case basis, we then find that there is commonly a clear warrant for the inferences in background facts, as required by the material theory of induction. We also see some similarities in how these facts are deployed to provide the warrant, and it is these similarities that sustain the sense that inferences somehow belong together. The similarities, however, are not strong enough to support a formal schema, but just a loose resemblance. Most importantly, once we have found the warrant for the inferences in background facts, we have enough warrant; there is no longer any need to search fruitlessly within the very notion of explanation itself for some unifying, special constituent that confers inductive powers upon explanation.

What remains is to identify the loose similarities that connect the inferences commonly identified as abductive or as inference to the best explanation. Drawing on the inventory of examples in science in the next chapter, the similarities are summarized in Section 8.7. Abductions or inferences to the best explanation in actual science are carried out in two steps with some distinctive notion of explanation playing no role in either.

The first is a comparative step. The favored hypothesis or theory is shown to do better than one or more foils. We are to prefer—but not necessarily infer to—the better of them. We might call this "Preference for the better explanation." The way the favored hypothesis or theory does better turns out to be simple. While the preferred hypothesis or theory accommodates the evidence, the foil might just be contradicted by the evidence. Or the foil might require additional posits, which do not themselves have evidential grounding. This lack amounts to what I will call the incurring of an "evidential debt" not taken by the favored hypothesis or theory. It is then easy to see how the evidential judgments of this first step are supported by material facts, for the still elusive general notion of explanation plays no role. We prefer the theory that is not contradicted by the evidence, or the theory that accommodates the evidence without overt lacunae of support in its individual parts.

The second step is more fraught. We are to suppose that better is best, and that best is good enough to warrant commitment. Preference becomes commitment. This step is commonly grounded in a presumption that no other theory can do better than those explicitly considered. The presumption is so hard to justify that this second step is often left tacit and sometimes even omitted completely. For the step commonly relies merely on our human imaginative powers to sustain the conclusion that there is no better account just beyond our horizon. Kyle Stanford (2006) has effectively and powerfully described this problem of "unconceived alternatives."

Section 8.8 presents a conjecture on why inference to the best explanation rose in prominence historically as an argument form in the twentieth century. Section 8.9 offers a concluding comparison of the formal and material approaches to abduction.

8.2. Scientists Explain

What gives inference to the best explanation solid credentials in philosophy of science is that scientists themselves often advertise the explanatory prowess of their theories and suggest it provides support for their theories. Here are two prominent examples.

Upland geese, Darwin (1876, pp. 142–43) reported, rarely go near water, but they have the same webbed feet that are of great utility to aquatic birds. This curious fact, Darwin noted, is readily explained by natural selection as a residual from ancestral aquatic geese. It is poorly explained by the hypothesis of independent creation. Why create geese with this unnecessary feature? Darwin made observations like this the explicit driver of his argument in *On the Origin of Species*. He concluded the final chapter of his book with a defense of the argument form, not just in biology but in ordinary life and the other sciences:

> It can hardly be supposed that a false theory would explain, in so satisfactory a manner as does the theory of natural selection, the several large classes of facts above specified. It has recently been objected that this is an unsafe method of arguing; but it is a method used in judging of the common events of life, and has often been used by the greatest natural philosophers. The undulatory theory of light has thus been arrived at; and the belief in the revolution of the earth on its own axis was until lately supported by hardly any direct evidence. It is no valid objection that science as yet throws no light on the far higher problem of the essence or origin of life. Who can explain what is the essence of the attraction of gravity? No one now objects to following out the results consequent on this unknown element of attraction; notwithstanding that Leibnitz formerly accused Newton of introducing "occult qualities and miracles into philosophy." (1876, p. 421)

In late 1915, Albert Einstein's general theory of relativity was still a highly speculative theory, operating at a level of abstraction and mathematical complexity remote from the other physical theories of his time. He needed an evidential coup to secure the theory. It came in mid-November 1915, when Einstein discovered to his delight that his new theory predicted the anomalous motion of Mercury. In a paper entitled "Explanation of the Perihelion Motion of Mercury from the General Theory of Relativity," he wrote:

> In the present paper, I find an important confirmation of this most radical theory of relativity; that is, it turns out that the secular rotation of Mercury's orbit in the direction of the orbital motion, discovered by Leverrier, which amounts to about 45" in a century, is explained qualitatively and quantitatively, without having to posit any special hypothesis at all. (1915, p. 831)

This success was so striking that it is one of the most used illustrations in subsequent work in confirmation theory. We shall return to both examples below.

The two examples of Darwin and Einstein make at least a *prima facie* case that there is an interesting inductive argument form at hand that is somehow associated with a notion of explanation. One would expect that logicians and philosophers of science would be able to seize upon these clues and deliver a rigorous and logically tight account of the argument form. Alas, the brief survey below of the philosophical literature reveals one that is stalled in preliminary and inadequate sketches of the argument form. Worse, its prospects are limited at the outset by a near-universal aversion to real examples in sciences. Instead, the literature favors examples in which the best explanation involves some human action, which makes the examples quite unlike the corresponding inferences in real science. The sections that follow will elaborate on this grim assessment.

8.3. Peirce and Abductive Inference

The philosophical literature attributes the first explicit discussion of abductive inference to Charles Peirce. The much-quoted statement on this comes from Peirce's 1903 Harvard Lecture, "Pragmatism as the Logic of Induction":

> Long before I classed abduction as an inference it was recognized by logicians that the operation of adopting an explanatory hypothesis—which is just what abduction is—was subject to certain conditions. Namely, the hypothesis cannot be admitted, even as a hypothesis, unless it be supposed that it would account for the facts of some of them. The form of the inference, therefore, is this:
>
> The surprising fact, C, is observed;
>
> But if A were true, C would be a matter of course,
>
> Hence, there is reason to suspect that A is true.
>
> Thus, A cannot be abductively inferred, or if you prefer the expression, cannot be abductively conjectured until its entire content is already present in the premises, "If A were true, C would be a matter of course." (1932, p. 189)

What is curious is the myopia in crediting Peirce. For Darwin's *On the Origin of Species* was already a tour-de-force of abduction. The inference form is used throughout the book.[2] As we saw in the passage quoted from Darwin above, he was aware of the distinctive character of the argument form he was using and offered a defense of it as something used generally in common life and other great scientific discoveries.[3] What more can we ask? The inference form is identified explicitly at the same time as it is used repeatedly and powerfully in the canonical demonstration of one of science's greatest discoveries. In contrast, Peirce's development is labored.

---

2 In the final edition (Darwin 1876), the word "explain" appears 108 times and "explanation" 44 times.

3 What of Charles Lyell's *Principles of Geology*? It contains a template for Darwin's argument in *On the Origin of Species*, and Darwin studied it and drew inspiration from it. While we and, presumably, Darwin saw the argument form there, I will argue in the next chapter that, curiously, Lyell did not.

While it has the superficial appearance of a logical schema, key terms are not given precise definitions. Just what is meant by "surprising" and "a matter of course"?[4]

8.4. Harman's Inference to the Best Explanation

Peirce's treatment also conforms to the nineteenth-century tradition of combining inductive methods with discovery methods. John Stuart Mill's methods were as much a way of discovering the causes of some phenomena as they were of supporting them inductively. The same is true of Peirce's account of abduction. This procedural aspect is lost by the time of Gilbert Harman's (1965) paper "Inference to the Best Explanation," from which the now popular label derives. His account of the inference is as follows:

> In making this inference one infers, from the fact that a certain hypothesis would explain the evidence, to the truth of that hypothesis. In general, there will be several hypotheses which might explain the evidence, so one must be able to reject all such alternative hypotheses before one is warranted in making the inference. Thus one infers, from the premise that a given hypothesis would provide a "better" explanation for the evidence than would any other hypothesis, to the conclusion that the given hypothesis is true. (1965, p. 89)

The account remains remote from a serviceable formal schema. What "explain" might mean is not made clear, and "better" is presented in scare quotes. Harman concedes that formulating a more precise account is an open problem:

> There is, of course, a problem about how one is to judge that one hypothesis is sufficiently better than another hypothesis. Presumably such a judgment will be based on considerations such as which hypothesis is simpler, which is more

4 Peirce's work is littered with citations of Darwin. I have not ascertained whether any of these credit Darwin's priority. Certainly, the credit is not given prominently.

> plausible, which explains more, which is less *ad hoc*, and so forth. I do not wish to deny that there is a problem about explaining the exact nature of these considerations; I will not, however, say anything more about this problem. (p. 89)

The paper is short, a mere eight pages. It has no well-developed examples, but many are mentioned by brief allusion. The only example from science is in this one sentence: "When a scientist infers the existence of atoms and subatomic particles, he is inferring the truth of an explanation for various data which he wishes to account for" (p. 89). Otherwise, all the examples mentioned pertain to human action:

> "a detective … decides that it *must* have been the butler" (p. 89)
>
> "we infer that a witness is telling the truth" (p. 89)
>
> "we infer from a person's behavior to some fact about his mental experience" (p. 89)
>
> "I read … that Stuart Hampshire is to read a paper at Princeton tonight" (p. 92)
>
> "obtaining knowledge from an authority" (p. 93)
>
> "knowledge of mental experience gained from observing behavior" (p. 93)

8.5. Thagard's Criteria

Paul Thagard's (1977) analysis is an exception to my dismal assessment of the philosophical literature on inference to the best explanation. It excels both in the range of real examples from science and in its dedication to clarifying just how inference to the best explanation works.

The range of examples deployed to illustrate and support the paper's claims is impressive. It includes

> Darwin's long argument in his *On the Origin of Species*;

> Lavoisier's case for the oxygen theory of combustion;
>
> The wave theory of light, as developed by Huygens in the seventeenth century; and by Young and Fresnel in the nineteenth century;
>
> Newton's explanation of the motion of planets and satellites;
>
> Halley's Newtonian prediction of the return a comet;
>
> Young's account of di-polarization;
>
> Fresnel's account of polarization through transverse waves;
>
> General relativity's treatment of the anomalous perihelion motion of Mercury, the gravitational bending of light, and the gravitational red shift of light;
>
> Quantum mechanics and its success with atomic spectra, magnetism, the solid state of matter, the photoelectric, and the Compton effect.

The list is so long as to be too ambitious for a single paper. The accounts given are brief and often amount to mere mentions. However, the laudable principle sustained is that Thagard's account is responsible to these real examples from science.

Thagard also recognizes that the inference form is in urgent need of elaboration and clarification; and he takes up the project. From the perspective of the material theory of induction, the project is ill-fated. For the arguments labeled "abductive" or "inference to the best explanation" form at best a loose unity. The individual arguments differ so much in their details that they can be grouped together only as long as the argument form is imperfectly specified. This means that any applicable notion of explanation must be kept vague enough so that it can be applied everywhere. Efforts to remove the vagueness in the notions of "explanation" and "better explanation" will require further notions and possibly many of them, if the existing range of individual arguments is to be accommodated.

Matters will get worse the further these efforts go, for each solution will generate new problems. An explosion of difficulties will be triggered. Yet the results of these efforts can never be secure. All it takes to overturn them is a new, troublesome instance of an abductive inference.

This fate befalls Thagard's project, as we shall now see. The project begins with a brief definition:

> To put it briefly, inference to the best explanation consists in accepting a hypothesis on the grounds that it provides a better explanation of the evidence than is provided by alternative hypotheses. We *argue* for a hypothesis or theory by arguing that it is the best explanation of the evidence. (p. 77)

Here, the key term "explanation" is left undefined. This serious oversight persists throughout the paper, until on the concluding page (p. 92) we find a begrudging admission that "Explanation is a pragmatic notion." Instead of defining the term, explicit analytic efforts are devoted to clarifying when one explanation—whatever the term may mean—is better than another. The clarification depends on three criteria: consilience, simplicity, and analogy. The difficulty, however, is that evaluations based on these criteria may pull in different directions. And so we see the multiplication of problems. The project has now replaced the problem of clarifying one notion with the problem of clarifying three notions.

The notion of consilience, drawn from the work of William Whewell,[5] is given the following gloss: "one theory is *more* consilient than another if it explains more classes of facts that the other does" (p. 79; emphasis in original). The problem then is to specify how we are to count classes of facts so that "more" has an unambiguous meaning. Of course, there is no simple solution, and the analysis stalls with inevitable difficulty: "In inferring the best explanation, what matters is not the sheer number of facts explained, but the variety, and variety is not a notion for which we can

5 Here, Thagard draws on his earlier (1977) where he identified Darwin's use of Whewell's notion.

expect a neat formal characterization" (p. 83). The notion of consilience then further mutates into a static and a dynamic notion.

The threat to a cogent notion of consilience is that any account can be made to embrace more facts if we are willing to make it more complicated. This is where Thagard's second criterion, simplicity, plays a role. Simplicity, he proposes, is measured by size and nature of auxiliary hypotheses needed by some theory to explain the facts. The fewer there are of these auxiliaries, the simpler and better the explanation. Needless to say, trying to give a more precise account of simplicity leads to further problems and the wry conclusion: "As has often been remarked, simplicity is very complex" (p. 88).

Finally, the third criterion, analogy, enters apparently through no pressing conceptual need but simply because the examples driving the analysis use it. Analogies, we are told, function to improve the explanations used. We have seen in Chapter 4 that efforts to explicate analogical inference face a similar difficulty of multiplying problems. Thagard's analysis only begins to probe this difficulty. After abandoning a classic definition of analogy, Thagard offers an alternative. If, for some entity *A*, property *S* explains why it has properties *P*, *Q*, and *R*, then we can project to other cases. That is, if another entity *B* has properties *P*, *Q*, and *R*, then we may "conclude that *B* has *S* is a promising explanation of why *B* has *P*, *Q*, and *R*" (p. 90). Of course, this characterization is only as good as the characterization of the notion of explanation, for which essentially nothing is offered.

At the end of the paper we are left with the unresolved problem of how to trade off the criteria:

> Consilience and simplicity militate against each other, since making a theory more consilient can render the theory less simple, if extra hypotheses are needed to explain the additional facts. The criterion of analogy may be at odds with both consilience and simplicity, if a radically new kind of theory is needed to account simply for all the phenomena. (p. 92)

Leaving the problem unsolved means that we cannot unambiguously apply the rule of inference to the best explanation. Far from recovering a universally applicable rule of inductive inference, we have failed even to arrive at an unambiguous rule.

The material theory of induction was introduced in response to the pervasiveness in formal accounts of inductive inference of difficulties like these. Seeing the burden of multiplying problems drag down his account, I truly sympathize with Thagard's concluding lament, "Application of the criteria of consilience, simplicity, and analogy is a very complicated matter" (p. 92).

8.6. Lipton's Monograph

Peter Lipton was the most prominent of recent proponents of inference to the best explanation, and his 2004 monograph has become a canonical source. His two works (2000, 2004) provide no formula or schema that would improve on those of Darwin, Peirce, Harman, or Thagard. But his detailed elaboration maps out just how open the problem set aside by Harman and Thagard remains. We have no notion of explanation or better explanation sufficiently well developed to convert what Lipton (2004) repeatedly calls the "slogan" of inference to the best explanation into formal schema.

Take the notion of explanation. Efforts to clarify it lead to the same multiplying problems we saw in Thagard's project. This is due to the fact that there are multiple competing accounts of explanation. Some of these accounts are surveyed in Lipton (2004, chap. 2). To explain a phenomenon might mean to subsume it under a covering law; or to display the factors that increase its probability; or to display the causes that bring it about. Again, an explanation may unify many phenomena, hitherto thought disparate. Each of these notions captures a sense of explanation applicable in some circumstance. A fully elaborated schema of abduction would then need to accommodate all of these further notions. What is a law as opposed to a general proposition? What is the origin of the probabilities? Just what do we mean by "cause"? How do we distinguish unification from mere conjunction? Needless to say, each of these is an unfinished project in its own right.

Prudently, Lipton does not take on the challenge of finding a schema that would embrace all of these senses of explanation. Rather, he develops the causal model of explanation, perhaps because it fits best with his favorite, elaborated example of Ignaz Semmelweis and his discovery of the cause of childbed fever (2004, chap. 3). However, Lipton concedes (2004. p. 3) that he can provide no analysis of the notion of causation and uses it as an unexplicated term.[6]

We press on. How are we to judge which explanation is better? We could, Lipton argues, adopt the most likely explanation. However, this would reduce abduction to a circularity: the most likely explanation would then be the one most likely to be true. Lipton introduces a distinct characteristic to replace "likeliest": we should infer to the "loveliest" explanation (p. 59, p. 121). According to Lipton, this would then guide us to the likeliest explanation. What makes an explanation "lovelier" is, loosely, that it provides the most understanding (p. 59). This derives in turn from what Lipton identifies as "explanatory virtues," which include "mechanism, precision, scope, simplicity, fertility or fruitfulness, and fit with background belief" (p. 122) as well as "unification" (p. 139). Once again, the singular problem of determining the lovelier explanation has multiplied into many, unresolved problems. We are quite far from any account of the virtues that would allow them a place in a formal schema of inductive inference.

Lipton does introduce what are, for present purposes, two important extensions to the notion of inference to the best explanation. The first is the recognition that an explanation must rise to some minimal level of success before we are authorized to infer to it. As a result, he is willing to relabel inference to the best explanation as "Inference to the Best Explanation if the Best is Sufficiently Good" (2004, p. 154) and "inference to the best of the available explanations, when the best one is sufficiently good" (2000, p. 187). Second, Lipton introduces a contrastive notion. It is restricted to causal explanation and its key assertion is the following:

> Difference Condition: to explain why P rather than Q, we must cite a causal difference between P and not-Q,

6 For a pessimistic view of any general account of causation that might serve his purposes, see Norton (2003).

> consisting of a cause of P and the absence of a corresponding event in the case of not-Q. (p. 42)

I will argue below the importance in all cases of distinguishing comparative judgments of which are better explanations from the absolute judgment that some explanation is best.

Unfortunately, Lipton's treatment of examples is superficial, with one exception. Many examples are named, but mostly there is little or no explanation of the content of the example. This is a serious problem, since we shall see in the next chapter that closer examination of canonical examples leads to conclusions other than those drawn by Lipton. Equally seriously, many of the examples are drawn from ordinary human situations. These are sufficiently unlike important examples in science that reliance on them is dangerous if one's goal is to understand inferences in science.

I have prepared a compendium of examples from science as a way of assessing their distribution over types:[7]

> A drought may explain a poor crop.
>
> The Big Bang explains background radiation.
>
> Stress, fatigue, etc. explains bridge collapse.
>
> Velocity of recession explains galactic red shift.
>
> Kinetic theory of gases explains thermal phenomena.
>
> Natural selection explains the traits of plants.
>
> Electronic theory explains current flow.
>
> Echolocation explains bat navigation.
>
> The same side of the moon faces us.
>
> Why the planets move in ellipses.

7 This list is a mix of quotes and paraphrasing. No page numbers are given since many examples are repeated over many pages.

Why leaves turn yellow in November.

Prior syphilis explains why someone contracted paresis.

Freudian wish-fulfillment explains a slip.

A field explains the deflection of a particle.

Chomsky infers language structure.

Lightning and thunder.

Perturbations in the orbit of Uranus explained by Neptune.

Mendeleyev predicts new elements.

Song employed by sparrows.

A double-blind test of a drug's efficacy.

Gregor Mendel's peas.

Millikan's oil drops.

Finally, there is an example from relativity theory:

> We are more impressed by the fact the special [*sic*] theory of relativity was used to predict [*sic*] the shift in the perihelion of Mercury than we would have been if we knew that the theory was constructed in order to account for that effect. (p. 172)

Others examples are essentially dependent on human actions and thus unlike real examples in science:

Why you didn't come to the party (headache).

Peculiar tracks in the snow in front of my house (snow-shoes).

A magician intuits the numbers I am thinking of.

I see a supposedly vacationing friend at the supermarket.

The rattle in the car.

Praise and punishment by Israeli airforce instructors.

Why a three-year-old threw his food on the floor.

Why Lipton went to see *Jumpers* rather than *Candide*.

Why Able rather than Baker got the philosophy job.

"Why did you order eggplant?"

Why Kate rather than Frank won the prize.

Why my horse won rather than yours.

Why Lewis went to Monash rather than Oxford in 1979.

A door opening triggers a bomb.

Why all men in the restaurant are wearing paisley ties.

The butler did it.

The patient has measles.

My front door has been forced open.

Why is my refrigerator not running.

Whether my car will start tomorrow.

Sherlock Holmes' dog that did not bark.

Movement of the mouse causes the movement of the cursor.

A crossword puzzle.

Successful navigation by means of a map.

Still others are intermediate between the two types of cases:

> Sticks in a bunch thrown in the air more likely horizontal.
>
> A spark causes a fire, but oxygen does not.
>
> Why mercury rises in a thermometer.
>
> Why people feel heat more when humidity is high.
>
> Kuhn infers normal science is governed by exemplars.
>
> Opium puts people to sleep.
>
> Data from flight recorder of crashed plane.
>
> Kahneman and Tversky's "Linda the Bank Teller"; people told of a taxi involved in a hit-and-run accident.
>
> A sympathetic powder that can cure wounds at a distance.
>
> Methods of predicting future performance on the London Metal Exchange.
>
> Persistence forecasting of the weather.

A scan of the lists indicates that the potentially misleading human examples have as much presence as the scientific examples.

Finally, we have one extended example in Lipton's text. It is the identification of the cause of childbed fever by Semmelweis in the 1840s in a Vienna maternity hospital. The primary narrative spans seventeen pages (pp. 74–90). The example is well known in philosophy of science through its inclusion in Hempel's (1965; chap. 2) widely read and highly accessible *Philosophy of Natural Science*. In brief, the maternity hospital had two divisions and, alarmingly, the death rate from childbed fever was markedly higher in one than in the other. Over a period of several years, Semmelweis checked all manner of differences between the two divisions in search of the cause. None was found until Semmelweis finally realized that the doctors and medical students in the higher mortality division only were delivering babies after performing autopsies elsewhere. He guessed that cadaveric material on the doctors' hands was the cause of the childbed fever. His guess was confirmed when he required the doctors to disinfect their hands with chloride of lime before delivering in the maternity ward, whereupon the differential death rate disappeared.

The case is a classic example of dedicated scientific detective work and the powerful use of evidence. However, as a case study intended to display the merits of inference to the best explanation specifically, the case study is a failure. For that, what is needed is a case study in which the evidential relations depend quite specifically on the distinctive merits peculiar to inference to the best explanation. It would do so in a way that makes it unlikely that any other account of inductive inference could do as well. This is not that case study, for explanation plays little if any role in the analysis. Rather, Semmelweis' investigation and analysis is a near perfect example of the application of Mill's methods.

The clearest application comes in the identification of the cause. Mill's method of difference applies when we have two instances, one in which the phenomenon of interest occurs and one in which it does not. If they differ only in one circumstance, that is the cause. This is precisely the case of Semmelweis. In the key experiment, the only change associated with the drop in mortality was that the doctors were disinfecting their hands from cadaveric material with chloride of lime. The eliminated cadaveric material was the cause.

We can see in Semmelweis' own narrative how his analysis was driven by just such considerations.

> As mentioned, the commissions identified various endemic factors as causes of the greater mortality rate in the first clinic. Accordingly, various measures were instituted, but none brought the mortality rate within that of the second clinic. Thus one could infer that the factors identified by the commissions were not causally responsible for the greater mortality in the first clinic. I assumed that the cause of the greater mortality rate was cadaverous particles adhering to the hands of examining obstetricians. I removed this cause by chlorine washings.

> Consequently, mortality in the first clinic fell below that of the second. I therefore concluded that cadaverous matter adhering to the hands of the physicians was, in reality, the cause of the increased mortality rate in the first clinic. Since the chlorine washings were instituted with such dramatic success, not even the smallest additional changes in the procedures of the first clinic were adopted to which the decline in mortality could be even partially attributed. (2008, pp. 7–8)

Clearly all that was at issue in Semmelweis' analysis was to find the difference that made a difference and identify it as the cause. Of course, one could embed Semmelweis' analysis in a larger narrative replete with discussion of how the cadaveric material explained the childbed fever, as Lipton did. However, this is unnecessary; Semmelweis' own analysis makes no essential use of explanatory notions.

The brief remarks above already indicate how well Semmelweis' methodology is captured by Mill's methods. Scholl (2013) has given a more thorough analysis of Semmelweis' methodology and finds extensive use of Mill's methods, including Mill's method of agreement and concomitant variation. Scholl (2015) argued for the failure of Lipton's attempts to impugn the understanding of Semmelweis' analysis as an application of Mill's methods.

8.7. Inference to the Best Explanation without Explanation: Two-Step Reconstruction

What do inferences commonly labeled abductive or inference to the best explain have in common? The examples of the next chapter are loosely bound together by a simple two-step scheme. The scheme does not require a sophisticated notion of explanation. Mere accommodation is all that is needed. Here, we may conjecture that Lipton was not just unlucky in choosing as his major example a case in which explanation proved to play no special role. While the Semmelweis example was an especially poor choice, it also reflects a problem that will be repeated in every example we will examine in the next chapter: the more closely we look at an example, the less important is the role of explanation as a distinct notion of philosophers.

Step 1. Preference for the better explanation.

What the examples developed in the next chapter have in common is that they all involve a comparison of a favored theory or hypothesis with one or more foils. The favored hypothesis is adequate to the evidence, most commonly in the sense that it deductively entails the evidence. The foils—that is, the alternatives—are judged inadequate in one of two ways:

1. Contradiction: the evidence at hand may directly contradict the alternative; or the evidence supplemented by specific background facts may contradict the alternative.
2. Evidential debt: to accept the alternative requires us to accept further assumptions for which we have no evidence.

The essential point is that the favoring invokes no explanatory notions, unless one accepts that the notions invoked here are a full—if thin—account of explanation. If the disfavoring consists of the alternative facing contradictions, it is simple logic. We prefer the logically consistent over the inconsistent. If the disfavoring is driven by evidential debt, a simple test will show that the presence of the evidential debt is fully responsible for the disfavoring. In the next chapter, in the case of the explanation of the anomalous perihelion motion of Mercury, we will see that if the evidential debt could have been discharged, a fully admissible hypothesis would have resulted.

Step 2. From comparative to absolute: better is best.

This first step just gives a reason to prefer one hypothesis or theory over another. This is not enough if we are to commit to the preferred hypothesis as the inference scheme requires. We need more and that comes from an assumption that no other hypothesis or theory can do better.

Under any account of inference to the best explanation—material or otherwise—this is the fragile step. Whereas the comparative judgments of Step 1 are explicit in the scientists' narrative, this absolute judgment is not.

How can this step be warranted? The surest case arises when background assumptions assure us that the hypotheses or theories we have considered are exhaustive. Then, there are no more credible candidates left, so the best of those considered must also be the best. These background assumptions are the assumptions that warrant the inference.

The most difficult case is the most common. It is when the inference from better to best is made, even though scientists have no clear grasp of the full range of hypotheses or theories possible. Then, at worst, the inference is unwarranted. Or, more charitably, there may be a tacit meta-argument at work. The argument works not at the level of theories but of theorizers. The assumption is that the theorizers are sufficiently inventive and perspicacious to have surveyed the full range of hypotheses or theories applicable, and that they have considered the most credible. Once again, the best of those considered then must be the unqualified best. These background assumptions over the power of theorizers warrants the inference from better to best.

8.8. Why Inference to the Best Explanation?

Given that that the full two-step inference faces such difficulties, why has it come to prominence over the last century or so? It is because it has helped theorists solve a vexing evidential problem. In earlier theorizing, theorists were often in the happy position that they could infer fairly directly from evidence to a theory. Newton, for example, could infer quickly from Kepler's third law of planetary motion to an inverse square law of gravitational attraction for the planets. He spoke confidently of *deductions* from the phenomena. While that now sounds extravagant, Newton's inferences from the phenomena employed background assumptions that made them deductive.[8] Even as late as 1929, Edwin Hubble could arrive at his Hubble law for the speed of recession of the galaxies merely by fitting a straight line to a plot of velocity-distance data for a subset of his data.

---

8 See, for example, Harper (2002).

By Hubble's time, the happy days of easily supported theories were passing. This was especially clear with Einstein's general theory of relativity. It was a theory of such enormous complexity that no similar inference from phenomena to theory was possible. The gap was just too great. While the problem is not as stark, others faced similar problems. Darwin could not infer directly from his mass of evidence in natural history to natural selection. The relationship between his evidence and theory was just too complicated.

How can theorists close the gap? Perhaps a direct inference cannot be made from evidence to a theory. But a theorist may sense that a theory fits the evidence so well that it must have something right. This sense can be viscerally strong and communicated fairly easily by recounting the details of the example. It is often expressed compactly by the claim that the theory explains the evidence. The task remaining for the inductive logician, however, is to take the loosely articulated but viscerally powerful sense and translate it into a transparent analysis of just how the evidence supports the theory. The project of translating the evidence into a precise, general, formal schema remains unfinished and, if the arguments of this chapter and book are upheld, will remain so. However, if a material warrant is sought on a case-by-case basis, a warrant can be found for each case in background facts.

8.9. Conclusion

Does inference to the best explanation provide a serviceable, general rule of inductive inference? Its failure to do so can be shown by a simple question: If we know that some hypothesis gives the best explanation of the evidence, should we infer to it? The answer, of course, is that without further details we simply cannot say. When we look more closely at the details, the strength of the inference becomes clearer. Since the strength of the inductive support can only be assessed, in the end, by looking at the details of the case at hand, we can see that inference to the best explanation is not a self-contained rule of inductive inference. It is at best a loose guide in urgent need of development. We have seen in this chapter that prospects for development are meager. Efforts to develop the rule lead to a multiplication of problems. Each solution presents more problems than it solves. The more we try to clarify the general argument form, the less clear it becomes.

That inference to the best explanation should be troubled in just this way is quite expected according to the material theory of induction. For, accordingly, there can be no universal formal rule covering all cases. At best, inferences grouped under the label "inference to the best explanation" form a loose unity that dissolves once we look more closely at each inference. The most precise assessment of the inductive strength of any particular argument comes only when we fully take into account the background facts that warrant the inference. In the final analysis, all that inference to the best explanation provides is an indication of a loose similarity with other arguments and nothing more.

REFERENCES

Darwin, Charles R. 1876. *On the Origin of Species by Means of Natural Selection, or The Preservation of Favoured Races in the Struggle for Life*. 6th ed. London: John Murray.

Day, Timothy and Harold Kincaid. 1994. "Putting Inference to the Best Explanation in Its Place," *Synthese* 98: pp. 271–95.

Douven, Igor. 2016. "Abduction." *Stanford Encyclopedia of Philosophy*. Winter 2016 Edition. Edited by Edward N. Zalta, https://plato.stanford.edu/archives/win2016/entries/abduction/.

Einstein, Albert. 1915. "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie." *Preussische Akademie der Wissenschaften, Sitzungsberichte* (part 2), pp. 831–39.

Harman, Gilbert H. 1965. "The Inference to the Best Explanation." *The Philosophical Review* 74: pp. 88–95.

Hempel, Carl G. 1966. *Philosophy of Natural Science*. Upper Saddle River, NJ: Prentice Hall.

Henderson, Leah. 2014. "Bayesianism and Inference to the Best Explanation." *British Journal for the Philosophy of Science* 65: pp. 687–715.

Hubble, Edwin. 1929. "A Relation Between Distance and Radial Velocity Among Extra-Galactic Nebulae." *Proceedings of the National Academy of Science* 15: pp. 168–73.

Iranzo, Valeriano (2008) "Bayesianism and Inference to the Best Explanation," *Theoria* 61: pp. 89–106.

Khalifa, Kareem, Jared Millson, and Mark Risjord. 2017. "Inference to the Best Explanation: Fundamentalism's Failures." In *Best Explanations: New Essays on Inference to the Best Explanation*, edited by K. McCain and T. Poston, pp. 80–96. Oxford: Oxford University Press.

Lipton, Peter. 2000. "Inference to the Best Explanation." In *A Companion to the Philosophy of Science*, edited by W. H. Newton-Smith, pp. 184–93. Malden, MA: Blackwell.

———. 2004. *Inference to the Best Explanation*. 2nd ed. London: Routledge.

Norton, John D. 2003. "Causation as Folk Science." *Philosophers' Imprint* 3(4), http://www.philosophersimprint.org/003004/. Reprinted in *Causation, Physics and the Constitution of Reality*, edited by H. Price and R. Corry, pp. 11–44. Oxford: Oxford University Press.

Peirce, Charles S. 1932. *Collected Papers of Charles Sanders Peirce*. Vol. 5, *Pragmatism and Pragmaticism*. Cambridge, MA: Harvard University Press.

Roche, William and Elliott Sober. 2013. "Explanatoriness is Evidentially Irrelevant, or Inference to the Best Explanation Meets Bayesian Confirmation Theory." *Analysis* 73: pp. 659–68.

Scholl, Raphael. 2013. "Causal Inference, Mechanisms, and the Semmelweis Case." *Studies in History and Philosophy of Science* 44: pp. 66–76.

———. 2015. "Inference to the Best Explanation in the Catch-22: How Much Autonomy for Mill's Method of Difference?" *European Journal for the Philosophy of Science* 5: pp. 89–110.

Semmelweis, Ignaz. 2008. "The Etiology, Concept and Prophylaxis of Childbed Fever: Classics in Social Medicine." *Social Medicine* 3: pp. 4–12.

Sober, Elliott. 1980. *Reconstructing the Pas*. Cambridge, MA: MIT Press

Stanford, Kyle. 2006. Exceeding Our Grasp: Science, History and the Problem of Unconceived Alternatives. Oxford: Oxford University Press.

Thagard, Paul R. 1977. "Discussion: Darwin and Whewell." *Studies in History and Philosophy of Science* 8: pp. 353–56.

———. 1978. "The Best Explanation: Criteria for Theory Choice." *Journal of Philosophy* 75: pp. 76–92.

Van Fraassen, Bas. 1977. "The Pragmatics of Explanation." *American Philosophical Quarterly* 14: pp. 143–50.

———. 1980. *The Scientific Image*. Oxford: Clarendon.