CHAPTER FOUR

Underdeterminism (I)

To understand how it is even possible to pursue metaphysics, it is necessary to begin by understanding the limits of what experience can tell us of the world,* and to understand how experiential (scientific) knowledge and metaphysics enjoy a symbiotic relationship.

4.1 The interconnectedness of science and metaphysics

In the first half of the twentieth century, there was a philosophical movement (c. 1920-45) which was openly hostile to metaphysics. The disaffection of the Logical Positivists (also known as Logical Empiricists) sprang principally from their antipathy to the highly speculative metaphysics of a number of nineteenth- and early-twentieth-century philosophers. Carnap mentions by name Fichte, Schelling, Hegel, Bergson, and Heidegger ([44], 80). He illustrates (1932), with a quotation from Heidegger, the sort of metaphysics which he is intent to eliminate.

> What is to be investigated is being only and – *nothing* else; being alone and further – *nothing;* solely being, and beyond being – *nothing. What about this Nothing? … Does the Nothing exist only because the Not,* i.e. *the Negation, exists?* Or is it the other way around? *Does Negation and the Not exist only because the Nothing exists? …* We assert: *the Nothing is prior to the Not and the Negation. …* Where do we seek the Nothing? How do we find the Nothing. … We know the Nothing. … *Anxiety reveals the Nothing. …* That for which and because of which we were anxious, was 'really' – nothing. Indeed: the Nothing itself – as such – was present. … *What about this Nothing? – The Nothing itself nothings.* ([44], 69; italics in the original; translation by Arthur Pap[1])

---

1. The translating of Heidegger's writings into English has always been

A.J. Ayer continued the attack (1936) and used for his own example of unacceptable metaphysics ([16], 36) a slightly edited version of the last sentence of this passage from F.H. Bradley:

> … pure spirit is not realized except in the Absolute. It can never appear as such and with its full character in the scale of existence. Perfection and individuality belong only to that Whole in which all degrees alike are at once present and absorbed. This one Reality of existence can, as such, nowhere exist among phenomena. And it enters into, but is itself incapable of, evolution and progress. ([33], 442)

Modern scholarship is not nearly so unsympathetic to these initially obscure-sounding pronouncements. Many modern writers claim to find in them intelligibility and significance not at all apparent to Carnap, to Ayer, or to their fellow Logical Positivists. Be that as it may, there is an undeniable difference between the style, the vocabulary, and the accessibility, on the one hand, of the metaphysics of Heidegger and Bradley, and on the other, of that of Strawson, for example, of whom we spoke briefly at the end of chapter 2. Suppose we compare the selection from Heidegger with one drawn from the first chapter of Strawson's *Individuals*. Both Heidegger and Strawson, we note, are endeavoring to explain what metaphysics is.

> We think of the world as containing particular things some of which are independent of ourselves; we think of the world's history as made up of particular episodes in which we may or may not have a part; and we think of these particular things and events as included in the topics of our common discourse, as things about which we can talk to each other. (Strawson [200], 15)

Metaphysics – at least as written by some philosophers – we see *can be* perfectly straightforward and readily comprehended.

With examples drawn from the most speculative extremes of meta-

---

problematic. What Pap translates as "being", Hull and Crick translate as "what-is" ([88], 358); and where Pap coins the verb "nothings" and translates Heidegger as saying "The Nothing itself nothings", Hull and Crick say "Nothing 'nihilates' of itself" ([88], 369).

physics, the Logical Positivists embarked on a program to try to devise linguistic and logical tests by which they could separate 'scientifically meaningful' statements from what they called "pseudo-statements". Some of them were so incautious, even, as to call all of metaphysics, not just that which gave rise in the first instance to their disapprobation, "literal nonsense".

Nowadays metaphysics again needs no apology. The Logical Positivists' attack on metaphysics was relatively short-lived. For a variety of technical reasons, their program to create what they called "a criterion of empirical significance" was to prove impossible to carry through. In due course, the Positivists came to realize the futility of their enterprise and abandoned it. Several philosophers (see e.g. Hempel [91]), including some of the Positivists themselves, carefully chronicled the successive attempts Positivists made along the way. Many explanations have been offered as to why the program was destined, ultimately, to fail.

The most common explanation for the failure is that there is no particular identifying feature of metaphysical statements. If a scientific theory happens to contain 'metaphysical' statements, then those statements cannot be distinguished from the other 'non-metaphysical' components of that theory. The source of the problem (from the Positivists' point of view) is that scientific theories are collections of diverse statements, and that sets of statements can only be tested *altogether*, i.e. one cannot, as a matter of logic, draw from a scientific theory testable implications whose truth or falsity redounds to the truth or falsity of individual members of the set (see, for example, Hempel [91], 129). In short, according to this explanation, there seems to be no way, either logically or linguistically, to isolate the 'metaphysical' components of a scientific theory from its 'non-metaphysical' components.

Such an explanation is, I think, incomplete. The problem lies deeper. The Positivists' program was doomed, not just because it is impossible to isolate the metaphysical components of a scientific theory, but rather, and more importantly, because metaphysical components are *essential* to any reasonable scientific theory. Scientific and metaphysical theorizing go hand in hand; it is impossible to do either one without the other. Science is impossible without *some* metaphysical presuppositions; metaphysics is impossible, or at the very least sterile in the extreme, unless informed by science (experience). Science and metaphysics are one of a kind; the difference is merely one of degree. The most fundamental presuppositions of science, those furthest removed from ‘observational’ data, tend to be regarded as ‘metaphysical’. It is these which change most slowly, which are so much a part of working scientists’ conception of the world that scientists scarcely attend to them in day-to-day work, spending most of their energies instead on that part of science which lends itself most easily to confirmation, disconfirmation, or revision in light of observation and experiment.

There is no question of logical or temporal priority in the interplay between our scientific and our metaphysical beliefs. Together they form a world-view. This world-view is in constant flux. As science progresses, our metaphysical views gradually change; and as metaphysical views change, adjustments are made within our science.

4.2 Case study: Francis Bacon’s account of induction

“What is heat?” The question is deceptively simple. One is tempted to think that its answer ought to be a relatively trivial matter to settle. After all, heat is not an unfamiliar, rare commodity. We encounter it, in varying degrees, throughout all our waking hours: in stoves, hot-water taps, light bulbs, automobile engines, noxious cigarettes, active muscles, etc. Surely all a scientist has to do, we might be tempted to believe, is to examine instances of many such cases and it will be quickly obvious what heat *is*. Francis Bacon (1561-1626), for one, explicitly said exactly this in *The New Organon* of 1620:

> … a nature [phenomenon] being given, we must first of all have a muster or presentation … of all known instances which agree in the same nature, though in substances most unlike. And such collection must be … without premature speculation … For example, let the investigation be into … heat. …
>
> 1. The rays of the sun, especially in summer and at noon.
>
> …
>
> 3. Fiery meteors.
> 4. Burning thunderbolts.
> 5. Eruptions of flames from the cavities of mountains.
> 6. All flames.
>
> …
>
> 9. Liquids boiling or heated.
>
> …
>
> 16. All bodies rubbed violently, as stone, wood, cloth, etc.

insomuch that poles and axles of wheels sometimes catch fire …

…

25. Aromatic and hot herbs … although not warm to the hand …, yet to the tongue and palate, being a little masticated, they feel hot and burning.

…

28. Other instances. ([20], 130-2)

Bacon thinks, however, that humankind lacks the requisite intelligence to infer the nature of heat from a table consisting solely of positive instances. The difficulty stems from the fact that, presented with any collection of items however much seemingly initially unlike one another, we can – with a little ingenuity – find not just one but several common features. A rug, a chair, and a tea bag – to invent just one example – may share any number of features in common: they may all be the same color; they may all be imported; they may all, to some degree, be manufactured; they may all be purchased goods; they may all be flammable; and so on. Drawing a single, correct, inference from a list of positive instances, Bacon thus believes, would be possible only by God and the angels, and perhaps by other higher intelligences (p. 151). Humankind, lacking the special faculties of divine intelligence, can reduce the number of detected commonalities only by supplementing the list of positive instances with lists of negative instances. Thus, for example, where the rays of the sun had been the first item in his own list of positive instances, he contrasts this with "the rays of the moon and stars and comets [which] are not found hot to the touch" (133), and similarly, for each of the other twenty-one specific items in his original list. And finally, he produces yet a third list, this time of some forty-one items, discussing a variety of instances in which heat comes in various degrees. For example, the twenty-fifth item in this third list reads: "Some ignited bodies are found to be much hotter than some flames. Ignited iron, for instance, is much hotter and more consuming than flame of spirit of wine [alcohol]" (147).

These methods of Bacon were to have a profound influence on philosophy. Two centuries later, we find them little changed, repeated in J.S. Mill's *Logic* (1843) as the Method of Agreement, the Method of Difference, and the Method of Concomitant Variations.

Having gathered his data, and having rejected a great number of hypotheses, e.g. the texture of materials, light or brightness, and rarity (pp. 154-5), Bacon states his conclusion:

> From a survey of the instances, all and each, of which the nature of heat is a particular case, [heat] appears to be motion. (156)

In spite of its modern sound, it is clear that Bacon's notion of the nature of heat is very unlike that of modern science. His subsequent comments reveal that he has not made the modern distinction between heat and temperature. Modern science tells us, for example, that two equal masses of water and iron at the same temperature contain different amounts of heat. It takes 8.4 times as much heat, for example, to raise 1 kg of water from 40°C to 41°C than it does to raise 1 kg of iron from 40°C to 41°C.[2] Nor does he conceive of heat as a quantity of fixed amount, i.e. he has no inkling of the law of the conservation of energy.

> … when heat is produced by the approach of a hot body, this … depends … on the nature of assimilation or self-multiplication. (157)

> … assimilation multiplies and transforms bodies and substances. … Heat does not diffuse itself, in heating a body, by communication of the original heat but simply by exciting the parts of the body to that motion which is the form of heat. (242)

My purpose is not, however, to fault Bacon's conclusions. My concern lies with his discussion of how he supposedly arrived at those conclusions.

At the most crucial point in all of this, at the juncture between having completed his review of his data and drawing his conclusions, Bacon offers no account whatsoever of how he proceeded to move from the former to the latter. Instead of an explanation, we find simply a label: Bacon calls the process of moving from data to conclusion an "induction" (130). Earlier, he had spent several pages (18-23) explaining that this was to be regarded as a new kind of induction, a 'legitimate' sort of induction.

---

2. This is to say, the *specific heat* of water is 8.4 times that of iron.

> … what the sciences stand in need of is a form of induction which shall analyze experience and take it to pieces, and by a due process of exclusion and rejection lead to an inevitable conclusion. … The testimony and information of the sense [i.e. of sensory perception] has reference always to man, not the universe; and it is a great error to assert that the sense is a measure of things. … The mind, when it receives impressions of objects through the sense, cannot be trusted to report them truly, but in forming its notions mixes up its own nature with the nature of things. … The intellect is not qualified to judge except by means of induction, and induction in its legitimate form. (20-3)

For Bacon, human senses can, and often do, deceive. The human intellect, either alone or in concert with the senses, is inadequate to the task of finding the route to Nature's "first principles … [which] lie at the heart and marrow of things" (20). The only way, according to Bacon, to proceed is by induction. And he thought, too, not only that his induction could generate new 'notions' but that it provided the only means to do so: "induction must be used … in the formation of notions" (99).

But while Bacon is expansive in his praise of induction, he is strangely silent in explaining precisely what it is supposed to be or how it works. He is completely silent about *how* induction might get us from observational data to explanatory hypothesis, or how one might learn the technique or teach it to someone else. There are no rules stated; there are not even any hints given. This seemingly minor oversight is of the utmost importance for our understanding how science is underdetermined,* and for identical reasons, how metaphysics is underdetermined.

Bacon did not explain how he was able to generate his conclusions from his data, not, as some might suppose, because he simply neglected to make the details of the step explicit. Rather the explanation for the omission is that no such account (of the step which generates conclusions from data) is even possible, and this for the reason that it is overwhelmingly likely that there is in fact no such step at all. Bacon never did, his claims to the contrary, generate his conclusion about the nature of heat from the survey he had conducted.

The statistics branch of mathematics has, since the late eighteenth century, provided us with a certain, steadily increasing, collection of inductive techniques. For example, modern polling techniques attempt to predict the behavior or choices of a wider population on the basis of sampling a subset*[3] of that population. In this instance, the inference is from the few to the many. Sometimes a statistical inference may go the other way, as, for example, from the knowledge that 81% of the children in a school system are black, one might infer that more than 60% of the children in some one school in the system are black. In these instances, what we find is that features of certain aggregates are posited to hold of other aggregates (often a subset or superset*[3] of the original aggregates). What it is essential to note in these and in many other sorts of inductive inferences sanctioned within statistics is that it is typically the *scope* of a feature which is being extrapolated, not the introduction or discerning of new features. Bacon's 'induction' is remarkably different. Indeed Bacon takes some pains to contrast his own kind of induction with statistical forms which he calls "simple enumeration" (98). Bacon claims that by using his considerably more powerful induction, he can glean from his data, taken collectively, certain features of each individual item, features which are anything but apparent in those individual items themselves. The inference thus is very unlike the inference that "Sally Jones has green eyes; her brother has green eyes; therefore all the other Jones's children as well have green eyes." In this latter instance, the data explicitly contains the information about each child examined that he/she has green eyes. Bacon's data is not at all like this.

Many, if not most, of his data items contain no *explicit* mention of motion. Consider item 17: "Green and moist vegetables confined and bruised together, as roses in baskets; insomuch that hay, if damp, when stacked, often catches fire"; or item 13: "All villous [long-haired] substances, as wool, skins of animals, and down of birds, have heat" (131). One looks in vain for any explicit or implicit mention here of "motion". Indeed, on the basis of 'data' such as this, if, that is, we were to consider wool as 'having heat', most of us would be positively insistent that whatever it is we are trying to explain, call it heat or whatever, is surely not a form of motion. What, after all, is more inert, i.e. motionless, than a clump of wool? It is clear that one cannot 'read off' of data such as this the conclusion that heat is motion. The trouble is that Bacon has adduced a conclusion intended to explain all of his data but in which there occurs a descriptive term "motion" which does not occur in all or, for that matter, in even very many of

3. For definition in Glossary, see under "set".

his premises. Indeed, Bacon's conclusion was not to prove convincing either to his contemporaries or to several subsequent generations of scientists. When a majority of scientists finally did begin to adopt the theory that heat is motion, in the mid-nineteenth century, it was for reasons remarkably different from the sorts of reasons (and reasoning) advanced by Bacon.

In contemporary philosophy, it is common practice to distinguish two, fundamentally different, senses of "induction". On the one hand we recognize the historical use of "induction" in the sense we have just seen promoted by Bacon. In this first sense, "induction" is said to describe the inferential step from data to a hypothesis explaining that data. On the other hand, "induction" is also used to describe the logic which endeavors to explain, and indeed in some cases to assign numerical values to, the amount of weight a certain body of evidence lends to the support or credibility of a hypothesis. In this latter case, there is no suggestion of one's being able to *generate* the hypothesis from the data; both hypothesis and data (evidence) are taken as 'given' (however each might have been arrived at), and the only matter to be examined by induction is the extent to which the evidence supports the hypothesis. Briefly, the distinction between the two senses of "induction" is reflected in the two phrases "logic of discovery" and "logic of justification".

Bacon clearly thought his inductive methods constituted a logic of discovery, that by the careful, systematic, unbiased collecting of data one could 'by induction' simply 'read off' from that data its explanation.

Beginning in the twentieth century, a number of philosophers – including, especially, K. Popper, H. Reichenbach, R. Carnap, and C. Hempel – have scorned this Baconian concept of induction. The contrary thesis, that there is a logic of discovery, had – until very recently – been defended by only a tiny handful of philosophers, principally J.S. Mill (1806-73), C.S. Peirce (1839-1914), and N.R. Hanson (1924-67).[4]

---

4. In 1962, the historian Thomas Kuhn published *The Structure of Scientific Revolutions* in which he clearly sides with the opinion that there is no logic of discovery. But he argues further that there is no logic of justification either, that the process of accepting or rejecting a scientific theory does not lend itself to appraisal or prediction by logical rules, that ultimately a great number of nonlogical factors influence scientists' decisions to abandon an old theory

Those philosophers who have argued against the very possibility of there being a Baconian logic of discovery have been able to use a so-called bottom-line argument, in effect saying: "If there really is a logic of discovery, show it to us, state its rules and principles so that we and anyone else can apply it to do useful work of scientific discovery."

Unexpectedly, the challenge was taken up. Researchers in Cognitive Science have, for the last thirty years, been actively engaged in precisely this pursuit.[5] They have been trying to discover the ways human beings actually go about making scientific discoveries and have been trying to emulate those methods in computer programs.

As each new science has developed it has revealed an underlying order where none had been previously noted. If we lacked scientific knowledge, most of what happens in the world would appear to us as capricious, random, or the presumed handiwork of a hidden supernatural intelligence. But science frequently reveals an underlying order: in the ways elements form compounds; in the ways living bodies fight diseases; in the ways characteristics are passed from parent to child; in the ways earthquakes come about; in the ways objects fall; etc. Even something as seemingly haphazard as the meandering of streams is found to be explainable (Einstein [65]). So, too, with the subject matter of Cognitive Science. What, prior to recent research in Cognitive Science, appeared to be one of the most capricious, undetermined, of all activities, viz. problem solving, has been revealed, under careful study, to have an unsuspected underlying

---

and to accept a new one. (We will see for ourselves below, when we come to the next case study, how scientists can resist abandoning one theory even when experimental data are produced which are found to be compelling refutation in the minds of some other scientists.) Arguments similar to those of Kuhn are found, too, in the writings of Paul Feyerabend (see e.g. [71]). To a limited extent, this modern historiographical repudiation of a logic of justification had been foreshadowed by Pierre Duhem and by Max Planck. Planck had written: "An important scientific innovation rarely makes its way by gradually winning over and converting its opponents. … What does happen is that its opponents gradually die out and the growing generation is familiarized with the idea from the beginning" ([154], 97; see also [155], 33-4). The views of Duhem, Planck, Kuhn, and Feyerabend have not, however, won unanimous endorsement from other philosophers and remain controversial.

5. Much of this research – theoretical and experimental – was both undertaken and inspired by Herbert A. Simon. See, e.g., [111].

order. Although they are not at all obvious to casual observation, the study under controlled conditions of problem solving is revealing sets of strategies and ploys used by virtually everyone. Problem solving turns out not to be a wild flailing about in a search for anything at all that 'works'. Problem solving is often methodical and systematic.

Those cognitive scientists who have lately revived the claim that there is a logic of discovery base their assertion on their equating scientific discovery with nothing but (a specialized kind of) problem solving: "A hypothesis that will be central to our inquiry is that the mechanisms of scientific discovery are not peculiar to that activity but can be subsumed as special cases of the general mechanisms of problem solving" (Langley et al. [119], 5).[6] But how legitimate is such an identification? Can scientific discovery realistically be equated with problem solving, or is there something more to scientific discovery than merely solving a problem?

Of course scientific discovery is some kind of problem solving. It would be pointless to deny that. Bacon had a problem: "What is heat?" And his eventual answer, "Heat is motion", might, in some sense, be regarded as a 'solution' to that problem. But was his 'route' to his 'solution' via a *logic*, as he claimed; or was it something else, as a great many other philosophers subsequently insisted?

The way Bacon conceived of inductive logic, and the model criticized by many of his successors, was as a logic of discovery understood to be something akin to a calculation. The idea was that one should be able to gather data, subject it to a calculation in accord with some specifiable formula or recipe, and generate therefrom a solution. Metaphorically we can conceive of a logic of discovery as being a kind of 'logic engine': one feeds in the data as input, one turns the crank, and the engine outputs the solution. Such a model suggests a possible realization, not in some visionary apparatus, but in something as mundane as a modern, large-scale digital computer. Indeed cognitive scientists[7] regard the test of their theories to reside just in

---

6. This book, *Scientific Discovery: Computational Explorations of the Creative Process*, contains an extensive bibliography of important work in this field.

7. In the remainder of this chapter I will use the expression "cognitive scientists" for those researchers in cognitive psychology, artificial intelligence (AI), etc. who advance the theory that scientific discovery is a kind of problem solving replicable (in principle) in a program for a digital computer. In

their (eventually) programming computers to act as engines of discovery.

Can such an engine of discovery be built? Equivalently, can computer programs be devised which will generate explanatory hypotheses from observational data? The cognitive scientists say, "Yes … in principle." Their opponents say, "No."

Critics of the claim that a logic of discovery is nothing but a kind of problem solving point to two major differences between what cognitive scientists offer as cases of problem solving and what often have been applauded as cases of scientific discovery. As in any new science, there is a great deal of optimism, a slighting of difficulties, and a tendency to exaggerate the significance of initial findings.[8] And, in this particular case, there has also been a marked penchant for the disputants to argue past one another, often because of subtle shifts in the meanings of central terms in the debate.

A substantial part of the writings of cognitive scientists lays out the experimental findings which reveal how persons will systematically search through what these scientists call "the hypothesis space" (more on this later, footnote 25, p. 187). Their intent is to show that persons do not search among alternative hypotheses aimlessly but do so using what have come to be called *heuristics*: rules of thumb, earlier successful techniques, etc.

But to argue, as some cognitive scientists frequently do, that these findings are good evidence in support of there being a logic of discovery is to overstate the case. For the philosophers who have argued against the logic of discovery do not contend that there is no rational way to select among alternative hypotheses for testing. Their claim, rather, has been that there is no rational means for generating the hypotheses in the first instance. Their objection is that these cognitive

---

using this term in this fashion I am of course distorting the ordinary meaning of the expression. It is just that there is no agreed-upon alternative expression, and I am loath to coin a neologism. I hope that no confusion will result.

8. In the 1950s and 1960s, for example, researchers in machine translation of natural languages and in machine (visual-)pattern recognition believed that within a few years they would be able to program computers to emulate human abilities. Subsequent developments, however, revealed that the problems they were tackling were very much more difficult than supposed at the outset. Initial progress was rapid, but soon gave way to steadily diminishing gains as the remaining problems grew harder and harder.

scientists are helping themselves to too much, are starting, as it were, too late in the day. Given a wealth of hypotheses, one might well argue that there is some strategy for moving among them to select candidates for testing. But that is to misrepresent the problem. The *real* problem lies earlier: in the formulating of the hypotheses initially.

But perhaps the point on which the cognitive scientist and the philosopher-critic will differ most is on the second, viz. the prospects of adducing a set of heuristics capable of yielding those special scientific discoveries which we regard as standing in an exalted, privileged niche: the scientific breakthroughs, the new scientific *theories*.

Much of what cognitive scientists regard with pride as being cases of scientific (re)discovery their critics will dismiss as cases of mere curve fitting.[9] These critics argue that what is needed to generate a genuinely new scientific theory – a theory which goes beyond being a single law and is instead a comprehensive way of looking at a large body of varied data – is precisely the abandonment of heuristics. What is needed are not rules of thumb, or familiar strategies for solving problems, but an act of creative imagination. As long ago as 1949, Herbert Butterfield laid the groundwork for an objection to regarding scientific discoveries, particularly those discoveries we call 'breakthroughs' or 'revolutionary', as being cases of (ordinary) problem solving:

> … of all forms of mental activity, the most difficult to induce … is the art of handling the same bundle of data as before, but placing them in a new system of relations with one another by giving them a different framework, all of which virtually means putting on a different kind of thinking-cap for the moment. … The supreme paradox of the scientific revolution is the fact that things which we find it easy to instill into boys at school … – things which would strike us as the ordinary natural way of looking at the universe, the obvious way of regarding the behaviour of falling bodies, for example – defeated the greatest intellects for centuries, defeated Leonardo da Vinci and at the

9. Trying to find computer programs to generate formulas to fit graphed data (curve fitting) is a task which engineers have been pursuing in industry since at least the 1950s. Fifty years ago, no one regarded such programs as modeling scientific discovery. It is only more recently that cognitive scientists have come to regard them in that way.

> marginal point even Galileo, when their minds were wrestling on the very frontiers of human thought with these very problems. ([40], 1-2)

No existing computer program (least of all the ones ironically named "BACON.1" to "BACON.6" [119]), nor any likely to be developed along lines currently being pursued by cognitive scientists, could possibly emulate or replicate Bacon's own thought processes (whatever they were) which led him to hypothesize that heat is motion. No realistically foreseeable computer program can bridge the gap between data which list a variety of hot things and the creation of the hypothesis "Heat is motion."

Or consider, as another example, Newton's second law of motion published in the *Principia* in 1687 (here reworded): "An object will experience an acceleration directly proportional to, and parallel to, the resultant of the total forces acting on it, and indirectly proportional to its mass." No amount of observing the world could ever provide data from which to generate such a law. For eons, presumably for all of time, prior to Newton's appearance on the scene, there had been massy objects, i.e. objects having mass. But mass, unlike weight, is a feature of the world which is *not* directly observable by any human sense. Newton did not observe the property mass, he posited it, i.e. he hypothesized it as part of a solution to a puzzle. He did that by inventing the *concept* of mass, or, if you find the notion of inventing a concept problematic, you could say that Newton *introduced* the concept of mass to science.[10] His posit was insightful and profitable beyond

---

10. There is an important distinction to be made between our having a concept (e.g. our having a concept of a unicorn, our having a concept of the superego) and there being anything in the world which exemplifies that concept (e.g. there actually being a unicorn, there actually being a superego). I personally happen to have a fairly 'realistic' attitude concerning the status of the referents of useful concepts in science. That is, if a concept in science seems to do the job, if it allows us to state useful and approximately true hypotheses, then I am inclined to regard that concept as referring to some actually existent thing or property in the world. However, a concept may prove useful, even necessary, to a scientific theory without there being anything in the world which it describes (refers to). For example, Newton also posited 'punctiform' masses, i.e. objects having mass but occupying only a mathematical point in space, that is, having zero depth, length, and

his ability to foresee. (It was also, we might note with some interest, eventually to be significantly emended by Einstein more than two centuries later, in 1905.) Newton's invention (introduction) of the concept of mass was a product of his fecund creative imagination, not of any superior powers of observation, not of his possessing a logic of discovery, and not of his utilizing some 'heuristic'.[11] It was as novel, and as free of being governed by a logic or set of recipes, as was, for

---

height. Few physicists, if any, believe that there actually are any punctiform masses in Nature, however useful the concept *punctiform mass* may be in their theories. Realism need not, then, be an unqualified belief in the actual existence of things or properties corresponding to every theoretical term of science. Scientific realism usually is something less than a one-to-one mapping of theoretical terms onto unique features of the world.

Although I have just incidentally admitted that I am a realist about mass, it is important to mention that there is considerable dispute about the wisdom of adopting such a realist position about the relations between the concepts of science and features of the world. One may, instead, adopt an instrumentalist attitude, arguing that scientific concepts may be justified by their successful role within a scientific theory and that one need not take the further step of believing that these concepts refer to bona fide features of the world. The dispute between scientific realists and instrumentalists runs very deep. I will return to a discussion of the concept of *property* in chapter 9; however, I will not pursue in this book the dispute between scientific realists and instrumentalists. Although it is an important dispute within metaphysics, it is not on the agenda for this particular book.

11. Langley et al. cite Clark Glymour's reconstruction ([78]) of the route by which Newton likely arrived at his second law of motion, $F = ma$, and assert that these steps are capable of being programmed so as to permit a computer to output the same formula given Newton's data as input ([119], 54-6). But Newton's accomplishment was not simply the stating of a mathematical relationship between certain variables, "F", "m", and "a". In Newton's hands, these were *interpreted* symbols, standing for, respectively, force, mass, and acceleration. What Newton could do, and computer programs of the sort described by Langley et al. cannot do, is to utilize the *concept of mass* in a comprehensive view, i.e. *theory*, of the world. To output the string of symbols "$F = ma$", as a computer might do, even as a solution to a particular problem we set for the computer using Newton's data, is not to have a theory or to have invented the *concept of mass*. The symbol "m" is not a concept. To describe a computer's outputting of "$F = ma$" as a 'rediscovery' of Newton's second law of motion is to caricature, indeed to misrepresent egregiously, Newton's accomplishment.

example, Beethoven's composing the *Waldstein* Sonata. Just as there are no known rules by which to write sublime music, there are no known rules by which to invent new scientific concepts or to generate scientific hypotheses in which these new concepts occur.[12]

The insistence, by many philosophers, that there are no such generative rules, either known or unknown, strikes many persons who are approaching philosophy for the first time as mistaken. These persons recall having been taught in high school something called 'the scientific method'. "Surely," they want to retort, "there is a method for generating scientific hypotheses."

If there were, then the history of science ought to have been very different from what it in fact has been. If there really were a logic of discovery, some logical procedure, some set of rules, by which one could get from observation to explanatory hypothesis, we should expect that the history of science would simply be a history of successive successes without any controversy or false starts. But the history of science is not at all of this latter sort. It is, instead, a history of a succession of guesses, of controversy, of disputes, of competing theories, of occasional successes, and of far more failures.

The world furnishes up to us its secrets extremely begrudgingly. Nature's 'deep secrets' are not written on the surface, as it were. No amount of careful *observation* can ever reveal the greater part of what we want to know. Observation of the world, no matter how carefully done, is an inadequate tool by itself for understanding the world. To understand the world, we need essentially to proceed by guessing, or if you like, by hypothesizing, and by testing those guesses (hypotheses).[13]

---

12. A third objection to the cognitive scientists' claim that there is logic of discovery, an objection which argues that heuristics are not effective algorithms, will be examined in the following section. Unlike the two objections just leveled, this third objection will be rejected, the counterargument being that although heuristics are not effective, to demand effective algorithms for a logic of discovery is to set an impossibly high requirement.

13. Later (in section 10.7, pp. 311ff.) we will examine a contemporary problem, "What is mind?", that is remarkably similar to Bacon's problem, "What is heat?", in that it, too, defies answer by recourse to simple observation or by an analysis which would construe it as a problem to be solved via heuristics. We will see how the answer to this contemporary puzzle can come about only through bold imaginative theory-construction.

4.3 Metaphysical strands in *The New Organon*

How, we might ask, did Bacon manage to make such a grievous error about scientific methodology*? How might we explain how he came to believe that one could 'read off' of Nature its secrets? I think it worthwhile, particularly given the wider purposes of this book, to dwell a bit longer on this remarkable episode in history.

It is easy, but nonetheless mistaken, to conceive of Bacon, in advancing his methods, as engaging solely in a piece of philosophizing. One may think of his *implementing* his methods, or his *illustrating* those methods, by his actually constructing the various tables we have sampled above, as his *doing science*. Used, as we have become, to insisting on a distinction between doing philosophy and doing science, we may be tempted to try to partition the material of *The New Organon* into nonoverlapping categories: the *description* of the methods is philosophy; the *practice* of those methods is science. Even if Bacon himself did not mark out his work in that way, we, in hindsight, working with more refined concepts, can.

I think that were we to do this, we would commit an error. For I think the best way to make sense of what Bacon was doing is to try to reconstruct what world-view he might have held which would lead him to advance the methods he did. Once we have done that, the result, I suggest, will defy categorization as either philosophy alone or as science alone. In short, if we try to imagine Bacon's world-view, we will discover that his methods were neither philosophy nor science alone, but an inseparable amalgam of the two.

One might think that it is possible to know *a priori** that Bacon's methods could not work. After all, the inductive leap from observational data to an explanatory hypothesis is universally acknowledged to be risky, i.e. not one guaranteed to reveal the truth. Even cognitive scientists who have argued that scientific discovery is a kind of problem solving, all of which proceeds via heuristics, have been careful to insist that heuristics are not effective algorithms*, that heuristics carry no guarantee of even a single solution, and hence no guarantee of a unique solution, still less any guarantee of providing 'the correct' solution. Without an ironclad guarantee, however, it might be supposed, strictly as a matter of logic, that any and every attempt to create a logic of discovery is doomed to failure. It might be thought, that is, that it is *logically* impossible that there should be a logic of discovery.

This pessimistic conclusion is too strong. If one makes it a matter of the very definition of the term "logic" that its results must always be certain, i.e. that logic is truth-preserving in the sense that applying its rules to a set of true premises can produce only true conclusions, then – just as a matter of definition – there can be no logic of discovery. But if we allow a weaker sense of "logic", by which we mean a set of stated, although not foolproof, rules (heuristics) by which to proceed, as when, for example, we talk of the 'logic' of making a medical diagnosis, it remains an entirely open question whether even a crude 'logic' of discovery is possible for the generating of explanatory hypotheses from observational data. It is this weaker sense of "logic" which philosophers such as Peirce and Hanson and present-day cognitive scientists have clearly had in mind when they have tried to defend the thesis that there is some 'logic' governing the activities of scientists in their search for laws.

It may well be that we never do succeed in devising a useful set of rules by which we can generate powerful explanatory hypotheses from observational data. But we will not know whether or not the goal is possible without our actively trying. Whether or not anything like a logic of discovery is possible depends on two critical factors: one, on the way the world is; and two, on our cleverness in making explicit canons (heuristics) for discovery. Although I am strongly of the opinion that a logic of discovery which goes beyond curve fitting and finding generalizations in data to being able to generate powerful explanatory theories is not realizable, my pessimism is not grounded in a priori or metaphysical principles. It stems, rather, from a conviction arising from my own experiential assessment of how complex the world is and how ingenious and multifaceted our explanations of the world have had to be in order to make sense of, and be able to explain, that complexity. Like cognitive scientists, I, too, believe that the question whether there is a logic of discovery is strictly an empirical* one (i.e. one to be decided solely by experience, not by a priori reason). But unlike many cognitive scientists I am convinced that scientific discovery is like ordinary problem solving only up to a point, that beyond that – when one comes to making scientific breakthroughs, doing what is sometimes called "revolutionary" science – what is called for is not the application of familiar heuristics, but the creating of unforeseen and radically new ways of explaining the old and familiar.

We cannot know what sort of world this is without actively exploring it. It might have been a terribly simple world, one in which Nature really does reveal (pretty much) all that is to be known simply by our observing it. It just may be that there is some possible world (here I anticipate chapter 6) in which the sorts of rudimentary methods Bacon advanced would prove very much more successful than they have proven in this world. The point is, however, that neither Bacon nor anyone else could know precisely which sort of world this one happens to be without actually trying out their methods to see whether, and if so to what extent, they worked. It turns out that this world is vastly more complicated than Bacon supposed. The subsequent course of science has found it more and more necessary to hypothesize all sorts of features hidden from direct observation (e.g. subatomic particles, electromagnetic fields, gravitational fields, free markets, capital, information content, and placebo effects) in order to explain those features which are observable. One could not know a priori that extraordinarily imaginative and creative hypothesizing along with the positing of arcane features would be needed for significant advancements in science. Such knowledge is attainable only by the verdict of experience, by actually trying simple methods to see whether they could be made to yield successful theories and by finding that they cannot.

To try to sort out in Bacon's methods what in particular was prompted by 'metaphysical' considerations and what by purely 'scientific' considerations is a hopeless task. Is the belief that Nature is relatively simple a scientific or a metaphysical belief? Our first, natural, response is to regard this question as a metaphysical one. But if one has – as Bacon had – a relatively simple science, and if that science, confining its observations pretty much to what unaided perception can furnish, produces results which are generally accepted and found useful, and if its explanations are regarded as satisfactory, then is the belief that Nature is simple not a belief warranted, not by metaphysics, but by science itself? Such a question strikes me as having no determinate answer. There is no determinate answer, I suggest, because doing science and doing metaphysics blend into one another to such a degree as to make any attempt at dissociation futile.

Bacon's views about the possibility of a logic of discovery were a product of a scientific outlook informed by late-sixteenth-century science. Sixteenth-century scientists and philosophers had no idea, nor could they have had an inkling, how *complex* the world is and how much the future course of scientific development would come to rely on positing a staggering complexity 'behind the appearances'. They could not have known, until they actually tried, and found wanting, methods which relied more heavily on observation than on creative imagination. To find excessive error in Bacon's manner of doing science is to believe that the criteria for judgment are ahistorical. If they are, then our own methods may come to be regarded as being as 'mistaken' as Bacon's. But if Bacon's 'methods' were a product of the late-sixteenth- and early-seventeenth-century world-views, his own views about the nature of explanation itself were rather more fore-sighted.

There is no authoritative account of the nature of explanation. What sorts of accounts are considered to be 'explanatory' change from time to time and place to place. Today many persons find it peculiar that throughout much of history so many persons were content with what we now, somewhat pejoratively, call 'teleological' explanations. For our ancestors, very often to explain why an event occurred was to state what purpose it served in the 'grand scheme of things'. Why was a person born? An answer might have been "to seek the good" or "to glorify God". Today an answer more likely will be in terms of antecedent events: two sexually mature adults had sexual intercourse; a sperm fertilized an egg; or some such account. We have, in modern times, switched our expectations about the very nature of explanation itself. We rarely offer or expect explanations, particularly within the 'non-life' sciences (such as physics, chemistry, astronomy, geology, and meteorology), to be in terms of purpose; rather we expect explanations in those sciences to cite causal factors. Only within the 'life' sciences, biology and medicine, does one still find teleological explanations – e.g. "the purpose of the kidneys is to filter impurities from the blood" – and even there a preponderance of explanations are causal, not teleological.[14] Bacon, himself, was one of the principal critics of the traditional teleological mode of explanation. He argued strenuously for the adoption within physics of the causal model instead:

> … the treating of final causes [i.e. the search for purpose] in physics has driven out the inquiry of physical ones, and made men rest in specious and shadowy causes, without ever searching in earnest after such as are real and truly physical. … "The leaves of trees are to defend the fruit from the sun and wind. The clouds are designed for watering the earth," etc. All … [such examples] … in physics are impertinent and … hinder the

14. In contemporary social science, many explanations which may at first appear to be teleological are often disguised causal explanations. To say of a person that she did something "with the goal (or purpose) of …" is to offer that person's having the goal as a cause of her behavior.

> sciences from holding on their course of improvement, and introduce a neglect of searching after physical causes. ([19], chap. VI, p. 97)

Bacon's views about the roles of teleological and causal explanations were, in due course, to prevail; they were, it turns out, farsighted, almost prescient. But we must be careful not to think that Bacon (and we) now have 'got it right', that earlier in history when persons were as likely as not to offer and accept explanations in terms of purpose, they had 'got it wrong'. We must be careful not to think that causal explanation is 'right' and teleological explanation is 'mistaken'.

Is it possible to know a priori whether there is purpose in Nature, or is this an empirical question? If empirical, what would show it to be true? to be false? If a priori, how could we know it to be true? to be false? My own inclination is to regard the question whether or not there is purpose in Nature as a metaphysical one, one which 'goes beyond' the possibility of experience to answer. But in saying that it is a metaphysical question, I do not mean that it can be settled a priori; quite the contrary, whether we choose to favor teleological or causal explanations depends to a very great extent on the manner in which we practice science and on whether that way is successful. We *adopt* the causal model (or 'paradigm' in Kuhn's terminology), not because there are persuasive a priori arguments in its favor, and not because there is compelling empirical data to warrant that belief. We favor causal explanations because, given our data, given the way science has developed, the causal paradigm suits our purposes, and guides our research, better.

The causal paradigm has gradually, over several centuries, nearly entirely displaced the teleological paradigm, but there were no crucial, definitive, empirical data uncovered, and there were few strong philosophical arguments offered, to have warranted the changeover. It is rather that, as science progressed, as more and more causal explanations were found, and were found useful, persons gradually came to abandon the one model of explanation for the other.

4.4 Case study: Rumford and the Calorists

Questions such as "What is heat?", "What is mind?", and "What is a person?" cannot be answered by *observing* Nature, neither casually nor in the most conscientious, scrupulously diligent manner possible. If one really could settle such a profound question as "What is heat?"

by *observing* Nature, then Bacon – with his elaborate lists – would have settled the matter. But subsequent theorizing about the nature of heat, over the course of nearly four centuries, right up to and including present-day research, provides compelling evidence of the insufficiency of Bacon's optimistic methods. Bacon's methods, although possibly suited for some world or other, an imagined world very much simpler than this one, were woefully inadequate to guide the developing course of science in the actual world.

Bacon published *The New Organon* in 1620, some twenty-eight years after Galileo had invented the first (crude) thermometer. These earliest thermometers lacked scales and, because they were open to the surrounding atmosphere, were significantly affected by changes in barometric pressure. It was not until 1641 that the first sealed thermometer was invented and not until the 1660s that standards emerged for calibrating the scales of thermometers ([174], 120-25). But once scientists had in hand serviceable instruments to measure temperature, the investigation of heat permanently switched from the sort of natural history practiced by Bacon to quantitative research. In the one hundred years after Bacon's death, scientists discovered that various materials had remarkably different specific heats (see footnote 2, p. 44) and discovered the phenomenon of latent heat (the heat of fusion, i.e. the heat needed to melt a substance, and the heat of vaporization, i.e. the heat needed to vaporize a substance). Where Bacon had merely produced lists of items 'having heat', his successors turned their efforts to measuring the amounts of heat needed to effect changes in substances.

While the developing quantitative and experimental methods seemed well suited to answering such questions as "How much heat is absorbed by one pound of ice in melting?", these same quantitative methods seemed unable to answer Bacon's initial, and 'deeper', question "What is heat itself?" Bacon's own answer had been, we have seen, that heat is motion. This theory is sometimes called the "dynamic" view, and later, toward the end of the nineteenth century, came to bear its modern name, the "kinetic" theory of heat. The dynamic theory was, in effect, that the heat of objects and of gases is due to the vibration of their constituent particles ("atoms" or "molecules" in modern terminology). These postulated vibrations were also sometimes referred to as "intestine [internal] tremors". But the trouble with Bacon's conclusions about the nature of heat was that those conclusions could not be seen by his successors, in spite of his claim that those conclusions were reached by an induction, to be in any way dictated by or generable from his data. Nor, for that matter, was his theory the sort which was much favored in the eighteenth century. For at that time, the scientific climate favored a static theory.

Gases, it was readily observed, were ‘elastic’: they resisted compression and would expand to fill their container. How was one to explain this phenomenon? According to Newton, all material objects (the particles of gases included) *attract* one another. Why then should a gas expand to fill its container, rather than collapse into a liquid or a solid? Obviously – so reasoned many scientists – there must be a repulsive force as well, a force opposing the gravitational attraction of the particles. Where did such a repulsive force originate? An ‘internal tremor’ seemed not especially promising as a source of repulsive forces. (To cite a modern analogy: the vibrating strings of a guitar do not seem to repel one another.) Instead, theorists turned to adapt what they could of Newtonian theory, the most successful physical theory that humankind had yet produced. Just as material particles attract one another under gravitational forces, there ‘must be’ – they reasoned – another kind of substance whose particles repel one another and which lie between the particles of matter.

The route to the theory is fairly obvious: if forces are pushes or pulls (the only sorts of forces recognized in Newtonian mechanics) and if material particles attract one another, and if gases expand rather than collapse, there ‘must’ then be other sorts of particles, nonmaterial ones, whose nature it is to repel one another, rather than to attract. It was understandable, then, that theorists should postulate another, nonmaterial, kind of stuff, a stuff which permeated gases and physical objects, and which tended to drive apart the material particles. This posited stuff came to be regarded as a kind of fluid and was called “caloric”. Inasmuch as most materials and gases expand when heated, it was an easy and obvious step to identify this caloric with heat itself. Why do material objects expand when heated? Simply because more caloric had been added to them and the additional caloric exercised a stronger repulsive force causing the expansion. There was no need in this theory for attributing any motion, vibrational or random, to the particles of caloric. Their sheer number, not their activity, was what accounted for expansion, sensations of elevated temperature, melting, etc. And thus the caloric theory was regarded as the ‘static’ theory of heat. On this theory, heat was a kind of stuff; it was not a vibration or tremor or motion of material particles.

There was at least one other major factor favoring the caloric theory. In areas of physics outside of heat – in light, in magnetism, and in electricity – all the best theories of the time were theories of special kinds of fluids, i.e. light, magnetism, and electricity were all thought to be accountable for in terms of special, subtle, weightless fluids. It would be an understatement, indeed something of a distortion, to say that positing the existence of caloric was done *on analogy* with the theories of light, electricity, and magnetism. Quite the contrary, the positing was done in accord with the overriding model of what the world was like. To posit a fluid to account for the nature of heat was as natural and as acceptable in the eighteenth century as it is in our own day to posit a virus as the cause of some particular disease. If today we were to describe a scientist who posited a virus as the cause of, let us say, multiple sclerosis, as proceeding by constructing an analogy with the explanation of the cause of poliomyelitis, we would, I think, feel that we had seriously underdescribed (if I may be permitted to coin such a word) the situation. Positing viruses nowadays as the causes of specific diseases is not so much constructing an analogy as it is simply following the normal, accepted, and expected practices of biological theorizing. Put another way, our positing a virus in this instance is in keeping with our world-view, call it "physical", call it "metaphysical"; it makes no difference. So too (and this statement *is* an analogy) was positing caloric in the eighteenth century. That was not a bold, analogical conjecture. It was, by that time, simply the obvious theory to promote. Such a hypothesis enjoyed, at that time, the fullest measure of scientific approbation and naturalness. It was, that is to say, fully in keeping with the then-current physical/metaphysical world-view.

Thus we find Joseph Black (1728-99), for one, arguing explicitly against Bacon's 'dynamic' theory, first on the grounds that it was counterintuitive,[15] and second, on the grounds that it was contrary to experimental findings:

> I cannot form to myself a conception of this internal tremor, that has any tendency to explain, even the more simple effects of heat, or those phenomena which indicate its presence in a body; and I think that Lord Verulam [Bacon] and his followers have been contented with very slight resemblances indeed, between those most simple effects of heat, and the legitimate

15. I will have more to say in chapter 6 (p. 105) about the role of so-called intuitions in informing our world-views.

> consequences of a tremulous motion. I also see many cases, in which intense heat is produced in this way, but where I am certain that the internal tremor is incomparably less than in other cases of percussion, similar in all other respects. Thus the blows, which make a piece of soft iron intensely hot, produce no [appreciable] heat in a similar piece of very elastic steel. ([30], 32-3)

Black has here raised what he regards as a crucial objection to Bacon's theory: if heat were motion, then in hammering equally two different pieces of iron, one soft and one elastic, the two pieces of iron should heat up equally. But they do not. Therefore, Black suggests, the theory that heat is motion is refuted.

In hindsight, living in an age where the dynamic (kinetic) theory has supplanted the caloric theory, and where the dynamic theory (in conjunction with quantum mechanics) can and does explain why some hammered materials heat up more than others, we may be inclined to regard Black's 'refutation' of the dynamic theory as disingenuous. But any such criticism would be anachronistic, in effect faulting Black for not having foreseen the subsequent development of science.

Black, like most eighteenth-century physicists, strongly preferred the caloric theory to – what he regarded as – Bacon's insupportable, vibrational theory. But one year before Black's death, Count Rumford read a paper (25 January 1798) before the Royal Society of London, describing a series of experiments which, some fifty years later, came to be regarded as strong evidence of the correctness of the vibrational theory and of the inadequacy of the caloric theory. But at the time, at the end of the eighteenth century and through much of the first half of the nineteenth, Rumford's experiments were either dismissed or their results believed to be accountable for within the prevailing caloric theory.

Rumford's name at birth had been "Benjamin Thompson". He was born in 1753 in Massachusetts, which was then still a colony of England. When the revolutionary war came, Thompson remained a loyalist, and when the British army evacuated Boston in 1776, he sailed for Europe. Although he was to maintain a correspondence with persons in America, and was to donate generous sums for scientific research there, he never returned. On the Continent, Thompson entered the service of the Elector of Bavaria, and in due course became the minister of war, gaining the title "Count Rumford" in 1791. In his capacity as minister of war, he became the superintendent of the military arsenal in Munich. And it was there that he undertook his most famous experiments on heat.

> Being engaged lately in superintending the boring of cannon …, I was struck with the very considerable degree of Heat which a brass gun acquires in a short time in being bored, and with the still more intense Heat (much greater than that of boiling water, as I found by experiment) of the metallic chips separated from it by the borer.
>
> … *whence comes* the Heat actually produced in the mechanical operation above mentioned?
>
> Is it furnished by the metallic chips which are separated by the borer from the solid mass of metal?
>
> If this were the case, then, according to the modern doctrines of latent Heat, and of caloric, the *capacity for Heat* of the parts of the metal, so reduced to chips, ought not only to be changed, but the change undergone by them should be sufficiently great to account for *all* the Heat produced.
>
> But no such change had taken place. ([176], 4-5)

Rumford begins by examining one possible hypothesis the calorists might have offered for the rise in temperature: the heat is being generated by pulverizing the metal. The idea here is that the total amount of heat in a large block of metal is greater than that in its smaller parts, and that in reducing the original to chips and shavings, the 'surplus' heat of the whole is released. But Rumford then reports on an experiment in which he compares the amount of heat furnished to a given mass of shavings and an equal mass of metal strips (taken from the same original block of brass), by submerging them first into boiling water and then into cold water to see how much heat each absorbs from the hot water and how much each in turn releases to the cold water. He finds no appreciable difference between the shavings and the larger strips.

These initial results, even if they hardly constitute definitive disproof of the caloric theory, are interpreted by Rumford as evidence that there is something seriously amiss in that theory. He begins to believe that the heat being generated is not anything 'latent' in the cannon itself, but is coming about through the conversion of the mechanical energy needed to turn the machinery. In short, Rumford now strongly suspects that heat is not a kind of fluid, but is, in his words, ‘excited by friction’. But how can he *prove* any of this?

Rumford then undertakes a series of four further experiments: (1) to measure quantitatively the amount of heat produced by friction (taking the precaution of insulating his apparatus); (2) to determine what the effect of excluding air would be (he finds none); (3) to see what effect there would be if the apparatus were surrounded with a water jacket (the water rises in temperature and eventually boils [see p. 5 above]); and (4) to test whether filling the bore with water will change the results (he finds that it does not). With these further experiments in hand, his conclusions are uncompromising:

> What is Heat? … Is there anything that can with propriety be called *caloric?*
>
> We have seen that a very considerable quantity of Heat may be excited in the friction of two metallic surfaces, and given off in a constant stream or flux *in all directions* without interruption or intermission, and without any sign of diminution or exhaustion.
>
> From whence came the Heat which was continually given off in this manner in the foregoing experiments? Was it furnished by the small particles of metal, detached from the larger solid masses, on their being rubbed together? This, as we have already seen, could not possibly have been the case.
>
> Was it furnished by the air? This could not have been the case; for, in three of the experiments, the machinery being kept immersed in water, the access of the air of the atmosphere was completely prevented.
>
> Was it furnished by the water which surrounded the machinery? That this could not have been the case is evident: *first*, because this water was continually *receiving Heat* from the machinery, and could not at the same time be *giving to*, and *receiving Heat from*, the same body; and, *secondly*, because there was no chemical decomposition of any part of this water. …
>
> Is it possible that the Heat could have been supplied by means of the iron bar to the end of which the blunt steel borer was fixed? or by the small neck of the gun-metal by which the hollow cylinder was united to the cannon? These suppositions appear more improbable even than either of those before mentioned; for Heat was continually going off, or *out of the machin-*

> *ery*, by both these passages, during the whole time the experiment lasted.
>
> And, in reasoning on this subject, we must not forget to consider that most remarkable circumstance, that the source of the Heat generated by friction, in these experiments, appeared evidently to be *inexhaustible*.
>
> It is hardly necessary to add, that anything which any *insulated* body, or system of bodies, can continue to furnish *without limitation*, cannot possibly be *a material substance;* and it appears to me to be extremely difficult, if not quite impossible, to form any distinct idea of anything capable of being excited and communicated in the manner Heat was excited and communicated in these experiments, except it be MOTION. ([176], 20-2)

It is interesting to compare the similarity of phrases, but the diametrically opposed views of Black and of Rumford, concerning the very possibility of entertaining the other's point of view. Black: "I cannot form to myself a conception of this internal tremor, that has any tendency to explain, even the more simple effects of heat." And Rumford: "it appears to me to be extremely difficult, if not quite impossible, to form any distinct idea of anything capable of being excited and communicated in the manner Heat was excited and communicated in these experiments, except it be motion." Two eminent scientists, writing at virtually the same time, are incapable – each confesses – of being able to subscribe to the opposing theory.

For the half-century following the publication of his experiments, Rumford's conclusion – that heat is a form of motion – was not only not accepted, it was positively and actively rebutted by calorists. His opponents were not crank scientists, but were among the best of their day.

On 5 June 1801 William Henry read a paper (actually written two years earlier, almost immediately after Rumford first made his experiments public) to the Manchester Literary and Philosophical Society, in which he raised serious objection to Rumford's theory: "… the Count has observed that water could not, at the same instant, be in the act of giving out and receiving heat. … But I cannot admit that the argument is demonstrative, in proving the evolved caloric not to be derived from external substances; for no absurdity is implied in supposing, that a body may be receiving caloric in one state, and giving it out in another" ([92], 606-7).

Even as late as 1856, some fifty-eight years after Rumford had read his first paper to the Royal Society, we can still find strenuous defenses of the caloric theory. Thomas Traill, the editor of the eighth edition of the *Encyclopaedia Britannica*, undertook himself to write the article on Heat in which he explicitly argued against the vibratory theory.

> The other opinion, which has been maintained by Bacon, Boyle, and several other philosophers,[16] considers heat as a mere quality of matter, and ascribes it to a vibratory movement among the intimate particles of bodies; an idea which was adopted by Rumford to explain his curious experiments on the excitation and communication of heat by friction. This opinion, however, seems vague and unsatisfactory. If we say that heat is motion amongst the particles of matter, still we have no explanation of the manner in which this motion is produced; for we cannot conceive any movement without an impulse, nor an impulse without material agent. … [If heat were to] consist in vibrations or motions of the particles of other matter, it should pervade elastic bodies with the greatest celerity; which we know not to be the fact. … If we mingle together equal quantities of water at different temperatures, the resulting temperature will be an exact mean between the extremes. But if heat consisted in such vibrations, there ought to have been a loss of heat, as in all other communicated motions. … Still more difficult is it to conceive how a permanent temperature could subsist among a great system of bodies, as the planets, if heat were nothing more than a vibration of the particles of bodies; for the original impulse ought to diminish with each communication. ([205], 260)

Among a variety of other objections, we can see here that Traill musters some strong counterevidence to Rumford's theory. Heat is supposed, on Rumford's theory, to be an internal mechanical vibration of a physical object. It presumably, then, ought to be conducted through physical objects with the same speed that mechanical im-

---

16. It is surprising to learn that the very term "scientist" is of very recent origin, having been coined by William Whewell (1794-1866) in 1840 (see Medawar [134], 9). As we can see, it had not achieved universal adoption by 1856.

pulses are transmitted through those objects. In steel, for example, mechanical impulses are transmitted at the speed of 4975 m/sec. Were you, then, to rap one end of a meter-long steel rod sharply, you would feel the impulse at the other end 0.0002 sec (two ten-thousandths of a second) later, i.e. virtually instantaneously. But were you, grasping one end of that same rod, to plunge the other into a fire, it would take some appreciable time, several minutes perhaps, before the end in your hand would grow noticeably warm.

Rumford's theory, we can see, met with opposition for a variety of reasons, not least because it did not offer a *quantitative* account of the nature of heat. Moreover, his theory seemed to contradict, without explanation, certain fundamentals of mechanics, the most basic and respected scientific theory of the day. But even that was not the end of it. For as Henry points out, it rested on certain quite unproven presuppositions, e.g. that a body could not be simultaneously gaining and losing caloric. Such assumptions, while they might have commended themselves with virtual a priori certainty to Rumford, were not in fact demonstrated by experiment or grounded in any theory accepted at that time, and were not nearly so 'self-evident' to Henry or other calorists.

The remarkable French philosopher-scientist-historian Pierre Duhem, writing in 1906, offered this perspective on such disputes.

> Now it may be good sense that permits us to decide between two physicists. It may be that we do not approve of the haste with which the second one upsets the principles of a vast and harmoniously constructed theory whereas a modification of detail, a slight correction, would have sufficed to put these theories in accord with the facts. On the other hand, it may be that we may find it childish and unreasonable for the first physicist to maintain obstinately at any cost, at the price of continual repairs and many tangled-up stays, the worm-eaten columns of a building tottering in every part, when by razing those columns it would be possible to construct a simple, elegant, and solid system.
>
> But these reasons of good sense do not impose themselves with the same implacable rigor that the prescriptions of logic do. There is something vague and uncertain about them; they do not reveal themselves at the same time with the same degree of clarity to all minds. Hence, the possibility of lengthy quarrels between the adherents of an old system and the partisans of a

> new doctrine, each camp claiming to have good sense on its side, each party finding the reasons of the adversary inadequate. ([60], 217)

4.5 The ineliminability of unproved presuppositions

The ensuing debate between the two schools of scientists – the kineticists (as they were eventually to be called) and the calorists – is not atypical. Quite the contrary: the sort of dispute we have seen in this instance has occurred, and will continue to occur, frequently in science. Controversies about scientific theories and the degree to which any given theory is confirmed or disconfirmed by experiment and observation are inevitable; they are, in fact, virtually mandated by certain logical principles.

The testing of scientific theories is not at all the straightforward, unambiguous, procedure it has often, historically, been portrayed to be. An experiment, and its overarching theory, which may be utterly convincing to one scientist, may be just as unconvincing to another. Such differences are not usually to be accounted for in terms of stubbornness, intellectual blindness, dishonesty, conservatism, or the like. Disputes between scientists usually arise, not because of psychological differences between personalities, but because of important principles at play in the *logic* of subjecting scientific theories to empirical testing.

It can be proven – relatively easily as a matter of fact – that for any set of data about the world (i.e. for any set of contingent* data), there must exist logically independent alternative sets of explanations for that data, indeed there are an infinite number of such alternative sets. Intuitively, in less technical vocabulary, this means that the 'fit' between theories and experimental and observational data is remarkably 'loose', and that for any proposed theory or explanation of a phenomenon, there must, theoretically, exist alternative theories or explanations which are compatible with the data. This is not to say, of course, that all such alternative explanations are equally probable or that they are equally attractive to us. The point is that experiment and observation are never themselves sufficient to eliminate all possible contenders among alternative explanations.

Rumford's conclusions were not convincing to his critics. The calorists had no difficulty whatsoever in homing in on all sorts of unproven presuppositions in his arguments. Clearly, Rumford's conclusions survived or floundered upon the correctness or incorrectness of these many, many presuppositions. But he was not in a position to test these presuppositions. Had he attempted that, his experimenting could have gone on forever.

The presuppositions we must bring to any of our experiments are virtually without limit, and there is no practical way of markedly reducing their number. Even as simple an ‘experiment’ as measuring a room for a carpet is encumbered, we find, by vast numbers of untested presuppositions. What sorts of presuppositions must be true for our measurements of the floor area to be correct? Our tape measures must be accurate; the walls of the room must meet at right angles, or, if not, we must have some means for measuring those angles; our tape measures must not change length as we move about in space; the area of the room must be calculable by some known formula; measurements of length must be independent of the time of day; the visual appearance of the room must have certain known relationships to its physical layout; etc. Were we to put some of these presuppositions themselves to the test, those very tests would themselves, in their turn, carry a number of untested presuppositions. For example, were we to test the angles of the walls, we should then have to ask whether our measuring instruments were accurate. How shall we test *them*? By using still other instruments. But what about the latter? We are faced with the potential of an infinite regress of presuppositions which it is impossible to complete.

Throughout the greater part of the twentieth century, and to a greater degree than any of his contemporaries, Karl Popper emphasized the role, the ineliminability, and the potential inexhaustibility of untested presuppositions in our doing of science. At first, the very existence of such untested and ineliminable presuppositions may be thought to give the lie to, indeed to make utterly impossible, the claim that science is ‘objective’. But Popper argued that this pessimistic conclusion is not forced upon us. In his view, objectivity does not – and more importantly, *could not* – consist in our being able to prove a theory to be true. The number of presuppositions in each of our theories seems to be without limit. They range from highly specific presuppositions of particular theories, e.g. that no physical object can take in and give off caloric at one and the same time, to the most general (often labeled “metaphysical”) presuppositions which ground virtually all our theories, e.g. that we will not wake up in an hour and discover that what we have taken to be reality was in fact nothing but a dream. If objectivity consisted in being able to prove a theory *true*, and if proving a theory true involved proving that every presupposition of that theory is true, then simply because the latter – proving the presuppositions true – would be an infinite task, nothing could be deemed to be objective. Popper's reply to this – one which I think is fundamentally sound – was to argue that this latter conception of objectivity is useless. It is useless because it never could apply to anything. Instead Popper urged that we conceive of objectivity, not as an *accomplishment*, but rather as an *attitude*: a critical frame of mind. Scientists are objective, not if they attend open-mindedly to their data and let it 'dictate' the theory, but rather if they admit to their presuppositions and recognize that their conclusions rest on those presuppositions and if they do not try to prejudge a priori or dogmatically what further tests of those presuppositions might reveal.

The logic of testing *scientific* theories is important, for illustrative purposes, because it is, as well, the logic of testing *any* theories, not just scientific, but the most mundane through to the most 'metaphysical'. There is no logic special or unique to the sciences. The logic of theory testing is the same for the child in her crib theorizing about the persistence of unperceived objects as it is for an adult theorizing about the efficacy of using a tape measure to fit a room for a carpet, as it is for the scientist theorizing about the nature of heat, and as it is for the metaphysician theorizing about the nature of space and time. What may differ is the degree to which the theorizing lends itself to empirical testing. But the logic is the same throughout, and the possibility of conclusive proof is not to be realized: there are only degrees of probability (about which – incidentally – there are enormous differences of opinion among researchers in the philosophy of logic).

To be sure, we do not actively *entertain* all, or even many, of the infinite number of potentially confounding factors when we proceed to do something as commonplace as measuring a room. But it is equally clear that these factors must *be* as we just described, if we are to have success. Only if, for example, tape measures do not change length as we move about the room, can our measuring the room work. As we go about our lives, doing what we commonly do, searching for misplaced scissors, measuring rooms, cooking meals, driving cars, turning on radios, etc., we do so in a context of making untold numbers of unproven assumptions. You may seem to recall having had a glass of milk last night. But how good is your memory? Is it perfect? Couldn't somebody have substituted soybean extract which you mistook for milk? Etc.

The point is the same when it comes to doing science, only the degree of uncertainty of the presuppositions is greater. You may reasonably be sure (but ought you to be utterly convinced?) that the walls of your room are square; but what, if anything, entitled Rumford to his – unproven – belief that an object could not simultaneously gain and lose caloric? What warranted his unproven belief that caloric could not be communicated to his apparatus by means other than those he had examined? Simply: he was not justified in these beliefs. He held them because – taken with his theory – they seemed to him to provide a better explanation of what was going on than did the competing caloric theory. But these very presuppositions did not seem compelling, indeed seemed false, to other equally rational scientists, the calorists.

Identical claims may be made for the practice of metaphysical theorizing. In metaphysics – just like commonplace theorizing, and just like scientific theorizing – theories are underdetermined by empirical data. We cannot *prove* that the world did not spring into existence ten minutes ago complete with fossil records, libraries, adult human beings with (apparent) memories, etc. But we theorize that such did not happen, and we do so because an alternative theory seems to work better. But there is nothing that can be regarded as a conclusive test of the truth of either theory.

Some philosophers and a somewhat greater number of cognitive scientists, however, have tried to take the sting out of underdeterminism by arguing that although there are potentially, from a *logical* point of view, an infinite number of different explanations of any given phenomenon, there are generally only a very few alternative explanations *psychologically* available. They will often cite in these arguments the fact that until Einstein's physics appeared, there was no good or reasonable alternative to Newtonian mechanics; that even now there are not many, if any, contenders against Einstein's special theory of relativity; and that there are few, if any, challengers to quantum mechanics. But the trouble with this defense and its complacent attitude about the possibility of underdeterminism is familiar: it focuses on too few examples taken from but one highly specialized area of human knowledge. What is, as we have seen, at best only a half-truth* about physics is surely not true of other areas of human interest. When we look outside of physics, we often do not find dominant, relatively unchallenged theories. Often we will find many, sometimes a bewildering variety of, alternative theories offered as answers to some of our most interesting and pressing concerns.

What is the concept of *causality*?
What are rights and obligations? How are we to recognize and/or agree to them?
What is a person?
How is knowledge possible?
What is a mind?
What is a soul?
Is there purpose in the universe?
Does God exist?
What are beliefs?
How does language work?
Can the language of science be translated into the language of logic?
To what extent is the atmosphere able to absorb industrial pollutants? Of what kinds?
How shall "gross national product" be defined?
Is punishment morally justified?
Does free will exist?
What makes a particular piece of art worthy or good?
Is *intelligence* a meaningful concept and, if so, can intelligence be measured in a society-independent manner?
Is a value-free science possible?
Are there historical forces?
How best should a society protect minority rights?
What moral justification is there for limiting immigration?
Do animals have rights?
Do males and females differ in their ability to do mathematics?
Etc.

Directing attention to just the first of this list, to the question regarding the analysis of *causality*, we find not just one, or even just a few, but a very great number of theories. One recent author ([94], 14-21) lists ten contemporary theories. And even at that, his list is incomplete.

Having a plethora of alternative theories, far from being 'psychologically unlikely', is in fact the norm. Underdeterminism is not just a logical possibility. It is in fact one of the most pervasive features of our attempts to make sense of the world. All of our theorizing, without exception, whether in science, philosophy, jurisprudence, etc., is underdetermined by the empirical data. We can never hope to 'read the truth off the world' as it were. All we can ever hope to do is to propose theories to try to make sense of the flood of data and to work at trying to improve these theories and to replace theories with better ones. If we do not recognize the underdeterminism in our theories, then we will unwittingly become dogmatists thinking all the while that we 'have seen the truth'.

Truth there well may be. But human beings, unfortunately, have no privileged access to truth when we try to construct scientific and philosophical theories about the world. In generating and adopting such theories there is for us no other method than the dialectical one of trial, error, trial, error, …[17]

---

17. Learning that a hypothesis is false rarely ends our trials. Our interest lies with highly specific hypotheses, ones which have what is called a high 'information content'. On a scale of 0.00 to 1.00, we desire hypotheses with information content close to 1.00 (see e.g. [34], 370-81). Learning that a hypothesis which has high informational content (e.g. 0.998) is in error is to learn the truth of a proposition whose informational content is very low (e.g. 0.002). For example, you may guess (hypothesize) that my telephone number is 555-9993. You put this hypothesis to a test (trial) and learn that it is false. You now have learned something true alright, but its informational content, and its practical value to you, is virtually nil. Inasmuch as there are – in principle – 10 million different 7-digit telephone numbers, learning (the truth) that 555-9993 is *not* my telephone number puts you only one ten-millionth closer to finding out what my telephone number actually is. Discovering that a hypothesis is false may be useful as a goad to further hypothesizing; but it is, in general, no substitute for learning that a hypothesis has withstood falsifying and may, therefore, be true.