CHAPTER 5

BACONIAN CONVERGENCE

How science's consensus on procedure, enforced by the iron rule, leads to discovery

IN 1618, Sir Francis Bacon was elevated to the post of Lord Chancellor of England, becoming one of the highest legal officials in the land and a close adviser to the king, James I. Just three years later in 1621, Bacon was denounced by his political enemies and put on trial for bribery. He escaped with his life but not his honor, retiring in disgrace. Somehow, in the middle of that short, tumultuous time, Bacon published one of the most significant books ever written on scientific inquiry: *The New Organon*, a blueprint for the knowledge machine that was to be constructed over the subsequent decades of the Scientific Revolution. Such was the book's fame that the poet Abraham Cowley, in a paean that prefaced Thomas Sprat's 1667 history of the Royal Society, declared Bacon to be "Lord Chancellor of the laws of nature."

Bacon admired the ancient Greek natural philosophers enormously, yet they had plainly failed, he saw, in their endeavors, their inquiries going "round in circles for ever, with meager, almost negligible, progress." Thus, "a new beginning has to be made from the most basic foundations": the old philosophical ways would have to give way to a novel method for discerning the deep structure of the natural world.

The primary casualty of Bacon's war on the past was to be Aristotle's *Organon*, meaning "the tools," a treatise on the rules of reasoning and the proper organization of knowledge regarded as the supreme work on these topics during the Middle Ages. It would be replaced, Bacon hoped, by his own work, whose ambitions were made clear on the title page: his *New Organon* would supersede Aristotle's ancient work.

When the book was published in 1620, the Scientific Revolution was barely underway. Galileo had used a telescope to observe the moons of Jupiter in 1609, calling into question the Aristotelian thesis that all celestial bodies orbit the center of the universe (which Jupiter most certainly was not). In the same year, Kepler published his first two laws of planetary motion, describing in mathematical terms the elliptical orbits of the planets around the sun. But the *New Organon* is not so much a part of the Scientific Revolution's first act as it is its Shakespearean prologue, laying the groundwork, heralding the action, anticipating the denouement, from the viewpoint of one who sees all that is to come—the playwright himself.

A method for discovering everything—that was what Bacon promised his readers. To show off its power, he proposed to apply it to a great mystery of the time, the nature of heat (the very same topic debated in the previous chapter by the Baconian characters Montague and Capulet).

The natural philosophers before him had failed to solve the problem of heat, as they had failed in so much else, but it was not, Bacon thought, for lack of clues. To understand nature, you must open your eyes to these clues. But before you begin, you must empty your mind.

You must put aside your personal prejudices, loyalties, and preferences. You must put aside the prejudices common to the human race, such as the tendency to oversimplify, to neglect to look for evidence against your pet theories, or to pay more attention to freakish than to everyday phenomena. You must put aside all prejudices inherent in your language—you must not assume, for example, that two things having the same name have a similar nature. (Just because a hairball is made of hair, don't expect a football to be made of feet.) You must put aside all philosophy and therefore any prior science that is based even in part on philosophical thinking, such as the science of the ancient Greeks. Indeed, you must put aside all learning that is not produced by the Baconian method. These distractions and temptations Bacon called "idols"; to allow them to influence your thinking was to worship false gods rather than to revere reason.

Can any human being hope to discard not only their particular idiosyncrasies, but also the suppositions and inclinations that are their birthright as an educated person, as a member of their culture and of their species? When at the height of his brilliant career Bacon was put on trial for corruption, he declared that he was "as innocent as any born upon St. Innocents Day"—because he had not allowed the bribes he had taken to sway his judgment. Perhaps he did indeed have a mind strong enough and a soul stern enough to resist any urge to favor those who had smiled on him, but the stories of Eddington and the eclipse, of Pasteur and spontaneous generation, of continental drift and the age of the earth suggest that scientists, even great scientists, are rather more pliable than that.

Never mind. Let's suppose that Bacon's mind is successfully purged of all inclination to venerate the idols; let's see how he pulls together the clues to infer the nature of heat.

His first step: assemble all positive instances of heat, that is, all types of circumstances in which heat is present: in fire, in the bodies of animals, when the rays of the sun are concentrated using a magnifying glass, when two solid bodies are rubbed together. And then there are flaming meteors, natural hot baths, spicy herbs, and more.

His second step: assemble all negative instances of heat, that is, all types of circumstances in which heat is not present. Of course, this list must be endless, so Bacon recommends the following alternative. For each of the positive instances, find similar circumstances in which heat is absent. For example: heat is absent in the bodies of dead animals or when two solid bodies are held together but not rubbed (and apparently in moonbeams, comets, and St. Elmo's fire—though as Bacon remarks, further investigation is needed).

The third and final step is to assemble all the ways in which heat varies with other quantities, for example, the way that objects get hotter the closer they are to a fire or the way that metals take longer to get hot than air but retain their heat longer.

With the third step completed, Bacon has the evidence he needs before him. What, he asks himself, can possibly explain all the assembled facts? What can explain why heat is present in live animals but not dead animals, why it is generated by friction, why metals soak up more heat than air? He considers many suggestions concerning the essential character of heat. Each explains at least a few of the positive and negative instances but fails to explain others and so is rejected: "Every contradictory instance destroys a conjecture about a form." Heat cannot be "light and brightness," because substances such as boiling water can be hot without being light or bright. Heat cannot be a substance, because when a hot iron warms another object, it loses heat but does not lose any weight. And so on.

At the end of the process, there is only one hypothesis standing, only one conjecture about the nature of heat that can explain every circumstance upon which Bacon has trained his gaze. It is the hypothesis that explains how friction generates heat: "The quiddity of heat is motion and nothing else," or, more exactly, heat is the disordered motion—the vibration—of the small particles of which all things are made. This idea, an early version of the kinetic theory, is, we now know, quite correct: a compelling advertisement for the Baconian method.

The *New Organon* recommends this same technique to investigate every natural phenomenon, from lightning to laryngitis to life itself: gather the conditions under which the phenomenon occurs, the conditions under which it does not occur, its patterns of change, and find the hypothesis that explains the lot—the occurrences, the nonoccurrences, the variation. That hypothesis is the theory you're looking for, the truth.

Bacon imagined that once scientists got to work using his discovery method to understand a phenomenon, they would quickly rule out all explanations but the correct one. That turned out to be rather optimistic. In the hundred years or so after Bacon laid down his prescriptions, the number of interesting, plausible explanations was multiplying rather than declining, as the great creative minds of the seventeenth and eighteenth centuries went to work attempting to make sense of the world. Bacon's triumphant revelation of the kinetic nature of heat, for example, was (as you know) challenged around 1800 by the development of the rival caloric theory, which was thought at that time to better account for the assembled facts. To scientists facing Capulet and Montague's predicament, Bacon's method provides no objective guidance in determining the nature of heat. They must, as we saw, fall back on their personal estimates of the likelihoods of various possibilities—their plausibility rankings—which inevitably means letting the idols, Bacon's icons of subjectivity, have their say.

The point of our visit to the early seventeenth century was not, however, to understand why copies of the *New Organon* are rarely spotted on the modern laboratory workbench. It was to understand how science eventually gets to the correct view of things. The difficulties in using the Baconian method arise not because there is something wrong in principle with Bacon's assumption that the truth and only the truth can explain everything, but because it is almost impossible to apply in practice: in the middle of scientific inquiry, there are typically several competing hypotheses of roughly comparable explanatory success.

At some point, nevertheless, the conflict begins to subside. The moment that Bacon longed for arrives: the treasury of evidence has become so rich that it singles out just one theory as supreme explainer, the sole theory capable of making sense of every observable pattern—in Bacon's terms, every "positive instance," every "negative instance," every pattern of variation. Viewed from this cosmic perspective, Bacon's idea looks not so different from Popper's. Each false theory has been confronted with some piece of evidence that it cannot explain, and in the light of its explanatory inadequacy has been tossed aside. Only the truth remains.

Popper rejected any role for plausibility rankings in scientific reasoning; they were, to his philosophical taste, contemptibly inductive. Let them back in, though, and you'll find that something interesting happens as science closes in on the last theory standing. As evidence accumulates, plausibility rankings begin to converge. Differences in opinion become less extreme. Consensus emerges as to which are the leading theoretical contenders and which are the also-rans, then eventually on which is the best of them all. There is not complete agreement, but there is ever less disagreement. This is *Baconian convergence.*

We do not have to wait indefinitely for Baconian convergence to do its work. The knowledge machine has, in just a few hundred years, discovered viruses, DNA, the nature of heat, the genes underlying most animals' basic body plans, the family relationship between humans and chimpanzees, the family relationship between English and Hindi, undersea volcanoes, tectonic plates, the moons of Jupiter, the rings of Saturn, the Andromeda galaxy, black holes, the atomic structure of gold, the difference between carbon and diamond, the function of the heart, and the architecture of neurons. In each case, the weight of observable fact came to overpower differences in plausibility rankings. Contrary opinion became quirky, then eccentric, then laughable. Evidently, Baconian convergence is real and well within scientists' reach.

Some circumspection is required. It is not always clear when convergence is complete. For a long time it seemed that Newton had gotten gravity basically right; then Einstein came along to upend that complacency. A new quantum theory of gravity might yet show that even Einstein did not have the full story. And science's truthward trajectory is often not a steady march. In the course of finding correct explanations, opinion may temporarily move away from literal truth, as the science of heat moved in the nineteenth century from the kinetic to the caloric theory. Baconian convergence is, in the short or even the medium term, a fitful process.

Furthermore, the observable facts themselves are not quite the bedrock that Bacon takes them to be: published evidence can be "bad" in various ways. Often, due to some problem with the instruments, it is misleading. The plates from Eddington's Brazilian astrographic telescope, for example, wrongly suggest a Newtonian gravitational bending angle for light. Sometimes evidence might even be manipulated, as historians have suspected of "facts" reported by Mendel, Haeckel, Millikan, and Newton. If Baconian convergence is to occur, science must somehow neutralize bad data.

Theorists who say that science is "self-correcting"—I quoted both Karl Popper and Atul Gawande to this effect in Chapter 2—believe that when it really matters, bad data is eventually recognized as such. In a sense, that is indeed what happened with the photographs from the Brazilian astrographic: now that we know that Einstein was right (or at least, more right than Newton), we can look back and say that the astrographic setup must have been faulty. But such reasoning is possible only because the scientific community has already converged on and endorsed general relativity. Corrections of this sort cannot, therefore, explain how convergence occurs in the first place.

What happened in the case of general relativity, and what tends to happen in science more generally, is that opinions converge not because bad data is corrected but because it is swamped. After Eddington, many determinations of the gravitational bending angle were made and many other tests of relativity conducted. The great preponderance of the resulting measurements fit Einsteinian physics better than Newtonian physics. Eddington's data, right or wrong, simply ceased to matter.

So we see that over the decades and centuries, agreement emerges again and again from the tissue of uncertainty and dissent that is the characteristic stuff of science—provided that everyone keeps arguing and the stock of observations continues to grow.

Written by Sir Francis Bacon, *Romeo and Juliet* would have had a happy ending. For all their rivalries, their contrasting temperaments, their contrary theoretical tastes, their disparate explanatory standards, Professors Montague and Capulet would be pressed together, eventually, by sheer weight of evidence. Their dialogue, strictly schooled by science's iron rule, would line by line, fact by fact, bend their wills to the world and so bring their minds at last into perfect congruence.

Two principal parts of my explanation of science's power, of my contribution to the Great Method Debate, are now in place: the iron rule's procedural consensus and the Baconian convergence that the consensus makes possible. There is one further element to the story.

In a basement in Ohio in 1887, Albert Michelson and Edward Morley sent two beams of light flying. One beam was traveling in the same direction as the earth's motion around the sun; the other was traveling at right angles to the first. After traversing identical distances, the beams were reflected back to their source and compared. If their waveforms, when superimposed, failed to overlap exactly, it was because one was taking longer to make its journey than the other—

which is precisely what Michelson and Morley expected. The aim was to measure the speed with which the solar system was traveling through the "ether," an invisible substance hypothesized by nineteenth-century physicists to be the bearer of light waves in the same way that water is the bearer of ocean waves and air is the bearer of sound waves. The faster the movement through the ether, the slower the total travel time of the beam traveling in line with the earth, so the greater the discrepancy between the two superimposed beams. One momentous consequence of such a measurement would be direct evidence, at last, for the ether's existence.

It was a finicky business. The distance by which one beam would be shifted relative to the other was expected to be on the order of a few hundred nanometers, or about 0.00001 inches. The smallest vibrations—passing horses or distant claps of thunder—could upset the measurement device. That is why it was situated in a basement, mounted on a huge sandstone block, floating in a trough of mercury (Figure 5.1). That is also, perhaps, why Michelson had, partway through the 10 years

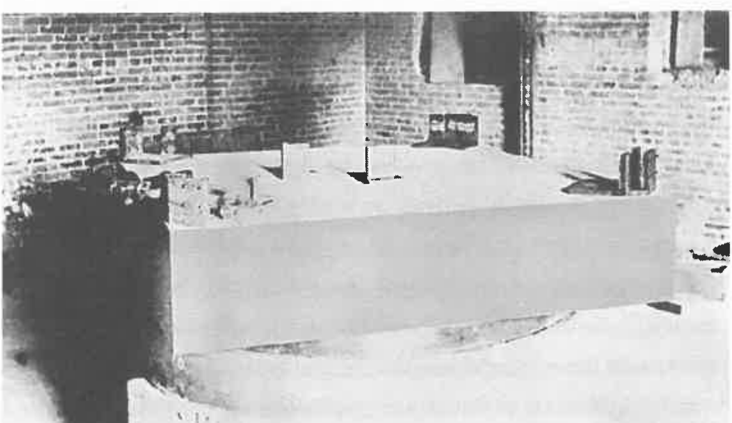

Figure 5.1. The apparatus that Michelson and Morley used to measure the speed of light in various directions relative to the earth's direction of travel.

he devoted to building ever more sensitive mechanisms to determine light's speed, suffered a nervous breakdown—so conjectured his collaborator Morley.

The end result was, as Michelson put it, "decidedly negative": the experiment had failed to detect any motion at all with respect to the ether. For a time, there was contention and confusion about both ether and the experimental setup. It was cured by Einstein, whose special theory of relativity dispensed with the ether and explained exactly what Michelson and Morley had unwittingly observed: that the velocity of incoming light is the same regardless of whether you are stationary with respect to the light source or traveling toward it at high speed. After two hundred glorious years, Newtonian physics was dead. That is the difference made by 0.00001 of an inch.

As you have already seen, Einstein's treatment of gravity was tested a few years later by scrutinizing similarly tiny details: Eddington's eclipse experiment measured a shift in the apparent positions of stars equal to about one-third of the apparent diameter of Mars at its smallest. The Gravity Probe B satellite, also testing general relativity, was designed to pick up discrepancies of about 0.00001 degrees per year. Newtonian and Einsteinian physics tell vastly different stories about the structure of the universe. But to see which of them is correct, you need to make almost infinitesimally small measurements.

That the secrets of the universe lie in minute structures, in nearly indiscernible details, in patterns that only the most sensitive, fragile, and expensive instruments can detect, is an insight so important that it deserves a name. I call it the *Tychonic principle*, after the sixteenth-century Danish astronomer Tycho Brahe, who, working just before the invention of the telescope, was the last and greatest "naked-eye" astronomer, using sextants and quadrants to pinpoint the positions of stars and the movements of planets to within 0.02 degrees. To achieve this level of accuracy, Tycho built an observatory called Stjerneborg ("castle of the

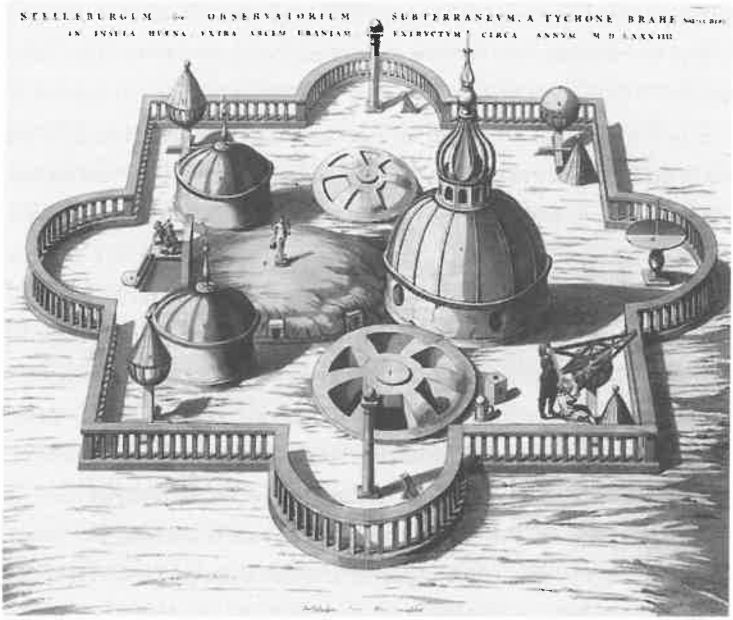

Figure 5.2. Stjerneborg, Tycho's subterranean observatory.

stars") entirely at basement level, seeking as Michelson and Morley later would to shelter from the inaccuracy inflicted by the bustle and noise of the world outside (Figure 5.2).

The Tychonic principle holds not only in fundamental physics and not only for strictly quantitative detail. To discover the intricacies of biological heredity and evolution or the way that multicellular organisms develop and grow from embryos, researchers must attend painstakingly to complex causal structures at the microscopic level. Similar attention to exact amounts and elaborate connections is required of neuroscientists, geologists, and archaeologists, of model builders in climate science and economics, and in a somewhat different way of anthropologists and sociologists.

Let me test your imagination with an outré thought: the Tychonic principle need not have been true. We might have lived in a universe where the fundamental laws of physics exerted only a rather slack control over the trajectories of particles, determining their movements to one or two decimal places but no further. Making measurements to the sixth decimal place in such a world would be a waste of time and money, no more conducive to knowledge than counting pebbles on the beach or drops of water in the sea. Equally, we might have lived in a universe where the forms of organisms were determined by vital spirits, or by the whimsy of the gods, rather than by intricately structured chains of molecules. But we don't; our world, it seems, is Tychonic through and through.

Bacon proposes that if we take explanatory power to be our guide, then empirical testing will ultimately single out the truth. The Tychonic principle says that much of that testing will be extraordinarily difficult to conduct. It is for this reason that the iron rule is essential to modern science's success. The effort required to build a store of observable fact sufficient for Baconian convergence in a Tychonic world is so great that humans can be persuaded, induced, or impelled to take on the project only under exceptional circumstances. The iron rule engineers those circumstances and so "forces scientists to investigate some part of nature in a detail and depth that would otherwise be unimaginable." Those are Thomas Kuhn's words. He attributed the compulsion to paradigms, not to a universal procedural consensus—otherwise, this was the best sentence he ever wrote.

THE IRON RULE tells scientists to pursue the truth by looking for the theory that best explains the observable facts. What is so groundbreaking, so revolutionary about that? Very little, some might say. We humans have known how to infer from what we see to what best explains it since our Paleolithic ancestors first grasped the significance of saber-tooth footprints in the snow.

Further, the value of this explanatory way of thinking was as obvious to ancient and medieval thinkers as it was to any prehistoric hunter. The first Greek philosophers—Thales, with his theory that everything is made of water; Heraclitus, who preferred fire; Anaximander, who thought that the fundamental constituent of things is the "boundless"—were all doing their best to account for what they saw in the world around them: rainbows, magnetism, fossils, epilepsy, the saltiness of the sea, the starriness of the sky, the exquisite adaptedness of living things to their natural habitats. Aristotle single-handedly added a dazzling array of phenomena to this list, from tornadoes to respiration to human hands and waterbirds' webbed feet, and offered systematic theories of physics, biology, and psychology to explain them. Indeed, he is sometimes held up as the first great scientist in history.

The eleventh-century Islamic philosopher Ibn Sīnā, called Avicenna in the West, laid down seven rules for medical experimentation. The English scholastic philosopher Robert Grosseteste developed the notion of a controlled experiment in the high Middle Ages, centuries before Bacon. And experiment was not only praised in principle but deployed in practice. The historian David Lindberg provides a partial list of experimenters at work before the Scientific Revolution that stretches from the ancient Roman Empire to Islamic Persia through the European Middle Ages, including, among many others, Ptolemy, Ibn al-Haytham (Alhazen), Kamāl al-Dīn, Theodoric of Freiberg, Rabbi Levi ben Gershon, Johannes de Muris, and Paul of Taranto.

What does the iron rule add to all this? Where is its novelty? On what grounds can I say that Aristotle was not a modern scientist when he manifestly valued the power of his theories to explain the phenomena that he so keenly observed?

For all its emphasis on explaining the world and its openness to experimentation, the natural philosophy of old could not establish the procedural consensus responsible for modern science's superlative knowledge-making power. To formulate the iron rule and thus to make modern science, something had to be added to the age-old doctrine that true theories can be recognized by their explanatory power. That supplement takes the form of four methodological innovations.

The first innovation is a reformulation of the material that would make up the iron rule—explanation itself. Before the Scientific Revolution, explanation was mixed with philosophical principles, and so explanatory power was a subjective matter, varying with the beholder's intellectual commitments and temperament. In modern science, the notion of explanation is free of philosophy and all other ideology; it is as pure as the elemental metal from which the iron rule takes its name. Consequently, explanatory power means the same thing to every modern scientist, regardless of upbringing and inclination, so that every scientist agrees on what the iron rule says, on what satisfies the rule's criterion for a meaningful scientific test.

The second innovation is a matter of place, of venue, of domain. The iron rule is focused not on what scientists think, like the traditional laws of logic formulated by Aristotle, but on what arguments they can make in their official communications. Their brains are left unfettered by the rule, then; it is their public pronouncements alone that are subject to its rigid specifications. As you will see, it is this innovation that makes possible both the third and the fourth innovations.

The third innovation is a special kind of objectivity that can be imposed on scientific debate without being upended by the essential subjectivity of scientific reasoning—a kind of objectivity that is consistent with everything I have said in previous chapters about scientists' moral frailty and the multifariousness of their plausibility rankings.

The iron rule's negative clause is its fourth innovation. The natural philosophers cared a great deal about their theories' power to explain natural phenomena. They also cared about their theories' philosophical integrity, theological purity, and formal beauty, and they were ready and

The Four Innovations That Made Modern Science

1. A notion of explanatory power on which all scientists agree
2. A distinction between public scientific argument and private scientific reasoning
3. A requirement of objectivity in scientific argument (as opposed to reasoning)
4. A requirement that scientific argument appeal only to the outcomes of empirical tests (and not to philosophical coherence, theoretical beauty, and so on)

willing to make a case for their views from every one of these perspectives. The iron rule, however, permits nothing but matters of explanatory power, nothing but a theory's ability to account for the observable, to determine the course of scientific argument. Theology, philosophy, even beauty are strictly off limits. Scientists, if they choose to dispute, are obliged to do so in the empirical manner, by running tests in accordance with the rule's prescriptions.

In this chapter and the last, you have seen how modern science's procedural consensus, established by the iron rule, spurs the creation and compilation of the sort of rich, intricate, revealing data that has enabled scientific progress in a Tychonic world. It is now time to see how the four innovations invest the iron rule with its ability to create that consensus.