CHAPTER 6

EXPLANATORY ORE

How explanation became objective, yielding the material to forge an iron rule that says the same thing to every scientist (the iron rule's first innovation)

THERE IS AN undiscovered island somewhere in the southern oceans, a scrap of land whose people have yet to make contact with the modern world. Call it Atlantis. If it does not exist, invent it: it is needed for philosophical purposes.

The Atlanteans, though few in number, have a scintillating civilization. Poets, architects, playwrights, and mathematicians mill about their court and their marketplaces, along with natural philosophers who aim to understand the workings of the visible world: the planets, light, falling objects, living things. But like the ancient Greeks—or the Ming dynasty Chinese or the Elizabethan English—the Atlanteans lack modern science.

After one navigational error too many, you find yourself washed up on the Atlantean shores. Your hosts are thrilled and perplexed by their mysterious visitor. You tell them about vaccines, antibiotics, and laparoscopic surgery. You show them your smartphone—a remarkable object, though the signal strength is questionable. You describe a world of effortlessly abundant food, half-mile-high towers, and flying machines.

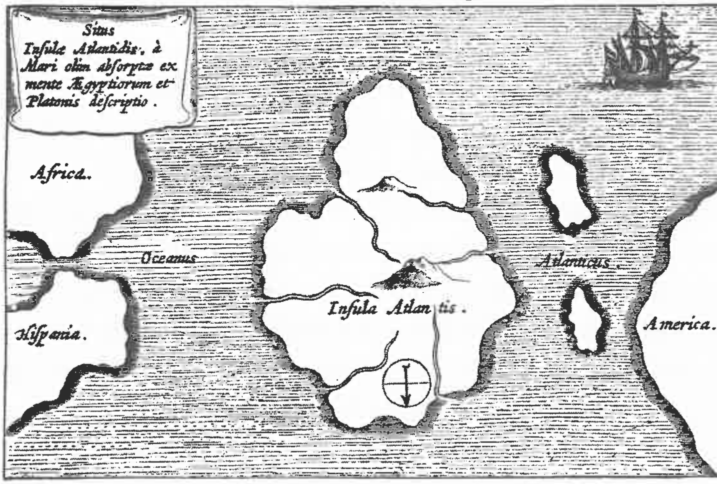

Figure 6.1. The location of Atlantis, as imagined by the scholar Athanasius Kircher (1602–1680). He places it a little too far to the north.

Better still, a world that understands the true nature of light, the principles that move the planets, and the code that programs life. The Atlanteans want it all. Can you perhaps teach them to do science?

Why not? It will help pass the time. The Atlanteans bring out their natural philosophers, eager pupils in the inaugural session of your first seminar: "Introduction to Modern Science."

Day one: the iron rule. There is no need to persuade the Atlantean philosophers that they should search for theories of great explanatory power; they've been doing exactly that for centuries. What's news is that exploiting this insight to make scientific progress will depend far more heavily on exacting measurement, on painstaking laboratory toil, than on deep thought. In our Tychonic world, it is by accounting for the fraction of an inch, the sliver of a degree, that the truth will make itself known.

Day two: you make an astounding discovery. The Atlanteans consider a theory to explain a fact whenever that fact rhymes with the theory. According to these instinctive poets of nature, the hypothesis that "things made of earthly matter seek to find the center of the universe" explains why "he who topples from a lofty crag comes out much the worse" but not why "straws tossed into a storm fly near and far away," since *worse* rhymes with *universe* but *away* does not.

What could the Atlanteans be thinking? Perhaps they imagine that the world is run according to the dictates of a "Great Poet in the Sky." This supernatural versifier has laid down certain edicts, such as "Things made of earthly matter seek to find the center of the universe," and ensures through direct intervention in the world that events occur only if they rhyme with the edicts. Although this is a bizarre view, it is not incomprehensible. "I see," you say to the Atlanteans. "Like us northerners, you Atlanteans believe that explaining an event is a matter of finding out what made it happen. But unlike us, you believe in gods who make things happen by way of rhyming rather than gravity."

No, that's not it at all, say the Atlantean philosophers. There is no celestial sonneteer. Indeed, there is nothing that "makes things happen." The world simply is. But its state of being can nevertheless be understood, its workings can be explained, by grasping how it all hangs together—as a poem. For the Atlanteans, understanding the universe is a matter of literary interpretation.

Bringing science to Atlantis is going to be harder than you supposed. You clear your throat. There are a few things you need to know about explanation, you say. Or at least about the kind of explanation that supplies the raw material for the iron rule. Most important of all, explaining events is a matter of finding their causes, not of rhyming. You cannot discover true theories by scrutinizing word endings! The Atlantean philosophers mutter among themselves, casting occasional pitying glances in your direction: such splendid technology; so little understanding. One speaks up: "No, I'm afraid it is *your* theories that must be incorrect." Who are the teachers and who is the pupil?

If only you could open your eyes and . . . ah, the real world! Electric lighting, anesthesia, and broadband internet! It was all a dark fantasy. Wasn't it?

It is not so far from the truth, contend the many historians of science who look back in time and see an archipelago of ideas about understanding, a chain of cultural islands stretching over the centuries, each with a proprietary and parochial concept of what it is for a theory to provide an adequate explanation of patterns of everyday events and experimental outcomes. Each era has its own rules for making sense of the world; there are, writes the contemporary historian Peter Dear, "no timeless, ahistorical criteria for determining what will count as satisfactory to the understanding."

Cast your mind back to Thomas Kuhn, early in his career, attempting to comprehend Aristotle's physics. He encountered some peculiar ideas. Aristotle regarded an object's location in space as though it were a property like color; moving, then, is a bit like changing color. As a consequence, there cannot be an empty location in space: locations are properties of things, so where there is a location, there is a thing, in the same way that where there is a color there has to be a thing to bear that color. In fact, according to Aristotle, there is no such thing as space; there are only "places," and they are all filled. (He was perhaps the first theorist of inner-city parking.) Further, characteristic patterns of motion, such as heavy things' tendency to fall to earth, are given an explanation that seems biological or even psychological: the "natural end state" for the heavier varieties of inert matter is to occupy the exact midpoint of the universe. All things, animal, vegetable, and mineral, seek out their natural states. Thus, inert matter heads directly for the center of the world, which is also the center of the earth. None of this, at first, made any sense to Kuhn; Aristotle's claims seemed so obscure as to be unintelligible.

Then he looked out the window and everything, as he said, "fell into place." He achieved the ability to see the world as Aristotle did, to operate with Aristotelian principles for explaining the world—principles that were so strange as to be almost Atlantean. "Now I could understand why [Aristotle] had said what he'd said." That experience inspired Kuhn's conception of scientific progress as a series of leaps from framework to framework, with each such "paradigm shift" bringing a new explanatory system that mystifies scientists educated in an older way of looking at the world.

A paradigm is more than an explanatory framework; it is a complete recipe for doing science, including goals and methods as well as modes of understanding. Modern historians are skeptical, as you saw earlier, that scientists' behavior conforms to any schema as totalizing as a Kuhnian paradigm. But like Peter Dear, many hold on to the explanatory component of Kuhn's picture even as they discard the rest, endorsing along with Kuhn an idea that might be called *explanatory relativism*. The thesis of explanatory relativism says that each age or school has its own standards for explanation and understanding and that one age's standards typically make little or no sense to the thinkers of another age.

In this formulation, "make no sense to" does not mean merely "appear implausible to." It means that explanations from another time do not seem to be explanatory at all—just as we consider the Atlanteans' rhyming scheme to be, however beautiful, quite incapable of conveying true understanding of why things happen. Untimely explanations, then, are off the table altogether, disqualified from inquiry before the game begins. Only a skilled historical interpreter, as Kuhn took himself to be, can appreciate their virtues, and even then perhaps only by way of some unbidden, almost mystical epiphany.

If the thesis of explanatory relativism is correct, then there can be no consensus, over time, about different theories' comparative explanatory success, since "success" means something completely different for each theory. In that case there would be no way to establish the sort of procedural consensus that, I have argued, turns the crank of the knowledge machine. Explanatory relativism is therefore incompatible with the iron rule; were it to be true, modern science could not exist.

It once was true. But then, during the Scientific Revolution, it was deposed. This chapter will explain how relativism undermines consensus and how it came itself to be undermined, clearing the way for a new form of inquiry governed by the iron rule.

IT IS THE fourth century BCE. You are wandering an island less fictional than Atlantis—ancient Greek Lesbos—in the company of Aristotle himself, observing the diverse aquatic life in its lagoon. (In the 340s BCE, the great philosopher spent a number of years on Lesbos and in the nearby mainland city of Assos, researching his treatises on biology.) There is a great deal on which the two of you might, by close and regular observation, come to agree: that some fish make grunting or piping noises, that male octopuses have a specialized tentacle for copulation, that dolphins do not lay eggs but rather give birth to live young, that the lagoon is rich in oysters, and that in winter many of its fish migrate into the warmer open sea. You might dissect a cuttlefish together, as we know Aristotle did, and find that the intestinal tract runs from the mouth to the stomach, then loops back to exit just below the head.

When you come to explain the cuttlefish's anatomy and behavior, however, the common ground beneath your feet starts to shift uncertainly. What does the cuttlefish heart do? You might think that it pumps blood, but Aristotle has other ideas: the heart is the primary source of blood, but also the seat of sensation—pleasure,

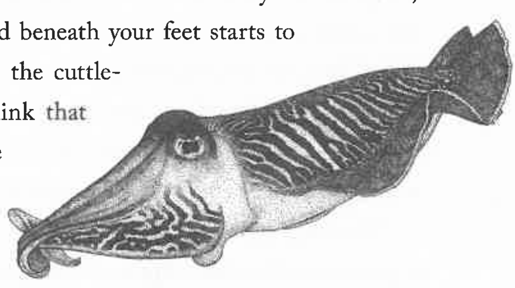

Figure 6.2. The common cuttlefish, *Sepia officinalis*.

pain, and sense experience—and in higher animals, the organ of thought. (On this matter Aristotle was at odds with many of his contemporaries, who located thought in the brain.) That is a disagreement, but no different from Capulet and Montague's disagreement about the nature of heat. Surely empirical testing will decide what's correct.

To help the process along, you have brought a surprise for your philosophical companion: a scalpel, dyes to accentuate cell walls and other biological boundaries, and a microscope. With the help of these newfangled instruments, you'll be able to understand the operation of a cuttlefish heart by examining the microscopic structure of its parts and figuring out how that structure, guided by the mechanical principles of physics, causes the behavior of the whole.

"Interesting," says Aristotle, stroking his well-tended beard. "But like the ideas of those old materialists Empedocles and Democritus"—here he refers to his predecessors and rivals, now long gone—"utterly implausible." He cocks a critical eyebrow. "Their picture of the world was so simple. And so simplistic: everything we see, they taught, is caused by the blind laws of physics pushing around dumb matter. If they were right, we would find monstrousness and malfunction everywhere. Perhaps just by chance an occasional collection of mechanical parts would work together like an organic body, feeding and breathing and turning out little duplicates of itself. But these would be by far the exceptions among the chaos." With a sudden, uncharacteristically loose and expansive gesture he takes in the hills, the water, the entire island. "Look at all this life! See how beautifully organized it is. Every plant, every animal, every branch, every limb, every organ, works carefully, regularly, precisely in its own way to thrive in its natural place in the world. It is anything but accidental. Physics alone cannot explain life."

The only way to explain the cuttlefish's characteristic form and activities, he continues, is to posit a property of the organism as a whole, its *psuche*, often translated as "soul." It is not what it sounds like: Aristotle's *psuche* is not an immaterial thing detachable from the animal in question. But nor is it a material part of the animal, like a leg, a brain, or a heart. It is rather something that subsists in the animal and causes it to have its overall form. The nature of the *psuche* is to organize things for the ultimate good of the animal. Thus, the animal's parts operate so as to perform functions that, working together, cause the animal to grow, flourish, and proliferate.

Every modern interpreter of the Aristotelian *psuche* finds in it a puzzle. It is, says Aristotle, the form of the animal in its body, which suggests that the *psuche* is simply the physical structure of the creature. That cannot be what Aristotle means, however, or else he is no better off than the materialists he criticizes, attributing the viability and efficiency of organisms to the fact that they happen by chance to have the ideal physical organization for the job. At the same time, it seems, the *psuche* is not something above and beyond the body. It is the physical form of the body, yet it also explains and sustains that form.

Reach for an understanding of the *psuche*, and it flits away into the depths. It will not be captured by a mind like yours. That is the characteristic symptom of explanatory relativism. The only intelligible explanation of life for Aristotle is to you unintelligible—barely better than the works of the explanatory poets of Atlantis, who sit in their libraries pondering rhymes for "cuttlefish."

So there stand two figures at odds on the shores of the lagoon: you with your microscope and Aristotle with his conception of natural order. You want to use your microscope to settle your differences. It will reveal a physical structure in cuttlefish tissue that is indiscernible to the naked eye, you say, yielding invaluable clues about the workings of the heart and other organs. "Futile," responds Aristotle. It is not that physical structure has no role to play in bodily functions (although Aristotle does in fact deny that organic tissue has microscopic structure). It is rather that physical structure is not at all the right kind of thing to account for biological function. Even if it were to explain how a cuttlefish moves, it could not explain why the movement is tuned so finely to promote life. Only *psuche* can do that.

In Aristotle's eyes, your program of inquiry—slicing cuttlefish parts into finer and finer pieces to hold up to the light—is not sophisticated science but plain butchery. Aristotle's program, meanwhile, looks not much better from the perspective of your own explanatory commitments. How can you detect a *psuche* and probe its properties and operations? Or more exactly—since it is easy to detect the physical form in which the *psuche* inheres—how can you detect and measure the explanatory aspect of the *psuche* that goes beyond mere physical form? To Aristotle, the answer is obvious: the appearances and activities of every organism are explained by its *psuche*. So evidence for the nature of the *psuche* is anywhere that a biologist might care to look. To you, with a different notion of biological explanation, that sounds not so much like empirical inquiry as an exercise in begging the question.

A disagreement about explanatory systems thereby leads to intellectual impasse. For Aristotle, the microscope is no more useful for biology than for astronomy: what it might detect, if anything, can be seen in advance to lack the power to explain life. For you, the explanatory power of the *psuche* is a mystery. You'd like to use your instruments to get an independent grip on the thing, but the only way to map its structure is through its explanatory ramifications, which are precisely what to you are so obscure.

By fomenting communication breakdowns of this sort, explanatory relativism presents an insuperable obstacle to the kind of procedural consensus essential for modern science, a consensus that assures scientists they are all playing the same game, freeing them from endless debates about the explanatory rules and so directing every last flicker of their investigative ardor toward observation and experimentation. The iron rule therefore outlaws relativism, cutting off all such disagreements by providing a fixed conception of explanatory power to which all scientists must subscribe—

modern science's first great innovation. Because scientists agree on what can explain what, they agree on how to go about testing any hypothesis, even a hypothesis couched in esoteric or controversial terms such as *psuche*, entropy, or the obscure terminology of quantum mechanics.

The explanatory accord was hard-won. To see how it came about, we will have to leave Aristotle on Lesbos and travel to northwestern Europe, two thousand years later. There we will entertain the ideas of two superlative thinkers. One was the last of the great natural philosophers, devising theories of matter and its behavior that conformed to an idiosyncratic, metaphysically regulated scheme of explanation yoked to the zeitgeist. The other destroyed these theories, and in so doing, shattered the dominion of philosophy over scientific explanation and overthrew explanatory relativism itself. The arc curving from the one intellectual life to the other is the story of the smelting of the new explanatory iron.

In 1618, as Francis Bacon worked on the manuscript of his *New Organon*, a young French student turned his back on law school and joined the Dutch army just as war erupted on the other side of Europe in Bohemia. Within a year he had left the Dutch and joined the allied forces of Maximilian of Bavaria, a decision that took him to the German states that made up the bulk of the Holy Roman Empire. There, most of the fighting and most of the dying in this devastating conflict would take place. By its end, the Thirty Years' War would have killed nearly half of the population in Bavaria, a third in Bohemia, and at least a fifth of the people of Central Europe as a whole—a greater percentage than any other European war in history to this day.

In the dull and peaceful months before the German princes were drawn into the struggle, the French soldier found himself at a loose end while billeted in the city of Ulm, on the Danube in what is now southern Germany. It was November; he was cold. He shut himself into a wellheated room and began to think; in the course of his cogitations, he later related, he "discovered the foundations of a marvelous science."

That night he dreamed three times. First came a dream of fear: specters, paralysis, a blasting wind in which he could not walk upright. Second came a dream of undirected power: fire and thunder. Third came a dream of books and learning, of the beginning of a journey, of the unity of all knowledge. In the following days, he resolved to devote his life to building his marvelous new science. René Descartes became a philosopher.

In short order, Descartes left the army and so left the war. He lived as a recluse in France and the Netherlands for the next three decades, while Europe chewed itself up around him. In his solitude, he wrote about mathematics, physics, philosophy, God, vision, thought, and, in the end, "the passions of the soul." He explained depth perception, the refraction of light, and the nature of the rainbow, and he pioneered the use of what we now call Cartesian coordinates in mathematics. ("Cartesian" means "proposed or invented by Descartes.") What was perhaps his grandest and most ambitious project, however, was also one of his first: a description of the structure of the universe, which he finished in 1633. Descartes called this book *The World.*

All the activity that makes up the world, proposed Descartes, from the simple orbits of the heavenly bodies to the intricate functioning of the human body, is the movements of chunks of matter, and these chunks interact in one way only: by direct contact, or in other words, by pushing each other around. The planets may look like they are out there hurtling through empty space. But it can't be so. If something moves, it is because something else forces it to move, and the sole way to transmit such a force is direct physical contact. The planets maintain their orbits around the sun, Descartes hypothesized, because what appears to us to be empty space is full of imperceptibly tiny particles that are themselves revolving around the sun, carrying the planets with them.

The universe, Descartes ventured, is packed with giant globules of matter rotating in this way; at the center of each is a sun around which its own proprietary planets revolve. To us, these suns are the distant stars (Figure 6.3).

Gravity on earth, since it is a matter of force, must also be caused by something in direct contact with the bodies that experience gravitation.

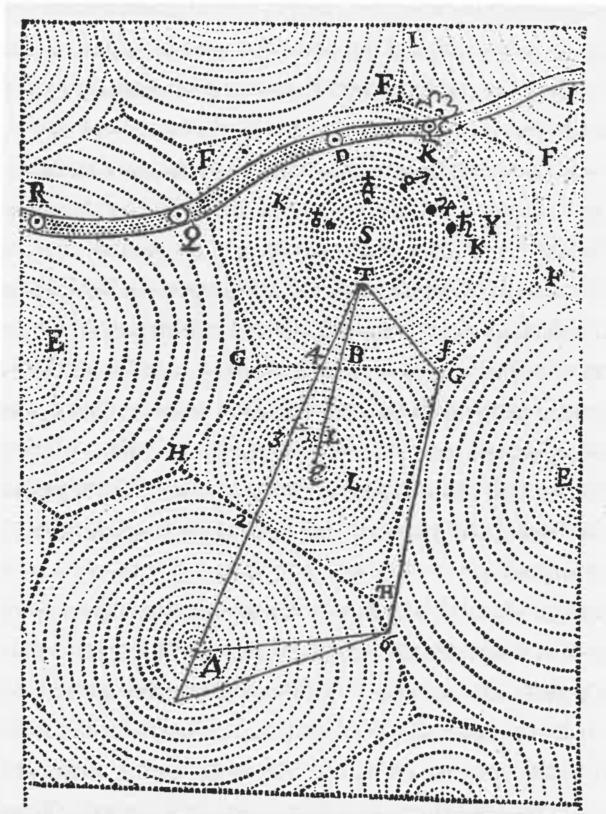

Figure 6.3. Descartes's universe. Each of the polygonal cells is a great rotating globe of matter with a star at its center. Our sun is marked *S*; the planets are labeled with their astrological signs (♂ for Mars, ♃ for Jupiter, and so on). A comet (⊙) blazes an irregular path through the "north" of the solar system.

Descartes dreamed up the following intricate story. All bodies, he reasoned, have a tendency to "fly away" from the earth by a kind of centrifugal force. Because air particles are more "agitated" than the particles of which heavy objects, such as rocks and humans, are composed, their centrifugal tendency is more powerful: your body is trying to fly away into the sky, but the air around you is trying even harder. Consequently, if you are not on solid ground, the air below rushes up past you, pushing you down into the space that it previously occupied; you fall, then, because the air is pushing you downward, and this downward pressure is greater than your body's tendency to move upward. The same mechanism keeps you from rising in the first place. It is as though you are in a crowd shoving your way toward the exit to escape a fire, but because the other members of the crowd are stronger and more determined than you, they elbow their way past, keeping you firmly in place—clamped to the ground.

The Cartesian explanation of gravity is ingenious but baroque, and indeed, rather awkward. If the awkwardness embarrassed Descartes, he didn't show it. Perhaps that was because he regarded his collision-driven physics as far more than a plausible empirical proposal. It had to be true, he believed, because he could prove that the world couldn't possibly work in any other way. His demonstration was not scientific in our modern sense, but rather philosophical and theological. It began with a metaphysical argument.

Space, he claimed, is by its very nature extended, that is, spread out. Empty space would therefore be an extended nothing, but "nothingness cannot possess any extension." It follows that there can be no empty space; space is necessarily filled with matter (as Aristotle, too, maintained). Indeed, in some sense, according to Descartes, matter and space are one and the same thing, described in two different ways. "It is no less impossible that there be a space that is empty," he wrote, "than that there be a mountain without a valley."

If space is completely packed, you might think, nothing could ever get started. But when God created the universe, said Descartes, turning from philosophy to theology, he set all its matter in motion, creating the great spinning globules pictured in Figure 6.3. From that point on, things followed the principles of Cartesian physics. The matter in a globule continued to rotate around its central point unless other matter got in the way—in which case there was an impact having the power to send the colliding matter off in new directions. (Because all space is filled with matter, Descartes's collisions are more like jostles and shoves than projectile impacts. I think of certain New York City subway lines at rush hour.)

Every departure from a steady state of motion, then—every meaningful change—is caused by direct contact. This is true, needless to say, of manifest collisions, such as a cannonball's crashing into a fortress wall. But with God and philosophy on his side, Descartes sought to show that change works that way for everything else as well. Gravitational attraction, for example, must somehow be a matter of our being pressed directly to the earth, even if there is no tangible sign of anything doing the pressing. Descartes therefore concocted the story in which gravity is effected by small particles pushing aside larger ones in their skyward stampede.

Likewise, visual perception must be powered by collisions. Light strikes the retina, Descartes proposed, which then signals the brain by way of some intricate hydraulic machinery that pumps "animal spirits" around the body. Collision underlies all other bodily functions for the same reason. Of necessity, every biological process is mechanical, and every mechanism is implemented by flows of matter interacting by way of impact.

The Cartesian and the Aristotelian views could hardly be more different—in their details but even more so in their fundamental explanatory principles. According to Aristotle, as we saw, each kind of matter has a natural motion. The natural motion of heavy matter, of things made of the elements earth and water, is toward the center of the universe, that is, toward the center of the globe on which we live. The natural motion of the stars and planets, which are made of a fifth element—closer to divinity than air, fire, earth, and water—is the circle. Thus, metal falls to the ground while the moon eternally orbits the earth; in both cases, they do what it is in their nature to do, unaided.

Descartes explains these same motions in terms of pushes and shoves. Even the planets, set on their way by God, require the pressure of other, unseen particles to keep them on their circular paths. To posit any other cause of motion, Descartes thinks, falls short of true science. Close examination of such a hypothesis, he would say, will reveal links in its proposed explanation that are no better than rhymes—pleasing to the mind, perhaps, but failing to establish a genuine explanatory connection between cause and effect.

The same goes for plant and animal biology. Descartes conceives of organic bodies as machines, while for Aristotle, as we saw, organisms are distinguished from machines by their possession of a *psuche*, an explanatory principle that subsists in the body's physical structure and explains the regularity and purposiveness of the actions brought about within and by that structure.

Descartes dismisses Aristotle's explanations as unintelligible; Aristotle would dismiss Descartes's as hopeless. To each the other's proprietary explanatory framework looks confused, empty, incomprehensible, absurd—as intellectually alien as the Atlantean rhyming scheme.

Yet at the most abstract level, Descartes and Aristotle have something of great importance in common. Both call upon higher powers—most notably, the power of philosophical reason—to determine what can legitimately explain the movements of matter and living things. It is this philosophizing that digs them into diametrically opposed positions. Their explanatory principles settled, their theories of physics and biology can only remain perpetually at odds.

What next? If Aristotle had surveyed the intellectual battlefields of the 1600s, he would surely have exclaimed, "Not again!" The atomists and other materialists, vanquished in fourth-century Athens, had returned and must be fought once more. It was just as Bacon said: the same arguments, going round in circles forever.

And so it might have continued, had not something new appeared to break the cycle and to direct the investigation of nature onto a wholly new trajectory, along which profound disagreements about explanatory standards would become quite unknown.

DESCARTES DIED IN 1650. The war in Europe was over. Germany lay in ruins; France was triumphant; England, having executed its king, Charles I, was a republic. Over the silent battlefields settled a peace born of exhaustion. In the universities and learned societies, by contrast, the gravity wars were poised to begin.

Isaac Newton was 7 years old. By the age of 11, he was ranked the second worst pupil in his class at King's School in Grantham, Lincolnshire. At some point, however—possibly in reaction to his bullying by the third worst pupil—he threw himself into his studies, which drew him to Cambridge University's Trinity College. Like Descartes, he preferred his own company, the better to work without interruption: "Truth is the offspring of silence and unbroken meditation." Unlike Descartes, who died in Sweden, he never left the country of his birth.

Indeed, Newton barely left the university, and he was still at Cambridge when, 25 years later and now the Lucasian Professor of Mathematics, he published his greatest work: the *Mathematical Principles of Natural Philosophy*, usually called (after its original Latin name) the *Principia*. The world had seen nothing like Newton's system before. It had the scope of the most ambitious physics, such as that of Aristotle and Descartes, the mathematical exactitude of the best astronomy, such as that of Ptolemy and Kepler, yet its core could be written down in a handful of simple formulas: Newton's three laws of motion and his law of universal gravitation.

Whereas Descartes held that all physical causation is by direct contact, Newton appealed to a "force of gravity" that pulls one object toward another massive object, apparently without any intervening mechanism. Many of his contemporaries saw in Newton's new force a gaping theoretical hole. How could gravitation be taken seriously unless there was a physical means, such as currents of interstellar fluid or collisions with tiny particles, to make it happen? Causal forces between planets made no sense without a tangible thing or stuff connecting them, communicating a pushing or a pulling or some other kind of visceral "oomph." The notion that causation could happen otherwise—that one planet could simply act on another at a distance, as though extending a ghostly grip of arbitrarily long range—seemed preposterous. The mathematical and metaphysical prodigy G. W. Leibniz, writing around the time that Newton issued the second edition of his magnum opus in 1713, dismissed Newton's theory as "barbaric physics," testimony to the sad fact that "people love to be returned to darkness."

In the decades after Newton, however, the idea of "action at a distance," of celestial bodies influencing one another's movements across empty space with no connecting causal links, became familiar and acceptable. If anything, it was causation by collision that seemed suspicious. Why should two objects, coming into contact, refuse to interpenetrate? Some sort of repulsive force, operating at a very short range, must prevent them from overlapping. If so, what appear to be genuine impacts are in fact merely unconsummated flirtations with physical contact: particles do not literally touch but at most come extremely close before irresistible countervailing forces fling them apart. When the great German philosopher Immanuel Kant wrote about physics a century later, in 1786, he claimed that it was incomprehensible that objects could act on each other except by way of brute forces, projected across space in accordance with something like Newton's law of gravity.

As I have related the story so far, this sounds like yet another chapter in the chronicle of explanatory relativism, yet another example of intellectual history as a series of discontinuous jumps from one explanatory framework to another. In the seventeenth century, causation had to be communicated by contact; anything else was darkness. Around 1700, according to the relativist story, Newton successfully pushed Descartes's conception of legitimate explanation aside, substituting his own—just as Descartes had earlier routed the Aristotelian approach to explanation. From then on, contact looked philosophically incoherent; action at a distance became the only logically acceptable way to make things happen.

That's how Kuhn spins the history of gravity in *The Structure of Scientific Revolutions*. But his relativist interpretation is wrong. Newton did not change his age's conception of explanation but rather something deeper still—he changed its conception of empirical inquiry.

To understand how, let us take a closer look at Newton's own interpretation of his method, laid out in a postscript to the *Principia*'s second edition of 1713. There Newton summarizes the fundamental properties of gravitational attraction—that it increases "in proportion to the quantity of solid matter" and decreases in proportion to distance squared—and then continues:

> I have not as yet been able to deduce from phenomena the reason for these properties of gravity, and I do not feign hypotheses. For whatever is not deduced from the phenomena must be called a hypothesis; and hypotheses, whether metaphysical or physical, or based on occult qualities, or mechanical, have no place in experimental philosophy. . . . It is enough that gravity really exists and acts according to the laws that we have set forth and is sufficient to explain all the motions of the heavenly bodies and of our sea.

In this famous passage, the essential aspects of the iron rule make a decisive appearance in world history.

Look at what Newton says about the mechanism for gravitation. You can see at once that he was not an ardent promoter of action at a distance. He allows the possibility that gravity is effected mechanically—by contact or collision—and also the possibility that it is effected by "occult qualities," that is, by some form of nonmechanical causation, such as action at a distance, the kind of thing that philosophers such as Leibniz found to be no more intellectually respectable than witchcraft. The mechanical and occult possibilities are, in this passage, put entirely on a par. Both are mere "hypotheses"; we presently have no way to tell which, if either, is true.

It does not matter. What does matter is that gravity "is sufficient to explain," with great quantitative accuracy, the observed phenomena: the orbits of planets, the paths of comets, the cycle of tides, and the arcs of deadly projectiles. To a philosopher such as Descartes, this modest proposal would make no sense. Gravity cannot explain anything unless it passes the test for being a genuine explainer—the collision test. To put the Cartesian view another way, to explain something means grasping its cause, and to grasp its cause means to comprehend the sequence of pushes or collisions that bring it about. Without that comprehension there is no understanding, no real explanation.

Newton, by contrast, pioneers in this passage a philosophically far shallower notion of explanation, on which a phenomenon is explained simply by deriving it from the causal principles of gravitational theory—that is, from the mathematical principles laid out in the *Principia*. Shallow explanation does not require the explainer to grasp the implementation of the principles. More important still, it does not require the principles to pass any philosophical test or to conform to the explanatory prescriptions of a Kuhnian paradigm. Gravity might turn out to be transmitted by impact, but equally it might turn out to be authentic action at a distance. Either way, Newton maintains, we have a "sufficient explanation" for the purposes of empirical inquiry, or as Newton calls it, "experimental philosophy." Newtonian explainers, like Popperian falsifiers, prove their worth by conforming closely not to intellectual precepts but to observed phenomena. The shallow conception of explanation thereby frees scientific theory builders to try just about anything, however ideologically abhorrent, in their attempts to explain the phenomena.

With this act of liberation, Newton escaped the endless circles of explanatory relativism and gave scientists a "timeless, ahistorical criterion" for explanatory power to serve as the raw material for an iron rule that dictates, in turn, a fixed criterion determining what counts as a legitimate empirical test. A scientific theory postulates some causal principles; what it explains is whatever can be logically derived from those principles. That is a standard for explanation that means the same thing in all places and all times.

To appreciate fully what is objective and what is subjective in scientific practice, however, we need to distinguish more carefully than Newton did between theories and theoretical cohorts. Theories, as you now know, explain very little on their own; very little can be derived from a theory without the assistance of auxiliary assumptions. Einstein's or Newton's theories of gravity, for example, did not on their own predict the positions of the images of stars on Eddington's photographic plates. For that, Eddington needed to add information about the positions of the telescopes, the times that the photographs were taken, the positions of the stars themselves, the assumption that the telescopes were functioning correctly, and certain theoretical posits about light. When surrounded by statements of background conditions and other hypotheses such as these, a theory forms a theoretical cohort and generates specific predictions and explanations.

The iron rule's Newtonian criterion for shallow explanatory power applies to cohorts rather than to bare, unaccompanied theories. It provides an objective and ideology-free test for whether a cohort explains a phenomenon: Can the phenomenon be derived from the cohort's causal facts and principles? Aristotle's centralizing urge; Descartes's democratic jostling; Newton's matter calling to matter across empty space—each qualifies as a causal principle that might make sense of gravitation, and each is warmly welcomed into science. All the iron rule asks is that such principles and their cohorts entail what is actually seen to occur and nothing that is seen not to occur. The modern scientific standard for explanation is as empirically demanding as it is philosophically lax.

Is the ineffable Aristotelian *psuche* scientifically admissible as an explanatory construct? Provided that it can be supplied with auxiliary assumptions to form a cohort that has definite observable consequences, yes. Even if *psuche* is in some sense unintelligible to the twenty-first-century mind, from the iron rule's point of view it is a potential scientific explainer all the same, just as long as its causal implications are, in a suitable theoretical context, explicit and sufficiently detailed.

Though permissive, the iron rule's conception of explanation is nevertheless not a matter of "anything goes." The Atlantean view of explanatory power does not qualify. Rhymes won't get you scientific glory; your theory has to tell a causal story.

By the time Newton died in 1727, it was already clear that he was, even among the great luminaries of the era, something special. At his funeral the pallbearers were dukes, earls, and the Lord Chancellor of England; the young French Enlightenment thinker Voltaire, who was present, reported that he was "buried like a king." Tributes to his brilliance lit up the eighteenth century with verbal fireworks. The Scottish philosopher David Hume described him as "the greatest and rarest genius that ever wrote for the ornament and instruction of the species"; Voltaire held him to be "the greatest genius that ever existed." The Marquis de l'Hôpital, a French mathematician, supposedly went further still: Newton was "a celestial intelligence entirely disengaged from matter."

Such intimations of near-divinity had begun with a poem that Edmond Halley wrote for the first edition of the *Principia*, ending with the line, "No closer to the gods can any mortal rise." The English artist George Bickham put the words into pictorial form in a 1732 engraving showing Newton as the sun surrounded by angels, muses, and putti. Around the same time, the Venetian painter Giovanni Battista Pittoni

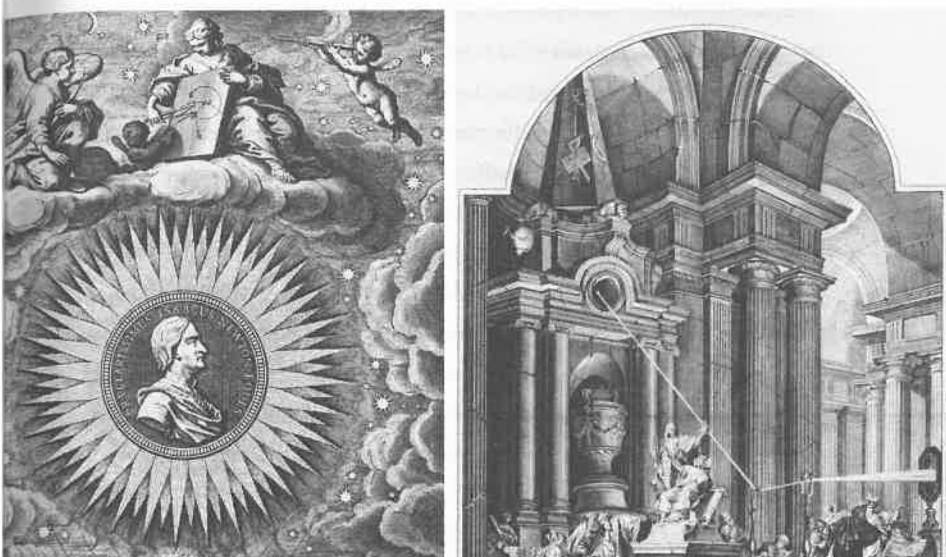

Figure 6.4. Newton at home with the gods. On the left, *Sir Isaac Newton*, by George Bickham the elder (1732); on the right, Giovanni Battista Pittoni's *Homage to Newton* (1727–1729).

conceived his *Homage to Newton*, in which an angel and the goddess of wisdom Minerva lead a procession of muses to Newton's shrine—a rococo confection featuring a colossal urn containing Newton's ashes and a reconstruction and commemoration of the optical experiment in which he used a prism to split light into its component colors (Figure 6.4).

What mattered above all for the spread of the iron rule's shallow conception of explanation, however, was not its advocate Newton's position as foremost among scientists, not even his metaphorical ascent into heaven, but, as the historian of science Mordechai Feingold puts it, his metamorphosis "into science personified." As a consequence, Feingold continues, "Newtonian science . . . became the model to emulate, the manifestation of 'superior knowledge' that summoned all other learning to reorient itself along similar lines." To be scientific simply was to be Newtonian.

So the investigation of nature changed forever. No longer were deep philosophical insights of the sort that founded Descartes's system considered to be the keys to the kingdom of knowledge. Put foundational matters aside, Newton's example seemed to urge, and devote your days instead to the construction of causal principles that, in their forecasts, follow precisely the contours of the observable world. The thinkers around and after Newton got the message, one by one.

Three centuries later, the iron rule's shallow conception of explanation continues to push science forward. There is perhaps no better illustration of its curious effectuality than the story of how the rule confronted the most perplexing scientific theory in history—quantum mechanics—and swallowed it whole.

We come to life suspended in fluid; we are then born into a world of solid things. These two ways that a substance can be, solid and fluid, have directed the scientific imagination from the earliest days. When Thales, around 580 BCE, hypothesized that the world was made of water, he opted for fundamental fluidity, as did his pupil Anaximander, who suggested air, which is also a kind of flowing stuff. The ancient Greek atomists, by contrast, thought that solidity lay at the bottom of it all, that every kind of matter, even water and air, is made of tiny, rigid, indivisible particles. The two views can be combined, as in the four-element theory of Empedocles the Sicilian: everything, he said, was composed of some mix of air, fire, earth, and water. Descartes, much later, found a different way of mixing solid and fluid: matter is rigid by its very nature but comes in many sizes, the larger chunks making up solid matter like the planets, while, at the other end of the scale, particles that can be infinitesimally tiny constitute a kind of dust that, like a fluid, flows into every available nook and crevice.

In the nineteenth century, more exotic forms of fluid and solid behavior were called upon to do explanatory service. As dramatized in my tale of Montague and Capulet, heat began the century as a fluid but ended it as the motion of solids—the disordered vibration of countless tiny particles. The same intellectual current revealed that gases, the most rarefied of tangible fluids, were themselves composed of microscopic solid particles traveling at high speed. Light, by contrast, was by 1900 thought to be a wave traveling through a fluid more subtle and ineffable than any gas, the electromagnetic ether. Inspired by this idea, a small but influential group of scientists proposed to analyze all solid matter as "ethereal." In the words of the German physicist Gustav Mie, writing in 1911:

> Elementary material particles . . . are simply singular places in the ether at which lines of electric stress of the ether converge; briefly, they are "knots" of the electric field in the ether.

After 2,500 years, it might have seemed that physics was sailing back to Miletus, where the ancient Greek philosopher Thales first made the bold surmise that all was liquid and liquidity was all.

The ship never arrived. By the 1930s, the great conceptual and metaphysical cornerstone that was the solid/fluid duality had been obliterated. Matter—now conceived as a mix of particles such as electrons and protons—was still in the picture, but thanks to quantum mechanics, it no longer behaved at all like the ordinary physical stuff of everyday experience, fluid or solid. It spent its days rather in a mysterious mode of being called "superposition."

Superposition was nigh unintelligible. It didn't matter—not to working scientists. What they cared about was shallow explanation, and quantum mechanics succeeded in that endeavor like nothing that had come before. The case of quantum mechanics, then, is the perfect exemplification of the supremacy, in modern science, of Newton's precept that, for the purposes of "experimental philosophy," derivation of the observable is the only kind of explanation that counts.

The first hints of superposition came in the initial decade of the twentieth century in a series of theories designed to make sense of puzzling phenomena involving radiation and light. Perhaps the most iconic of these was Albert Einstein's explanation of the photoelectric effect, in which light striking certain substances causes the emission of electrons. Light, Einstein posited, which usually behaves like a wave, in this case behaves like a stream of particles—"photons." Descartes had argued that light is made of particles. Nineteenth-century scientists had shown that it must be a wave. Now it turned out that it was not one or the other, but somehow both at once.

When the first full versions of quantum mechanics were formulated in the 1920s, it became clear that it was not only light that lived a double life: streams of electrons, universally assumed to be ordinary particles, were quite capable of behaving like waves. In fact, all matter appeared to be some sort of wave-particle chimera, as though ancient mythology had all along been truer to the hidden structure of the universe than ancient physics—as though the centaur, half horse and half man, or the sphinx, half woman and half lion, were better models of reality than Aristotelian substance or the atom.

It was a philosophical emergency. The nature of reality was at stake. What was quantum physics trying to tell us about the world? The question was intensely discussed in letters and in lab corridors, most of all those of Niels Bohr's Institute for Theoretical Physics in Copenhagen. Bohr was an enthusiast for quantum mechanics, which he had helped to pioneer; Einstein was a skeptic in spite of his own early contributions. The two scientists famously debated the adequacy of the theory at the fifth Solvay Conference on Physics and Chemistry in Brussels in 1927 and then again at the sixth in 1930. The physicist Paul Ehrenfest gave a sense of the spectacle:

> Bohr from out of philosophical smoke clouds constantly searching for the tools to crush one example after the other. Einstein like a jack in the box; jumping out fresh every morning. Oh, that was priceless.

Einstein had earlier written to another founder of quantum theory, Max Born: "Quantum mechanics is certainly imposing. But an inner

Figure 6.5. Participants in the fifth Solvay Conference, 1927, in Brussels. Bohr is in the second row on the very right; Einstein is at the center of the first row.

voice tells me that it is not yet the real thing." Ninety years later, physicists still have no alternative. It is beginning to look like quantum mechanics is indeed the real thing. The debates about its meaning continue. Yet the science rolls imperturbably on, as quantum theory has been extended since 1930 to apply to electromagnetic force, the interior of the atomic nucleus, and (though an undertaking still in its infancy) gravitation itself.

That is a remarkable scientific achievement. It is also a remarkable social achievement. In the world of ancient Greece or the pre-Newtonian seventeenth century, competing philosophies meant competing sciences. But quantum mechanics did not split into philosophical schools. Rather, even as Bohr, Einstein, and many other key figures at the Solvay Conference in 1927 philosophized furiously, the theory remained a unified set of ideas that "developed rapidly, disseminated very quickly, and met almost no resistance." The new textbooks written to explain quantum mechanics to students barely mentioned the disputes at all.

The interpretation of quantum mechanics was, in short, considered to be so much philosophical superstructure, perched on top of but hardly integral to the science below. The same is true today. Yet it is not as if the strangeness of the theory is invisible to scientists. Murray Gell-Mann, the discoverer of quarks, called quantum mechanics a "mysterious, confusing discipline." According to Roger Penrose, one of the late twentieth century's foremost mathematical physicists, quantum mechanics "makes absolutely no sense." "I think I can safely say that nobody understands quantum mechanics," remarked Richard Feynman. How can a theory be widely regarded both as incomprehensible and also as the best explanation we have of the physical world we live in?

The answer lies in the shallow conception of causal explanation, in which it is derivation rather than comprehension that is paramount. To inspect that answer more closely, let me take you wading in the shallows of quantum explanation. Though entirely nontechnical, the overview that follows will take a few pages. But I promise you that it's worth your time. What we'll see is a contrast between the obscure *nature* of superposition and its clear *consequences*. The natural philosophers, such as Aristotle and Descartes, cared about the consequences of their theories' causal principles, of course, but they also cared a great deal about their nature—about what the principles were saying about the metaphysical foundation of the world's causal structure. Modern science, by contrast, is oblivious to the principles' nature. Let it be as opaque as action at a distance was to Newton's adversary Leibniz, as Aristotle's *psuche* is to us: it is immaterial. All that concerns the iron rule is what kinds of things the causal principles are capable of bringing about. I will show you how quantum theory derives accurate predictions from a notion, superposition, that is quite beyond our human understanding.

Matter, says quantum mechanics, occupies the state called superposition when it is not being observed. An electron in superposition occupies no particular point in space. It is typically, rather, in a kind of "mix" of being in many places at once. The mix is not perfectly balanced: some places are far more heavily represented than others. So a particular electron's superposition might be almost all made up from positions near a certain atomic nucleus and just a little bit from positions elsewhere. That is the closest that quantum mechanics comes to saying that the electron is orbiting the nucleus.

As to the nature of this "mix"—it is a mystery. We give it a name: superposition. But we can't give it a philosophical explanation. What we can do is to represent any superposition with a mathematical formula, called a "wave function." An electron's wave function represents its physical state with the same exactitude that, in Newton's physics, its state would be represented by numbers specifying its precise position and velocity. You may have heard of quantum mechanics' "uncertainty principle," but forget about uncertainty here: the wave function is a complete description that captures every matter of fact about an electron's physical state without remainder.

So far, we have a mathematical representation of the state of any particular piece of matter, but we haven't said how that state changes in time. This is the job of Schrödinger's equation, which is the quantum equivalent of Newton's famous second law of motion $F = ma$, in that it spells out how forces of any sort—gravitational, electrical, and so on—will affect a quantum particle. According to Schrödinger's equation, the wave function will behave in what physicists immediately recognize as a "wavelike" way. That is why, according to quantum mechanics, even particles such as electrons conduct themselves as though they are waves.

In the early days of quantum mechanics, Erwin Schrödinger, the Austrian physicist who formulated the equation in 1926, and Louis de Broglie, a French physicist—both eventual Nobel Prize winners—wondered whether the waves described by quantum mechanics might be literal waves traveling through a sea of "quantum ether" that pervades our universe. They attempted to understand quantum mechanics, then, using the old model of the fluid. This turned out to be impossible for a startling reason: it is often necessary to assign a wave function not to a single particle, like an electron, but to a whole system of particles. Such a wave function is defined in a space that has three dimensions for every particle in the system: for a 2-particle system, then, it has 6 dimensions; for a 10-particle system, 30 dimensions. Were the wave to be a real entity made of vibrations in the ether, it would therefore have to be flowing around a space of 6, or 30, or even more dimensions. But our universe rather stingily supplies only three dimensions for things to happen in. In quantum mechanics, as Schrödinger and de Broglie soon realized, the notion of substance as fluid fails completely.

There is a further component to quantum mechanics. It is called Born's rule, and it says what happens when a particle's position or other state is measured. Suppose that an electron is in a superposition, a mix of being "everywhere and nowhere." You use the appropriate instruments to take a look at it; what do you see? Eerily, you see it occupying a definite position. Born's rule says that the position is a matter of chance: the probability that a particle appears in a certain place is proportional to the degree to which that place is represented in the mix.

It is as though the superposition is an extremely complex cocktail, a combination of various amounts of infinitely many ingredients, each representing the electron's being in a particular place. Taste the cocktail, and instead of an infinitely complex flavor you will—according to Born's rule—taste only a single ingredient. The chance of tasting that ingredient is proportional to the amount of the ingredient contained in the mixture that makes up the superposition. If an electron's state is mostly a blend of positions near a certain atomic nucleus, for example, then when you observe it, it will most likely pop up near the nucleus.

One more thing: an observed particle's apparently definite position is not merely a fleeting glimpse of something more complex. Once you see the particle in a certain position, it goes on to act as though it really is in that position (until something happens to change its state). In mixological terms, once you have sampled your cocktail, every subsequent sip will taste the same, as though the entire cocktail has transformed into a simple solution of this single ingredient. It is this strange disposition for matter, when observed, to snap into a determinate place that accounts for its "particle-like" behavior.

To sum up, quantum mechanical matter—the matter from which we're all made—spends almost all its time in a superposition. As long as it's not observed, the superposition, and so the matter, behaves like an old-fashioned wave, an exemplar of liquidity (albeit in indefinitely many dimensions). If it is observed, the matter jumps randomly out of its superposition and into a definite position like an old-fashioned particle, the epitome of solidity.

Nobody can explain what kind of substance this quantum mechanical matter is, such that it behaves in so uncanny a way. It seems that it can be neither solid nor fluid—yet these exhaust the possibilities that our human minds can grasp. Quantum mechanics does not, then, provide the kind of deep understanding of the way the world works that was sought by philosophers from Aristotle to Descartes. What it does supply is a precise mathematical apparatus for deriving effects from their causes. Take the initial state of a physical system, represented by a wave function; apply Schrödinger's equation and if appropriate Born's rule, and the theory tells you how the system will behave (with, if Born's rule is invoked, a probabilistic twist). In this way, quantum theory explains why electrons sometimes behave as waves, why photons (the stuff of light) sometimes behave as particles, and why atoms have the structure that they do and interact in the way they do.

Thus, quantum mechanics may not offer deep understanding, but it can still account for observable phenomena by way of shallow causal explanation, the kind of explanation, favored by Newton, that the iron rule cares about and uses to supervise modern science's procedural consensus. The theory was accepted so readily in the 1920s and 1930s because, for all the philosophical arguments surrounding its interpretation, it provided a clear and well-defined system for providing shallow explanations that had no serious rivals. Had Newton rather than Bohr debated Einstein at the Solvay conferences, he would perhaps have proclaimed:

> I have not as yet been able to deduce from phenomena the nature of quantum superposition, and I do not feign hypotheses. It is enough that superposition really exists and acts according to the laws that we have set forth and is sufficient to explain all the motions of the microscopic bodies of which matter is made.

Ehrenfest would have enjoyed that.

Explanatory standards change with the prevailing philosophical outlook of the age; this thesis of explanatory relativism is inherent in Kuhn's vision of scientific inquiry as a parade of paradigms, each with a different way of making sense of the world. The thesis holds a grain of temporary truth. For millennia, explanatory standards did change: as the metaphysical wheel turned, so the explanatory cogs turned with it, from Aristotelian teleology to Cartesian collision.

Late in the seventeenth century, the philosophically unified machinery of inquiry in which the wheels and cogs rotate as one was spiked. Newton was the saboteur, and thus he was the chief architect of modern science's first great innovation. Rather than deep philosophical understanding, Newton pursued shallow explanatory power, that is, the ability to derive correct descriptions of phenomena from a theory's causal principles, regardless of their ultimate nature and indeed regardless of their very intelligibility. In so doing, he was able to build a gravitational theory of immense capability, setting an example that his successors were eager to follow.

Predictive power thereby came to override metaphysical insight. Or as the historian of science John Heilbron, writing of the study of electricity after Newton, put it:

> When confronted with a choice between a qualitative model deemed intelligible and an exact description lacking clear physical foundations, the leading physicists of the Enlightenment preferred exactness.

So it continued to be, as the development and acceptance of quantum mechanics, as unerring as it is incomprehensible, goes to show. The criterion for explanatory success inherent in Newton's practice became fixed for all time, founding the procedural consensus that lies at the heart of modern science. From the raw material of shallow explanation, the iron rule was forged.