

BOOKS OCTOBER 5, 2020 ISSUE

HOW DOES SCIENCE REALLY WORK?

Science is objective. Scientists are not. Can an “iron rule” explain how they’ve changed the world anyway?

By **Joshua Rothman**

September 28, 2020



Michael Strevens argues that “shallow explanation” can be singularly powerful. Illustration by Alexander

Glandien



0:00 / 29:06

Audio: Listen to this article.

When I was a kid, I'd sometimes spend the day with my dad in his lab, at the National Institutes of Health. For a few hours I'd read, while eating vending-machine crackers and drinking Diet Coke. I'd spend the rest of the time at a lab bench, pipetting—using a long glass eyedropper to draw water out of one set of test tubes and drip it, carefully, into another.

I was seven, eight, maybe nine years old. Still, the lab was an interesting place for me. I understood, loosely, that my dad was investigating addiction in the brain. He believed that it depended on the way certain chemicals bind to certain receptors. To study this, the scientists in his lab performed experiments on rats, then killed them and analyzed their brains. On one of my visits, a lab tech named Victor reached into a centrifuge and removed a large container filled with foamy pink liquid. “Brain juice!” he said, pretending to drink it.

Often, though, we were there on weekends, and were the only ones in the lab. The corridors were dim and quiet, the rooms mostly dark and deserted; the metal and linoleum surfaces were beige, gray, white, and green, relieved, occasionally, by a knob or button made of vivid red or blue plastic. Hulking machines stood on the counters—ugly but, according to my dad, incredibly expensive. Chemical showers and eyewash stations loomed; sometimes, in a distant room, a dot-matrix printer burred. In the sci-fi novels I devoured, labs were gleaming and futuristic. But my dad's seemed worn-in, workaday, more “Alien” than “2001.” I knew that the experiments done there took years and could come to nothing. As I pipetted, I watched my dad in his office, poring over statistical printouts—a miner in the mountains of knowledge.

Later, in college and afterward, I got to see the glamorous side of science. Some researchers had offices with sweeping views, and schedules coordinated by multiple assistants. They wore tailored clothes, spoke to large audiences, and debated ideas in fancy restaurants. Their rivalries, as they described them, evoked titanic struggles from the history of science—Darwin versus Owen, Galileo

versus the Pope—in which rationalist grit overpowered bias and folly. Science, in this world, was a form of exploratory combat, in which flexible minds stretched to encompass the truth, pushing against the limits of what was known and thought. It was an enterprise that demanded total human engagement. Even aesthetics mattered. “You live and breathe paradox and contradiction, but you can no more see the beauty of them than the fish can see the beauty of the water,” Niels Bohr tells Werner Heisenberg, in Michael Frayn’s quantum-physics play, “Copenhagen.”

Reading, seeing, learning all of this, I wanted to be a scientist. So why did I find the actual work of science so boring? In college science courses, I had occasional bursts of mind-expanding insight. For the most part, though, I was tortured by drudgery. In my senior year, I bonded with my biology professor during field work and in the lab, but found the writing of lab reports so dreary that, after consulting the grading rubric on the syllabus, I decided not to do them. I performed well enough on the exams to get a D—the minimum grade that would allow me to graduate.

Recorded history is five thousand years old. Modern science, which has been with us for just four centuries, has remade its trajectory. We are no smarter individually than our medieval ancestors, but we benefit, as a civilization, from antibiotics and electronics, vitamins and vaccines, synthetic materials and weather forecasts; we comprehend our place in the universe with an exactness that was once unimaginable. I’d found that science was two-faced: simultaneously thrilling and tedious, all-encompassing and narrow. And yet this was clearly an asset, not a flaw. Something about that combination had changed the world completely.

In “The Knowledge Machine: How Irrationality Created Modern Science” (Liveright), Michael Strevens, a philosopher at New York University, aims to identify that special something. Strevens is a philosopher of science—a scholar charged with analyzing how scientific knowledge is generated. Philosophers of science tend to irritate practicing scientists, to whom science already makes complete sense. It doesn’t make sense to Strevens. “Science is an alien thought form,” he writes; that’s why so many civilizations rose and fell before it was invented. In his view, we downplay its weirdness, perhaps because its success is so fundamental to our continued existence. He promises to serve as “the P. T. Barnum of the laboratory, unveiling the monstrosity that lies at the heart of modern science.”

In school, one learns about “the scientific method”—usually a straightforward set of steps, along the lines of “ask a question, propose a hypothesis, perform an experiment, analyze the results.” That

method works in the classroom, where students are basically told what questions to pursue. But real scientists must come up with their own questions, finding new routes through a much vaster landscape.

Since science began, there has been disagreement about how those routes are charted. Two twentieth-century philosophers of science, Karl Popper and Thomas Kuhn, are widely held to have offered the best accounts of this process. Popper maintained that scientists proceed by “falsifying” scientific claims—by trying to prove theories wrong. Kuhn, on the other hand, believed that scientists work to prove theories right, exploring and extending them until further progress becomes impossible. These two accounts rest on divergent visions of the scientific temperament. For Popper, Strevens writes, “scientific inquiry is essentially a process of disproof, and scientists are the disprovers, the debunkers, the destroyers.” Kuhn’s scientists, by contrast, are faddish true believers who promulgate received wisdom until they are forced to attempt a “paradigm shift”—a painful rethinking of their basic assumptions.

Working scientists tend to prefer Popper to Kuhn. But Strevens thinks that both theorists failed to capture what makes science historically distinctive and singularly effective. To illustrate, he tells the story of Roger Guillemin and Andrew Schally, two “rival endocrinologists” who shared a Nobel Prize in 1977 for discovering the molecular structure of TRH—a hormone, produced in the hypothalamus, that helps regulate the release of other hormones and so shapes many aspects of our lives. Mapping the hormone’s structure, Strevens explains, was an “epic slog” that lasted more than a decade, during which “literally tons of brain tissue, obtained from sheep or pigs, had to be mashed up and processed.” Guillemin and Schally, who were racing each other to analyze TRH—they crossed the finish line simultaneously—weren’t weirdos who loved animal brains. They gritted their teeth through the work. “Nobody before had to process millions of hypothalami,” Schally said. “The key factor is not the money, it’s the will . . . the brutal force of putting in sixty hours a week for a year to get one million fragments.”

Looking back on the project, Schally attributed their success to their outsider status. “Guillemin and I, we are immigrants, obscure little doctors, we fought our way to the top,” he said. But Strevens points out that “many important scientific studies have required of their practitioners a degree of single-mindedness that is quite inhuman.” It’s not just brain juice that demands such commitment.

Scientists have dedicated entire careers to the painstaking refinement of delicate instruments, to the digging up of bone fragments, to the gathering of statistics about variations in the beaks of finches. Uncertain of success, they toil in an obscurity that will deepen into futility if their work doesn't pan out.

"Science is boring," Strevens writes. "Readers of popular science see the 1 percent: the intriguing phenomena, the provocative theories, the dramatic experimental refutations or verifications." But, he says,

behind these achievements . . . are long hours, days, months of tedious laboratory labor. The single greatest obstacle to successful science is the difficulty of persuading brilliant minds to give up the intellectual pleasures of continual speculation and debate, theorizing and arguing, and to turn instead to a life consisting almost entirely of the production of experimental data.

The allocation of vast human resources to the measurement of possibly inconsequential minutiae is what makes science truly unprecedented in history. Why do scientists agree to this scheme? Why do some of the world's most intelligent people sign on for a lifetime of pipetting?

Strevens thinks that they do it because they have no choice. They are constrained by a central regulation that governs science, which he calls the "iron rule of explanation." The rule is simple: it tells scientists that, "if they are to participate in the scientific enterprise, they must uncover or generate new evidence to argue with"; from there, they must "conduct all disputes with reference to empirical evidence alone." Compared with the theories proposed by Popper and Kuhn, Strevens's rule can feel obvious and underpowered. That's because it isn't intellectual but procedural. "The iron rule is focused not on what scientists think," he writes, "but on what arguments they can make in their official communications." Still, he maintains, it is "the key to science's success," because it "channels hope, anger, envy, ambition, resentment—all the fires fuming in the human heart—to one end: the production of empirical evidence."

Strevens arrives at the idea of the iron rule in a Popperian way: by disproving the other theories about how scientific knowledge is created. The problem isn't that Popper and Kuhn are completely wrong. It's that scientists, as a group, don't pursue any single intellectual strategy consistently. Exploring a number of case studies—including the controversies over continental drift, spontaneous generation,

and the theory of relativity—Strevens shows scientists exerting themselves intellectually in a variety of ways, as smart, ambitious people usually do. Sometimes they seek to falsify theories, sometimes to prove them; sometimes they're informed by preëxisting or contextual views, and at other times they try to rule narrowly, based on the evidence at hand.

Like everybody else, scientists view questions through the lenses of taste, personality, affiliation, and experience. In 1912, a young meteorologist and champion balloonist named Alfred Wegener proposed that the continents had once fit together but then drifted apart. His theory, which drew on a global survey of coastlines and continental shelves, made sense of the fact that the same sorts of rocks and fossilized animals often appeared on distant shores. Opponents of Wegener's theory, led by the eminent paleontologist George Gaylord Simpson, pointed out that he had no explanation for how the continents had moved. A rational non-scientist might have stayed neutral until more evidence had come in. But geologists had a professional obligation to take sides. Europeans, Strevens reports, tended to back Wegener, who was German, while scholars in the United States often preferred Simpson, who was American. Outsiders to the field were often more receptive to the concept of continental drift than established scientists, who considered its incompleteness a fatal flaw.

Strevens's point isn't that these scientists were doing anything wrong. If they had biases and perspectives, he writes, "that's how human thinking works." His point is that, despite their heated partiality, the papers they published consisted solely of data about rocks. Ultimately, in fact, it was good that the geologists had a "splendid variety" of somewhat arbitrary opinions: progress in science requires partisans, because only they have "the motivation to perform years or even decades of necessary experimental work." It's just that these partisans must channel their energies into empirical observation. The iron rule, Strevens writes, "has a valuable by-product, and that by-product is data."

Science is often described as "self-correcting": it's said that bad data and wrong conclusions are rooted out by other scientists, who present contrary findings. But Strevens thinks that the iron rule is often more important than overt correction. He tells the story of Arthur Eddington, an English astronomer who, in 1919, sailed to the island of Príncipe, off the west coast of Africa, to observe and photograph the position of a group of stars during a total eclipse of the sun. Eddington's observations were expected to either confirm or falsify Einstein's theory of general relativity, which predicted that the sun's gravity would bend the path of light, subtly shifting the stellar pattern. For reasons having

to do with weather and equipment, the evidence collected by Eddington—and by his colleague Frank Dyson, who had taken similar photographs in Sobral, Brazil—was inconclusive; some of their images were blurry, and so failed to resolve the matter definitively. Eddington pressed ahead anyway: the expedition report he published with Dyson contained detailed calculations and numerical tables that, they argued, showed that Einstein was right.

At the time, many physicists and astronomers were skeptical of the findings. Everyone knew that Eddington “wanted very much for Einstein’s theory to be true,” Strevens writes, “both because of its profound mathematical beauty” and because of Eddington’s “ardent internationalist desire to dissolve the rancor that had some Britons calling for a postwar boycott of German science.” (As a Quaker and an avowed pacifist, Eddington believed that scientific progress could be “a bond transcending human differences.”) All the same, Eddington was never really refuted. Other astronomers, driven by the iron rule, were already planning their own studies, and “the great preponderance of the resulting measurements fit Einsteinian physics better than Newtonian physics.” It’s partly by generating data on such a vast scale, Strevens argues, that the iron rule can power science’s knowledge machine: “Opinions converge not because bad data is corrected but because it is swamped.”

Why did the iron rule emerge when it did? Strevens takes us back to the Thirty Years’ War, which concluded with the Peace of Westphalia, in 1648. The war weakened religious loyalties and strengthened national ones. Afterward, he writes, what mattered most “was that you were English or French”; whether you were Anglican or Catholic became “your private concern.” Two regimes arose: in the spiritual realm, the will of God held sway, while in the civic one the decrees of the state were paramount. As Isaac Newton wrote, “The laws of God & the laws of man are to be kept distinct.” These new, “nonoverlapping spheres of obligation,” Strevens argues, were what made it possible to imagine the iron rule. The rule simply proposed the creation of a third sphere: in addition to God and state, there would now be science.

In the single-sphered, pre-scientific world, thinkers tended to inquire into everything at once. Often, they arrived at conclusions about nature that were fascinating, visionary, and wrong. Looking back, we usually fault such thinkers for being insufficiently methodical and empirical. But Strevens tells a more charitable story: it was only natural for intelligent people who were free of the rule’s strictures

to attempt a kind of holistic, systematic inquiry that was, in many ways, more demanding. It never occurred to them to ask if they might illuminate more collectively by thinking about less individually.

It's in this context, Strevens suggests, that we should understand the story of René Descartes, the philosopher and mathematician who, among other things, invented the system of plotting points and lines on a grid. In his first book, "The World," completed in 1633, Descartes, who was then in his late thirties, offered a sprawling account of the universe, explaining how vision works, how muscles move, how plants grow, how gravity functions, and how God set everything spinning in the first place. Today, the ambition of treatises like "The World" strikes us as absurd. But Strevens imagines how, to someone in Descartes's time, the iron rule would have seemed "unreasonably closed-minded." Since ancient Greece, it had been obvious that the best thinking was cross-disciplinary, capable of knitting together "poetry, music, drama, philosophy, democracy, mathematics," and other elevating human disciplines. We're still accustomed to the idea that a truly flourishing intellect is a well-rounded one. And, by this standard, Strevens says, the iron rule looks like "an irrational way to inquire into the underlying structure of things"; it seems to demand the upsetting "suppression of human nature." (Perhaps it's as compensation that, today, so many scientists seem to pursue their hobbies—woodworking, sailing, ballroom dancing—with such avidity.) Descartes, in short, would have had good reasons for resisting a law that narrowed the grounds of disputation, or that encouraged what Strevens describes as "doing rather than thinking."

In fact, the iron rule offered scientists a more supple vision of progress. Before its arrival, intellectual life was conducted in grand gestures. Descartes's book was meant to be a complete overhaul of what had preceded it; its fate, had science not arisen, would have been replacement by some equally expansive system. The iron rule broke that pattern. Strevens sees its earliest expression in Francis Bacon's "The New Organon," a foundational text of the Scientific Revolution, published in 1620. Bacon argued that thinkers must set aside their "idols," relying, instead, only on evidence they could verify. This dictum gave scientists a new way of responding to one another's work: gathering data. But it also changed what counted as progress. In the past, a theory about the world was deemed valid when it was complete—when God, light, muscles, plants, and the planets cohered. The iron rule allowed scientists to step away from the quest for completeness.

The consequences of this shift would become apparent only with time. In 1713, Isaac Newton appended a postscript to the second edition of his “Principia,” the treatise in which he first laid out the three laws of motion and the theory of universal gravitation. “I have not as yet been able to deduce from phenomena the reason for these properties of gravity, and I do not feign hypotheses,” he wrote. “It is enough that gravity really exists and acts according to the laws that we have set forth.” What mattered, to Newton and his contemporaries, was his theory’s empirical, predictive power—that it was “sufficient to explain all the motions of the heavenly bodies and of our sea.”

Descartes would have found this attitude ridiculous. He had been playing a deep game—trying to explain, at a fundamental level, how the universe fit together. Newton, by those lights, had failed to explain anything: he himself admitted that he had no sense of how gravity did its work or fit into the whole; he’d merely produced equations that predicted observations. If he’d made progress, it was only by changing the rules of the game, redefining wide-ranging inquiry as a private pastime, rather than official business. And yet, by authorizing what Strevens calls “shallow explanation,” the iron rule offered an empirical bridge across a conceptual chasm. Work could continue, and understanding could be acquired on the other side. In this way, shallowness was actually more powerful than depth.

We seem to be crossing a similar bridge today. Quantum theory—which tells us that subatomic particles can be “entangled” across vast distances, and in multiple places at the same time—makes intuitive sense to pretty much nobody. Niels Bohr and Werner Heisenberg, who argued in Copenhagen (and in “Copenhagen”), agreed on one interpretation of the theory, according to which the universe is essentially probabilistic; Albert Einstein took the opposite view. Eight decades later, it’s still unclear what the theory means. The confusion most of us feel about it is echoed, in a higher register, among physicists, who argue about whether there are many worlds or one.

Without the iron rule, Strevens writes, physicists confronted with such a theory would have found themselves at an impasse. They would have argued endlessly about quantum metaphysics. Following the iron rule, they can make progress empirically even though they are uncertain conceptually. Individual researchers still passionately disagree about what quantum theory means. But that hasn’t stopped them from using it for practical purposes—computer chips, MRI machines, G.P.S. networks, and other technologies rely on quantum physics. It hasn’t prevented universities and

governments from spending billions of dollars on huge machines that further explore the quantum world. Even as we wait to understand the theory, we can refine it, one decimal place at a time.

Compared with other stories about the invention and success of science, “The Knowledge Machine” is unusually parsimonious. Other theorists have explained science by charting a sweeping revolution in the human mind; inevitably, they’ve become mired in a long-running debate about how objective scientists really are. One group of theorists, the rationalists, has argued that science is a new way of thinking, and that the scientist is a new kind of thinker—dispassionate to an uncommon degree. As evidence against this view, another group, the subjectivists, points out that scientists are as hopelessly biased as the rest of us. To this group, the aloofness of science is a smoke screen behind which the inevitable emotions and ideologies hide.

Strevens offers a more modest story. The iron rule—“a kind of speech code”—simply created a new way of communicating, and it’s this new way of communicating that created science. The subjectivists are right, he admits, inasmuch as scientists are regular people with a “need to win” and a “determination to come out on top.” But they are wrong to think that subjectivity compromises the scientific enterprise. On the contrary, once subjectivity is channelled by the iron rule, it becomes a vital component of the knowledge machine. It’s this redirected subjectivity—to come out on top, you must follow the iron rule!—that solves science’s “problem of motivation,” giving scientists no choice but “to pursue a single experiment relentlessly, to the last measurable digit, when that digit might be quite meaningless.”

On one level, it’s ironic to find a philosopher—a professional talker—arguing that science was born when philosophical talk was exiled to the pub. On another, it makes sense that a philosopher would be attuned to the power of how we talk and argue. If it really was a speech code that instigated “the extraordinary attention to process and detail that makes science the supreme discriminator and destroyer of false ideas,” then the peculiar rigidity of scientific writing—Strevens describes it as “sterilized”—isn’t a symptom of the scientific mind-set but its cause. Etiquette is what has created the modern world.

Does Strevens’s story have implications outside of science? Today, we think a lot about speech—about its power to frame, normalize, empower, and harm. In our political discourse, we value unfiltered authenticity; from our journalism, we demand moral clarity. Often, we bring our whole

selves into what we say. And yet we may be missing something important about how speech drives behavior. At least in science, Strevens tells us, “the appearance of objectivity” has turned out to be “as important as the real thing.” Perhaps speech codes can be building materials for knowledge machines. In that case, our conversations can still be fiery and wide-ranging. But we should write those lab reports, too. ♦

Published in the print edition of the October 5, 2020, issue, with the headline “The Rules of the Game.”



*Joshua Rothman, the ideas editor of [newyorker.com](https://www.newyorker.com), has been at *The New Yorker* since 2012.*

More: [Science](#) [Scientists](#) [Theories](#) [Histories](#) [National Institutes of Health](#)
