

# WHY IS IT SO HARD TO BE RATIONAL?

*The real challenge isn't being right but knowing how wrong you might be.*

**By Joshua Rothman**

August 16, 2021



Part of being “metarational” is knowing when to let someone else do the thinking. Illustration by Francesco Ciccolella

I met the most rational person I know during my freshman year of college. Greg (not his real name) had a tech-support job in the same computer lab where I worked, and we became friends. I planned to be a creative-writing major; Greg told me that he was deciding between physics and economics. He'd choose physics if he was smart enough, and economics if he wasn't—he thought he'd know within a few months, based on his grades. He chose economics.

We roomed together, and often had differences of opinion. For some reason, I took a class on health policy, and I was appalled by the idea that hospital administrators should take costs into account when providing care. (Shouldn't doctors alone decide what's best for their patients?) I got worked up, and developed many arguments to support my view; I felt that I was right both practically and morally. Greg shook his head. He pointed out that my dad was a doctor, and explained that I was engaging in "motivated reasoning." My gut was telling me what to think, and my brain was figuring out how to think it. This felt like thinking, but wasn't.

The next year, a bunch of us bought stereos. The choices were complicated: channels, tweeters, woofers, preamps. Greg performed a thorough analysis before assembling a capable stereo. I bought one that, in my opinion, looked cool and possessed some ineffable, tonal *je ne sais quoi*. Greg's approach struck me as unimaginative, utilitarian. Later, when he upgraded to a new sound system, I bought his old equipment and found that it was much better than what I'd chosen.

In my senior year, I began considering graduate school. One of the grad students I knew warned me off—the job prospects for English professors were dismal. Still, I made the questionable decision to embark on a Ph.D. Greg went into finance. We stayed friends, often discussing the state of the world and the meta subject of how to best ascertain it. I felt overwhelmed by how much there was to know—there were too many magazines, too many books—

and so, with Greg as my Virgil, I travelled deeper into the realm of rationality. There was, it turned out, a growing rationality movement, with its own ethos, thought style, and body of knowledge, drawn heavily from psychology and economics. Like Greg, I read a collection of rationality blogs—Marginal Revolution, Farnam Street, Interfluidity, Crooked Timber. I haunted the Web sites of the Social Science Research Network and the National Bureau of Economic Research, where I could encounter just-published findings; I internalized academic papers on the cognitive biases that slant our thinking, and learned a simple formula for estimating the “expected value” of my riskier decisions. When I was looking to buy a house, Greg walked me through the trade-offs of renting and owning (just rent); when I was contemplating switching careers, he stress-tested my scenarios (I switched). As an emotional and impulsive person by nature, I found myself working hard at rationality. Even Greg admitted that it was difficult work: he had to constantly inspect his thought processes for faults, like a science-fictional computer that had just become sentient...

Often, I asked myself, How would Greg think? I adopted his habit of tracking what I knew and how well I knew it, so that I could separate my well-founded opinions from my provisional views. Bad investors, Greg told me, often had flat, loosely drawn maps of their own knowledge, but good ones were careful cartographers, distinguishing between settled, surveyed, and unexplored territories. Through all this, our lives unfolded. Around the time I left my grad program to try out journalism, Greg swooned over his girlfriend’s rational mind, married her, and became a director at a hedge fund. His net worth is now several thousand times my own.

Meanwhile, half of Americans won’t get vaccinated; many believe in conspiracy theories or pseudoscience. It’s not that we don’t think—we are constantly reading, opining, debating—but that we seem to do it on the run, while squinting at trolls in our phones. This summer, on my phone, I read a blog post by the economist Arnold Kling, who noted that an unusually large number of books about rationality were being published this year, among them

Steven Pinker's "Rationality: What It Is, Why It Seems Scarce, Why It Matters" (Viking) and Julia Galef's "The Scout Mindset: Why Some People See Things Clearly and Others Don't" (Portfolio). It makes sense, Kling suggested, for rationality to be having a breakout moment: "The barbarians sack the city, and the carriers of the dying culture repair to their basements to write." In a polemical era, rationality can be a kind of opinion hygiene—a way of washing off misjudged views. In a fractious time, it promises to bring the court to order. When the world changes quickly, we need strategies for understanding it. We hope, reasonably, that rational people will be more careful, honest, truthful, fair-minded, curious, and right than irrational ones.

And yet rationality has sharp edges that make it hard to put at the center of one's life. It's possible to be so rational that you are cut off from warmer ways of being—like the student Bazarov, in Ivan Turgenev's "Fathers and Sons," who declares, "I look up to heaven only when I want to sneeze." (Greg, too, sometimes worries that he is rational to excess—that he is becoming a heartless boss, a cold fish, a robot.) You might be well-intentioned, rational, and mistaken, simply because so much in our thinking can go wrong.

("RATIONAL, adj.: Devoid of all delusions save those of observation, experience and reflection," Ambrose Bierce wrote, in his "Devil's Dictionary.") You might be rational and self-deceptive, because telling yourself that you are rational can itself become a source of bias. It's possible that you are trying to appear rational only because you want to impress people; or that you are more rational about some things (your job) than others (your kids); or that your rationality gives way to rancor as soon as your ideas are challenged. Perhaps you irrationally insist on answering difficult questions yourself when you'd be better off trusting the expert consensus. Possibly, like Mr. Spock, of "Star Trek," your rational calculations fail to account for the irrationality of other people. (Surveying Spock's predictions, Galef finds that the outcomes Spock has determined to be impossible actually happen about eighty per cent of the time, often because he assumes that other people will be as "logical" as he is.)

Not just individuals but societies can fall prey to false or compromised

rationality. In a 2014 book, “The Revolt of the Public and the Crisis of Authority in the New Millennium,” Martin Gurri, a C.I.A. analyst turned libertarian social thinker, argued that the unmasking of allegedly pseudo-rational institutions had become the central drama of our age: people around the world, having concluded that the bigwigs in our colleges, newsrooms, and legislatures were better at appearing rational than at being so, had embraced a nihilist populism that sees all forms of public rationality as suspect. COVID deniers and climate activists are different kinds of people, but they’re united in their frustration with the systems built by experts on our behalf—both groups picture élites shuffling PowerPoint decks in Davos while the world burns. From this perspective, the root cause of mass irrationality is the failure of rationalists. People would believe in the system if it actually made sense.

And yet modern life would be impossible without those rational systems; we must improve them, not reject them. We have no choice but to wrestle with rationality—an ideal that, the sociologist Max Weber wrote, “contains within itself a world of contradictions.” We want to live in a more rational society, but not in a falsely rationalized one. We want to be more rational as individuals, but not to overdo it. We need to know when to think and when to stop thinking, when to doubt and when to trust. Rationality is one of humanity’s superpowers. How do we keep from misusing it?

**W**riting about rationality in the early twentieth century, Weber saw himself as coming to grips with a titanic force—an ascendant outlook that was rewriting our values. He talked about rationality in many different ways. We can practice the instrumental rationality of means and ends (how do I get what I want?) and the value rationality of purposes and goals (do I have good reasons for wanting what I want?). We can pursue the rationality of affect (am I cool, calm, and collected?) or develop the rationality of habit (do I live an ordered, or “rationalized,” life?). Rationality was obviously useful, but Weber worried that it was turning each individual into a “cog in the machine,” and life into an “iron cage.” Today, rationality and the words around it are still shadowed with Weberian pessimism and cursed with double meanings. You’re

rationalizing the org chart: are you bringing order to chaos, or justifying the illogical?

The Weberian definitions of rationality are by no means canonical. In “The Rationality Quotient: Toward a Test of Rational Thinking” (M.I.T.), from 2016, the psychologists Keith E. Stanovich, Richard F. West, and Maggie E. Toplak call rationality “a torturous and tortured term,” in part because philosophers, sociologists, psychologists, and economists have all defined it differently. For Aristotle, rationality was what separated human beings from animals. For the authors of “The Rationality Quotient,” it’s a mental faculty, parallel to but distinct from intelligence, which involves a person’s ability to juggle many scenarios in her head at once, without letting any one monopolize her attention or bias her against the rest. It’s because some people are better jugglers than others that the world is full of “smart people doing dumb things”: college kids getting drunk the night before a big exam, or travellers booking flights with impossibly short layovers.

Galef, who hosts a podcast called “Rationally Speaking” and co-founded the nonprofit Center for Applied Rationality, in Berkeley, barely uses the word “rationality” in her book on the subject. Instead, she describes a “scout mindset,” which can help you “to recognize when you are wrong, to seek out your blind spots, to test your assumptions and change course.” (The “soldier mindset,” by contrast, encourages you to defend your positions at any cost.) Galef tends to see rationality as a method for acquiring more accurate views. Pinker, a cognitive and evolutionary psychologist, sees it instrumentally, as “the ability to use knowledge to attain goals.” By this definition, to be a rational person you have to know things, you have to want things, and you have to use what you know to get what you want. Intentions matter: a person isn’t rational, Pinker argues, if he solves a problem by stumbling on a strategy “that happens to work.”

Introspection is key to rationality. A rational person must practice what the neuroscientist Stephen Fleming, in “Know Thyself: The Science of Self-Awareness” (Basic Books), calls “metacognition,” or “the ability to think about

our own thinking”—“a fragile, beautiful, and frankly bizarre feature of the human mind.” Metacognition emerges early in life, when we are still struggling to make our movements match our plans. (“Why did I do that?” my toddler asked me recently, after accidentally knocking his cup off the breakfast table.) Later, it allows a golfer to notice small differences between her first swing and her second, and then to fine-tune her third. It can also help us track our mental actions. A successful student uses metacognition to know when he needs to study more and when he’s studied enough: essentially, parts of his brain are monitoring other parts.

In everyday life, the biggest obstacle to metacognition is what psychologists call the “illusion of fluency.” As we perform increasingly familiar tasks, we monitor our performance less rigorously; this happens when we drive, or fold laundry, and also when we think thoughts we’ve thought many times before. Studying for a test by reviewing your notes, Fleming writes, is a bad idea, because it’s the mental equivalent of driving a familiar route. “Experiments have repeatedly shown that testing ourselves—forcing ourselves to practice exam questions, or writing out what we know—is more effective,” he writes. The trick is to break the illusion of fluency, and to encourage an “awareness of ignorance.”

Fleming notes that metacognition is a skill. Some people are better at it than others. Galef believes that, by “calibrating” our metacognitive minds, we can improve our performance and so become more rational. In a section of her book called “Calibration Practice,” she offers readers a collection of true-or-false statements (“Mammals and dinosaurs coexisted”; “Scurvy is caused by a deficit of Vitamin C”); your job is to weigh in on the veracity of each statement while also indicating whether you are fifty-five, sixty-five, seventy-five, eighty-five, or ninety-five per cent confident in your determination. A perfectly calibrated individual, Galef suggests, will be right seventy-five per cent of the time about the answers in which she is seventy-five per cent confident. With practice, I got fairly close to “perfect calibration”: I still answered some questions wrong, but I was right about how wrong I would be.

There are many calibration methods. In the “equivalent bet” technique, which Galef attributes to the decision-making expert Douglas Hubbard, you imagine that you’ve been offered two ways of winning ten thousand dollars: you can either bet on the truth of some statement (for instance, that self-driving cars will be on the road within a year) or reach blindly into a box full of balls in the hope of retrieving a marked ball. Suppose the box contains four balls. Would you prefer to answer the question, or reach into the box? (I’d prefer the odds of the box.) Now suppose the box contains twenty-four balls—would your preference change? By imagining boxes with different numbers of balls, you can get a sense of how much you really believe in your assertions. For Galef, the box that’s “equivalent” to her belief in the imminence of self-driving cars contains nine balls, suggesting that she has eleven-per-cent confidence in that prediction. Such techniques may reveal that our knowledge is more fine-grained than we realize; we just need to look at it more closely. Of course, we could be making out detail that isn’t there.

**K**nowing about what you know is Rationality 101. The advanced coursework has to do with changes in your knowledge. Most of us stay informed straightforwardly—by taking in new information. Rationalists do the same, but self-consciously, with an eye to deliberately redrawing their mental maps. The challenge is that news about distant territories drifts in from many sources; fresh facts and opinions aren’t uniformly significant. In recent decades, rationalists confronting this problem have rallied behind the work of Thomas Bayes, an eighteenth-century mathematician and minister. So-called Bayesian reasoning—a particular thinking technique, with its own distinctive jargon—has become de rigueur.

There are many ways to explain Bayesian reasoning—doctors learn it one way and statisticians another—but the basic idea is simple. When new information comes in, you don’t want it to replace old information wholesale. Instead, you want it to modify what you already know to an appropriate degree. The degree of modification depends both on your confidence in your preëxisting knowledge and on the value of the new data. Bayesian reasoners begin with

what they call the “prior” probability of something being true, and then find out if they need to adjust it.

Consider the example of a patient who has tested positive for breast cancer—a textbook case used by Pinker and many other rationalists. The stipulated facts are simple. The prevalence of breast cancer in the population of women—the “base rate”—is one per cent. When breast cancer is present, the test detects it ninety per cent of the time. The test also has a false-positive rate of nine per cent: that is, nine per cent of the time it delivers a positive result when it shouldn’t. Now, suppose that a woman tests positive. What are the chances that she has cancer?

When actual doctors answer this question, Pinker reports, many say that the woman has a ninety-per-cent chance of having it. In fact, she has about a nine-per-cent chance. The doctors have the answer wrong because they are putting too much weight on the new information (the test results) and not enough on what they knew before the results came in—the fact that breast cancer is a fairly infrequent occurrence. To see this intuitively, it helps to shuffle the order of your facts, so that the new information doesn’t have pride of place. Start by imagining that we’ve tested a group of a thousand women: ten will have breast cancer, and nine will receive positive test results. Of the nine hundred and ninety women who are cancer-free, eighty-nine will receive false positives. Now you can allow yourself to focus on the one woman who has tested positive. To calculate her chances of getting a true positive, we divide the number of positive tests that actually indicate cancer (nine) by the total number of positive tests (ninety-eight). That gives us about nine per cent.

Bayesian reasoning is an approach to statistics, but you can use it to interpret all sorts of new information. In the early hours of September 26, 1983, the Soviet Union’s early-warning system detected the launch of intercontinental ballistic missiles from the United States. Stanislav Petrov, a forty-four-year-old duty officer, saw the warning. He was charged with reporting it to his superiors, who probably would have launched a nuclear counterattack. But Petrov, who in all likelihood had never heard of Bayes, nevertheless employed

Bayesian reasoning. He didn't let the new information determine his reaction all on its own. He reasoned that the probability of an attack on any given night was low—comparable, perhaps, to the probability of an equipment malfunction. Simultaneously, in judging the quality of the alert, he noticed that it was in some ways unconvincing. (Only five missiles had been detected—surely a first strike would be all-out?) He decided not to report the alert, and saved the world.

Bayesian reasoning implies a few “best practices.” Start with the big picture, fixing it firmly in your mind. Be cautious as you integrate new information, and don't jump to conclusions. Notice when new data points do and do not alter your baseline assumptions (most of the time, they won't alter them), but keep track of how often those assumptions seem contradicted by what's new. Beware the power of alarming news, and proceed by putting it in a broader, real-world context.

In a sense, the core principle is *mise en place*. Keep the cooked information over here and the raw information over there; remember that raw ingredients often reduce over heat. But the real power of the Bayesian approach isn't procedural; it's that it replaces the facts in our minds with probabilities. Where others might be completely convinced that G.M.O.s are bad, or that Jack is trustworthy, or that the enemy is Eurasia, a Bayesian assigns probabilities to these propositions. She doesn't build an immovable world view; instead, by continually updating her probabilities, she inches closer to a more useful account of reality. The cooking is never done.

**A**ppplied to specific problems—Should you invest in Tesla? How bad is the Delta variant?—the techniques promoted by rationality writers are clarifying and powerful. But the rationality movement is also a social movement; rationalists today form what is sometimes called the “rationality community,” and, as evangelists, they hope to increase its size. The rationality community has its own *lingua franca*. If a rationalist wants to pay you a big compliment, she might tell you that you have caused her to “revise her priors”—that is, to alter some of her well-justified prior assumptions. (On her

mental map, a mountain range of possibilities has gained or lost probabilistic altitude.) That same rationalist might talk about holding a view “on the margin”—a way of saying that an idea or fact will be taken into account, as a kind of tweak on a prior, the next time new information comes in.

(Economists use the concept of “marginal utility” to describe how we value things in series: the first nacho is delightful, but the marginal utility of each additional nacho decreases relative to that of a buffalo wing.) She might speak about “updating” her opinions—a cheerful and forward-looking locution, borrowed from the statistical practice of “Bayesian updating,” which rationalists use to destigmatize the act of admitting a mistake. In use, this language can have a pleasingly deliberate vibe, evoking the feeling of an edifice being built. “Every so often a story comes along that causes me to update my priors,” the economist Tyler Cowen wrote, in 2019, in response to the [Jeffrey Epstein](#) case. “I am now, at the margin, more inclined to the view that what keeps many people on good behavior is simply inertia.”

In Silicon Valley, people wear T-shirts that say “Update Your Priors,” but talking like a rationalist doesn’t make you one. A person can drone on about base rates with which he’s only loosely familiar, or say that he’s revising his priors when, in fact, he has only ordinary, settled opinions. Google makes it easy to project faux omniscience. A rationalist can give others and himself the impression of having read and digested a whole academic subspecialty, as though he’d earned a Ph.D. in a week; still, he won’t know which researchers are trusted by their colleagues and which are ignored, or what was said after hours at last year’s conference. There’s a difference between reading about surgery and actually being a surgeon, and the surgeon’s priors are what we really care about. In a recent interview, Cowen—a superhuman reader whose blog, [Marginal Revolution](#), is a daily destination for info-hungry rationalists—told Ezra Klein that the rationality movement has adopted an “extremely culturally specific way of viewing the world.” It’s the culture, more or less, of winning arguments in Web forums. Cowen suggested that to understand reality you must not just read about it but see it firsthand; he has grounded his priors in visits to about a hundred countries, once getting caught in a shoot-

out between a Brazilian drug gang and the police.

Clearly, we want people in power to be rational. And yet the sense that rationalists are somehow unmoored from direct experience can make the idea of a rationalist with power unsettling. Would such a leader be adrift in a matrix of data, more concerned with tending his map of reality than with the people contained in that reality? In a sketch by the British comedy duo Mitchell and Webb, a government minister charged with ending a recession asks his analysts if they've considered "killing all the poor." "I'm not saying do it—I'm just saying run it through the computer and see if it would work," he tells them. (After they say it won't, he proposes "blue-skying" an even more senseless alternative: "Raise V.A.T. and kill all the poor.") This caricature echoes a widespread skepticism of rationality as a value system. When the Affordable Care Act was wending its way through Congress, conservatives worried that similar proposals would pop up on "death panels," where committees of rational experts would suggest lowering health-care costs by killing the aged. This fear, of course, was sharpened by the fact that we really do spend too much money on health care in the last few years of life. It's up to rationalists to do the uncomfortable work of pointing out uncomfortable truths; sometimes in doing this they seem a little too comfortable.

In our personal lives, the dynamics are different. Our friends don't have power over us; the best they can do is nudge us in better directions. Elizabeth Bennet, the protagonist of "Pride and Prejudice," is intelligent, imaginative, and thoughtful, but it's Charlotte Lucas, her best friend, who is rational. Charlotte uses Bayesian reasoning. When their new acquaintance, Mr. Darcy, is haughty and dismissive at a party, she gently urges Lizzy to remember the big picture: Darcy is "so very fine a young man, with family, fortune, everything in his favour"; in meeting him, therefore, one's prior should be that rich, good-looking people often preen at parties; such behavior is not, in itself, revelatory. When Charlotte marries Mr. Collins, an irritating clergyman with a secure income, Lizzy is appalled at the match—but Charlotte points out that the success of a marriage depends on many factors, including financial ones, and

suggests that her own chances of happiness are “as fair as most people can boast on entering the marriage state.” (In modern times, the base rates would back her up: although almost fifty per cent of marriages end in divorce, the proportion is lower among higher-income people.) It’s partly because of Charlotte’s example that Lizzy looks more closely at Mr. Darcy, and discovers that he is flawed in predictable ways but good in unusual ones. Rom-com characters often have passionate friends who tell them to follow their hearts, but Jane Austen knew that really it’s rational friends we need.

In fact, as Charlotte shows, the manner of a kind rationalist can verge on courtliness, which hints at deeper qualities. Galef describes a typically well-mannered exchange on the now defunct Web site ChangeAView. A male blogger, having been told that one of his posts was sexist, strenuously defended himself at first. Then, in a follow-up post titled “Why It’s Plausible I’m Wrong,” he carefully summarized the best arguments made against him; eventually, he announced that he’d been convinced of the error of his ways, apologizing not just to those he’d offended but to those who had sided with him for reasons that he now believed to be mistaken. Impressed by his sincere and open-minded approach, Galef writes, she sent the blogger a private message. Reader, they got engaged.

The rationality community could make a fine setting for an Austen novel written in 2021. Still, we might ask, How much credit should rationality get for drawing Galef and her husband together? It played a role, but rationality isn’t the only way to understand the traits she perceived. I’ve long admired my friend Greg for his rationality, but I’ve since updated my views. I think it’s not rationality, as such, that makes him curious, truthful, honest, careful, perceptive, and fair, but the reverse.

**I**n “Rationality,” “The Scout Mindset,” and other similar books, irrationality is often presented as a form of misbehavior, which might be rectified through education or socialization. This is surely right in some cases, but not in all. One spring, when I was in high school, a cardinal took to flying at our living-room window, and my mother—who was perceptive, funny, and

intelligent, but not particularly rational—became convinced that it was a portent. She'd sometimes sit in an armchair, waiting for it, watchful and unnerved. Similar events—a torn dollar bill found on the ground, a flat tire on the left side of the car rather than the right—could cast shadows over her mood for days, sometimes weeks. As a voter, a parent, a worker, and a friend, she was driven by emotion. She had a stormy, poetic, and troubled personality. I don't think she would have been helped much by a book about rationality. In a sense, such books are written for the already rational.

My father, by contrast, is a doctor and a scientist by profession and disposition. When I was a kid, he told me that Santa Claus wasn't real long before I figured it out; we talked about physics, computers, biology, and "Star Trek," agreeing that we were Spocks, not Kirks. My parents divorced decades ago. But recently, when my mother had to be discharged from a hospital into a rehab center, and I was nearly paralyzed with confusion about what I could or should do to shape where she'd end up, he patiently, methodically, and judiciously walked me through the scenarios on the phone, exploring each forking path, sorting the inevitabilities from the possibilities, holding it all in his head and communicating it dispassionately. All this was in keeping with his character.

I've spent decades trying to be rational. So why did I feel paralyzed while trying to direct my mother's care? Greg tells me that, in his business, it's not enough to have rational thoughts. Someone who's used to pondering questions at leisure might struggle to learn and reason when the clock is ticking; someone who is good at reaching rational conclusions might not be willing to sign on the dotted line when the time comes. Greg's hedge-fund colleagues describe as "commercial"—a compliment—someone who is not only rational but timely and decisive. An effective rationalist must be able to short the mortgage market today, or commit to a particular rehab center now, even though we live in a world of Bayesian probabilities. I know, rationally, that the coronavirus poses no significant risk to my small son, and yet I still hesitated before enrolling him in daycare for this fall, where he could make friends. You

can know what's right but still struggle to do it.

Following through on your own conclusions is one challenge. But a rationalist must also be “metarational,” willing to hand over the thinking keys when someone else is better informed or better trained. This, too, is harder than it sounds. Intellectually, we understand that our complex society requires the division of both practical and cognitive labor. We accept that our knowledge maps are limited not just by our smarts but by our time and interests. Still, like Gurri's populists, rationalists may stage their own contrarian revolts, repeatedly finding that no one's opinions but their own are defensible. In letting go, as in following through, one's whole personality gets involved. I found it possible to be metarational with my dad not just because I respected his mind but because I knew that he was a good and cautious person who had my and my mother's best interests at heart. I trusted that, unlike the minister in the Mitchell and Webb sketch, he would care enough to think deeply about my problem. Caring is not enough, of course. But, between the two of us, we had the right ingredients—mutual trust, mutual concern, and a shared commitment to reason and to act.

The realities of rationality are humbling. Know things; want things; use what you know to get what you want. It sounds like a simple formula. But, in truth, it maps out a series of escalating challenges. In search of facts, we must make do with probabilities. Unable to know it all for ourselves, we must rely on others who care enough to know. We must act while we are still uncertain, and we must act in time—sometimes individually, but often together. For all this to happen, rationality is necessary, but not sufficient. Thinking straight is just part of the work. ♦