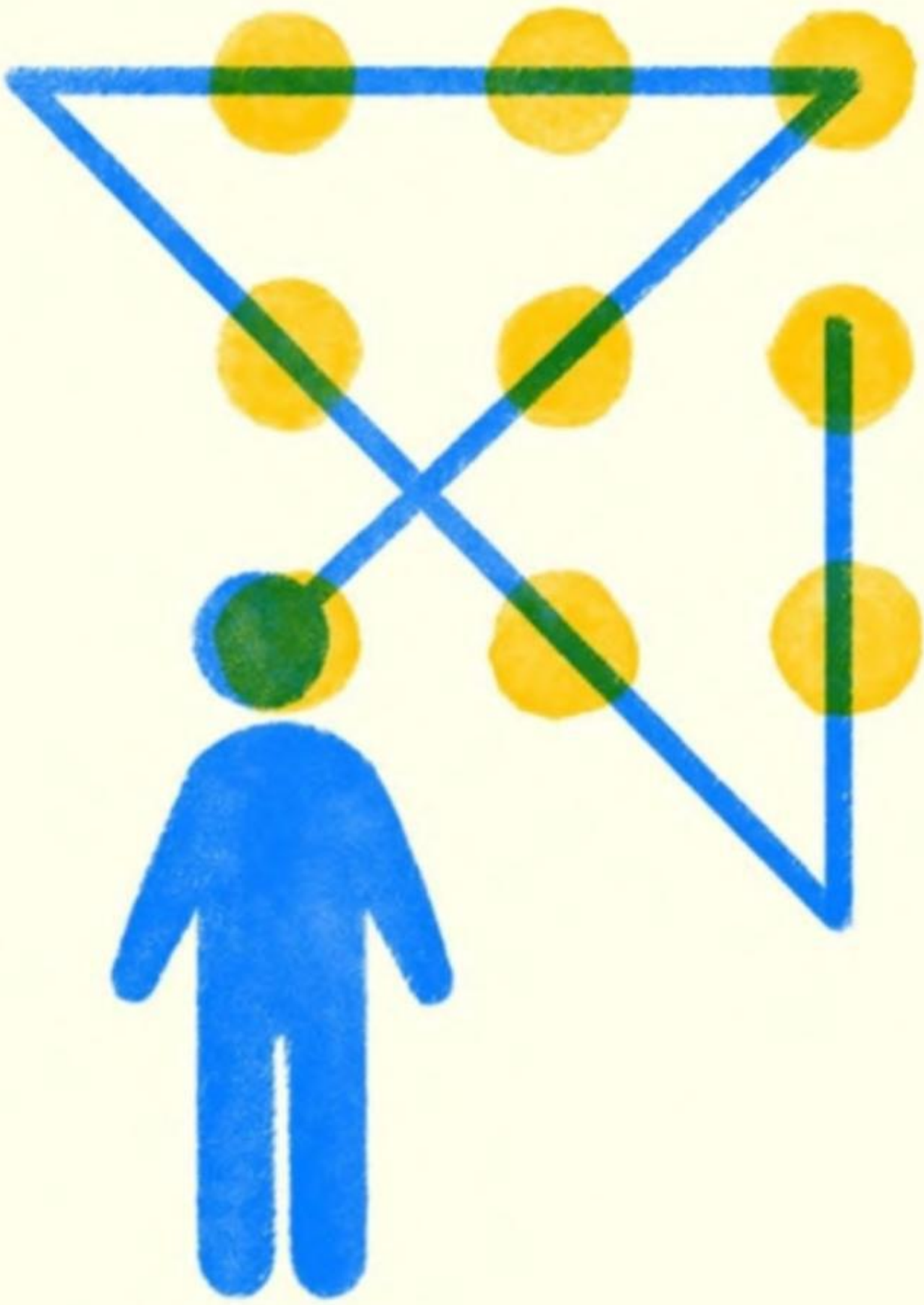
Human Advantage in an Age of Technology and Turmoil

framers



Kenneth Cukier

Viktor Mayer-Schönberger

Francis de Véricourt

framers

human advantage in an age of technology and turmoil

Kenneth Cukier,
Viktor Mayer-Schönberger,
Francis de Véricourt

DUTTON

DUTTON

An imprint of Penguin Random House LLC
penguinrandomhouse.com

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Names: Cukier, Kenneth, author. | Mayer-Schönberger, Viktor, author. | Véricourt, Francis de, author.
Title: Framers: human advantage in an age of technology and turmoil / Kenneth Cukier,
Viktor Mayer-Schönberger, Francis de Véricourt.
Description: [New York, New York]: Dutton, [2021] | Includes bibliographical references and index.
Identifiers: LCCN 2020051813 (print) | LCCN 2020051814 (ebook) |
ISBN 9780593182598 (hardcover) | ISBN 9780593182604 (ebook)
Subjects: LCSH: Cognitive maps (Psychology) | Frames (Sociology) | Reasoning (Psychology)
Classification: LCC BF314 .C85 2021 (print) | LCC BF314 (ebook) |
DDC 153—dc23
LC record available at https://lccn.loc.gov/2020051813
LC ebook record available at https://lccn.loc.gov/2020051814

ISBN 9780593185049 (export edition)

Printed in the United States of America
ScoutAutomatedPrintCode

Book design by Nancy Resnick

contents

1

decisions

the source of human power is neither muscle nor mind but models

Some threats are sudden and unexpected. Others are slow and smoldering. Both represent cognitive blind spots for which societies are unprepared. Whether pandemics or populism, new weapons or new technologies, global warming or gaping inequalities, how humans respond marks the difference between survival and extinction. And how we act depends on what we see.

Each year, more than 700,000 people around the world die from infections that antibiotics once cured but no longer do. The bacteria have developed resistance. The number of deaths is rising fast. Unless a solution is found, it is on track to hit ten million a year, or one person every three seconds. It makes even the tragedy of Covid-19 pale by comparison. And it is a problem that society itself has produced. Antibiotics work less and less well due to overuse: the very drugs that could once staunch the bacteria have turned them into superbugs.

We take antibiotics for granted, but before penicillin was discovered in 1928 and mass-produced more than a decade later, people

routinely died from broken bones or simple scratches. In 1924, the sixteen-year-old son of American president Calvin Coolidge got a blister on his toe while playing tennis on the White House lawn. It became infected, and he died within the week—neither his status nor wealth could save him. Today, almost every aspect of medicine, from a C-section to cosmetic surgery to chemotherapy, relies on antibiotics. If their power were to wane those treatments would become far riskier.

From her colorful, plant-strewn office in Cambridge, Massachusetts, Regina Barzilay, a professor of artificial intelligence at MIT, envisioned a solution. Conventional drug development mostly focuses on finding substances with molecular "fingerprints" similar to ones that work. That generally performs well, but not for antibiotics. Most substances with similar compositions have already been examined, and new antibiotics are so close in structure to existing ones that bacteria quickly develop resistance to them, too. So Barzilay and a diverse team of biologists and computer scientists, led by Jim Collins, a professor of bioengineering at MIT, embraced an alternative approach. What if, instead of looking for structural similarities, they focused on the effect: Did it kill bacteria? They reconceived the problem not as a biological one but an informational one.

Charismatic and confident, Barzilay doesn't come across as a typical nerd. But then, she is accustomed to defying categories. She grew up under communism in what is now Moldova, speaking Russian; was educated in Israel, speaking Hebrew; and attended grad school in America. In 2014, as a new mother in her early forties, she was diagnosed with breast cancer, which she survived after difficult treatments. This ordeal led her to change her research in order to focus on artificial intelligence in medicine. As her research gained attention, a MacArthur "genius grant" followed.

Barzilay and the team got to work. They trained an algorithm on

more than 2,300 compounds with antimicrobial properties, to find if any inhibited the growth of *E. coli,* a noxious bacterium. Then the model was applied to around six thousand molecules in the Drug Repurposing Hub and later to more than one hundred million molecules in another database to predict which might work. In early 2020 they struck gold. One molecule stood out. They named it "halicin" after HAL, the renegade computer in *2001: A Space Odyssey.*

The discovery of a superdrug to kill superbugs made headlines around the world. It was hailed as a "video killed the radio star" moment for the superiority of machine over man. "AI Discovers Antibiotics to Treat Drug-Resistant Diseases," boomed a front-page headline in the *Financial Times.*

But that missed the real story. It wasn't a victory for artificial intelligence but a success of human cognition: the ability to rise up to a critical challenge by conceiving of it in a certain way, altering aspects of it, which open up new paths to a solution. Credit does not go to a new technology but to a human ability.

"Humans were the ones who selected the right compounds, who knew what they were doing when they gave the material for the model to learn from," Barzilay explains. People defined the problem, designed the approach, chose the molecules to train the algorithm, and then selected the database of substances to examine. And once some candidates popped up, humans reapplied their biological lens to understand why it worked.

The process of finding halicin is more than an outstanding scientific breakthrough or a major step toward accelerating and lowering the cost of drug development. To succeed, Barzilay and the team needed to harness a form of cognitive freedom. They didn't get the idea from a book, from tradition, or by connecting obvious dots. They got it by embracing a unique cognitive power that all people possess.

Mental Models and the World

Humans think using mental models. These are representations of reality that make the world comprehensible. They allow us to see patterns, predict how things will unfold, and make sense of the circumstances we encounter. Reality would otherwise be a flood of information, a jumble of inchoate experiences and sensations. Mental models bring order. They let us focus on essential things and ignore others—just as, at a cocktail party, we can hear the conversation that we're in while tuning out the chatter around us. We craft a simulation of reality in our minds to anticipate how situations will play out.

We use mental models all the time, even if we are not aware of them. But there are moments when we are acutely conscious of how we size up a situation, and can deliberately maintain or change our perspective. This often happens when we need to make a high-consequence decision, such as whether to switch jobs, become a parent, buy a home, close a factory, or build a skyscraper. In those instances, it can become apparent to us that our decisions are not simply based on the reasoning we apply, but on something more foundational: the particular lens through which we look at the situation—our sense of how the world works. That underlying level of cognition consists of mental models.

The fact that we need to interpret the world in order to exist in it, that how we perceive reality colors how we act within it, is something that people have long known but take for granted. It is what made Regina Barzilay's achievement so impressive. She conceived of the problem in the right way. She applied a mental model, shifting her focus from the structure of the molecule (that is, the mechanism by which it worked) to its function (that is, whether it worked

at all). By framing the problem differently, she and the team achieved a discovery that had eluded others.

Barzilay was a framer. By correctly framing the situation, she could unlock new solutions.

The mental models that we choose and apply are frames: they determine how we understand and act in the world. Frames enable us to generalize and make abstractions that apply to other situations. With them, we can handle new situations, rather than having to relearn everything from scratch. Our frames are always operating in the background. But we can stop and deliberately ask ourselves which frame we are applying, and whether it is the best fit for the circumstances. And if it's not, we can choose another frame that is better. Or, we can invent a new frame altogether.

Framing is so fundamental to human cognition that even those who study the workings of the mind rarely focused on it until relatively recently. Its importance was overshadowed by other mental capabilities, such as sensing and memory. But as people have become more aware of the need to improve their decision-making, the role of frames as fundamental to choosing and acting well has moved from the background to center stage. We now know that the right frame applied in the right way opens up a wider range of possibilities, which in turn leads to better choices. The frames we employ affect the options we see, the decisions we make, and the results we attain. By being better at framing, we get better outcomes.

Many of society's thorniest problems involve, at their core, a friction over the way an issue is framed. Should America build a wall with other countries or a bridge? Should Scotland remain within the United Kingdom or declare independence? Should China's "one country, two systems" policy toward Hong Kong emphasize the first or last part of that phrase? People can look at the same situation and see different things because they frame it differently.

When San Francisco 49ers quarterback Colin Kaepernick took a knee during the national anthem in 2016 to call attention to racism and police violence, some saw it as a respectful form of quiet, symbolic protest. After all, he didn't turn his back or raise a fist—or a finger. Others saw it as a grotesque disrespect for the country, an antagonistic publicity stunt by a mediocre player that brought the culture wars to one of the few areas of American life that hadn't yet been spoiled by them. The argument wasn't over what happened but what it meant. It was a Rorschach test: what people saw in it depended on what frame they brought to it.

Each frame lets us see the world from a distinct vantage point. Frames magnify certain elements and minimize others. The capitalist frame shows us commercial opportunity everywhere; the communist frame reduces everything to class struggle. The industrialist looks at a rain forest and sees timber that is valuable today, while the environmentalist sees the "lungs of the planet," vital for long-term survival. Should people be mandated to wear a mask in public amid a pandemic? In the United States those using the health frame stated "yes, absolutely," while those applying the freedom frame cried "hell no!" Same data, different frames, opposite conclusions.

Sometimes our frames don't fit the reality to which we apply them. There is no such thing as a "bad" frame per se (save for one exception that we'll raise later), but there are certainly cases of misframing, where a given frame doesn't fit very well. In fact, the path of human progress is littered with the carcasses of misused frames. Take the fifteenth-century anatomy book *Fasciculus Medicinae*. It associated body parts with zodiac signs—a pleasing symmetry between the heavens above and the organs within. But that frame never cured anyone, and it was cast aside as more useful frames came along.

We make similar mistakes in our own day. In 2008 Nokia led the

world in mobile phone sales. When Apple introduced the iPhone, few thought it would take off. The trend was to make handsets smaller and cheaper, but Apple's was bulkier, pricier, and buggier. Nokia's frame came from the conservative telecom industry, valuing practicality and reliability. Apple's frame came from the breathlessly innovative computing industry, valuing ease of use and the extensibility of new features via software. That frame turned out to be a better fit for the needs and wants of consumers—and Apple dominated the market.

Misapplying frames can have horrendous consequences. In the 1930s the Soviet Union followed Lysenkoism, a theory of plant genetics. It was based on Marxist-Leninist ideology, not botany. Among its precepts was that crops can be planted close together because, according to communist theory, members of the same class live in solidarity and do not compete for resources.

Taking a communist frame from economics and applying it to farming was lunacy, but the country's leaders made it the basis of their agriculture policy. Its proponent, Trofim Lysenko, had the favor of Stalin himself. Scientists who questioned his findings were fired, imprisoned, exiled, or executed. The great Russian biologist Nikolai Vavilov criticized Lysenko's science and was sentenced to death. As for the result of Lysenkoism? Although the country increased the area of land under cultivation a hundredfold, yields fell as crops died or rotted. The misapplied frame contributed to tragic famines that cost millions of lives.

If a frame doesn't fit, the good news is that we can use a different one, or invent a new and better one. Some new frames are responsible for world-changing breakthroughs. Darwin's theory of evolution provided an explanation of the origins of life without reliance on religion. Similarly, Newtonian physics explained the movement of physical objects in space for centuries, but over time there were

phenomena that it couldn't account for. Einstein reframed physics by showing that time, long considered constant, is actually relative.

It is easiest to see the value of frames in the sciences, where they are (or at least should be) explicit, and where researchers document the mental inputs they use to reach their conclusions. Yet when it comes to the vast challenges that humanity faces today, we often fail to notice which frames we are applying. Understanding the power of framing in all domains is vital. We need to see problems differently in order to solve them. The centerpiece of our response to our most difficult dilemmas—whether at the level of the individual, community, country, or indeed civilization—lies within us: our unique human ability to frame.

But we need to get better at it. This book explains how.

Seeing What Isn't There

Over the past few decades, a revolution in cognitive psychology and decision theory has placed mental models at the center of how people live and think. Framing commonly happens subconsciously. But people who consistently make good decisions, or are in positions where they need to make high-stakes ones, are aware of framing and their ability to reframe. This affects the options they see and the actions they take.

When a venture capitalist scrutinizes an investment, a military officer thinks through an operation, or an engineer approaches a technical problem, they have to frame the issue. Need to decide whether to build a wind park in one location or another, or go with a solar farm instead? The information we gather is only part of the decision-making process. More important, in many respects, is how one sizes up the situation itself: how one frames it.

Yet framing is not only for high-stakes matters. It affects our everyday lives as well. We are continually confronted with questions that require having a model of the world in our mind. How can I get along better with my partner? How can I impress my boss? How can I rearrange my life to be healthier? And wealthier? Framing is just as essential for these types of questions. It undergirds our thoughts, affecting what we perceive and how we think. By making our frames apparent and learning how to deliberately choose and apply them, we can improve our lives and our world.

Put simply: we can turn framing from a basic feature of human cognition into a practical tool we can use to make better decisions.

Our mind uses frames to capture the most salient aspects of the world, and filter out the others—we couldn't comprehend life in all of its intricate complexity otherwise. By mentally modeling the world, we keep it manageable and thus actionable. In this sense, frames simplify reality. But they aren't dumbed-down versions of the world. They concentrate our thinking on the critical parts.

Frames also help us to learn from single experiences and come up with general rules that we can apply to other situations—including ones that have not yet happened. They enable us to know something about the unobserved and even the unobservable; to imagine things for which no data exists. Frames let us see what isn't there. We can ask "What if?" and foresee how different decisions might play out. It is this ability to envision other realities that makes possible individual achievement and societal progress.

Humans have long looked to the sky and wished to fly. We now do so—but not like a bird. Yet no amount of data and processing power could have taken a plethora of bicycle parts and conceived of an airplane, as the Wright brothers did in 1903. A mental model was needed; a frame. Likewise, humans dreamed of seeing inside the body without cutting the skin. And today we do, but with technologies like X-rays, not with our naked eyes. For that too a new

conceptualization was required, a frame for how to use the electromagnetic radiation known as Röntgen rays in 1895.

Some of the things we use every day are the result of changing the way they were initially framed—sometimes comically so. The telephone was first thought about as a way to listen to music remotely: people would dial in to hear a concert. The phonograph was considered to be a way to communicate messages: a company president might send audio memos on grooved cylinders to far-flung managers. Only when these uses were flipped around did the technologies catch on. Thomas Edison in the early 1900s believed motion pictures would replace classrooms—a vision only realized a century later when Zoom became the new schoolhouse.

The term *framing* is well established in the social sciences. The psychologists Daniel Kahneman and Amos Tversky eloquently explained how different characterizations of outcomes influence decision-making—which they called the "framing effect," and described it as a flaw in human reasoning. Though we share the same term, the meaning here is somewhat different: not how something is positioned but a deliberate act of harnessing mental models to elicit options prior to making a decision. Although the misframing of a situation can certainly lead to flawed decisions, framing is a valuable and empowering human capability. It lets us make sense of the world and reshape it. We would not be who we are, as individuals or as a species, without it.

It may be tempting to understand instances of reframing as a paradigm shift; that is, a fundamental change in the prevailing concepts and practices in a domain. In 1962 Thomas Kuhn, a philosopher of science, argued that paradigm shifts drive scientific progress. But the comparison is inexact. Every paradigm shift is a reframing, such as when Copernicus's heliocentric view overturned Ptolemy's geocentric model of Earth and the sun. But not every reframing is a para-

digm shift—reframing happens comparatively frequently. Sometimes it changes society's concept of the world; far more often it leads to some small but significant change in our individual lives. In each case, a successful act of reframing leads to better decisions.

Working with frames may sound complicated and difficult. It does require skill. Yet humans are surprisingly good at it. We have been doing it for tens of thousands of years, even though we don't always realize it.

Frames are more than a person's individual perspective—they are cognitive templates. But the concept of perspective is a useful metaphor. Before the Italian architect Filippo Brunelleschi began drawing with geometric perspective around 1420, artists painted the world as flat and positioned objects relative to their implied importance. Taking their cue from Brunelleschi, artists learned to depict depth and draw scenes as they actually appeared. Comparing pictures from before and after this change is one way to realize the potency of switching to a new frame.

We are all framers. We make predictions, from the mundane to the spectacular. In this way, we constantly perform a sort of mental time-travel. Some of us are better at it than others, and there are ways in which we can all improve. In fact, we must.

Dreaming with Constraints

Frames help to do two tasks really well, which we outline in chapter 2. First, in novel situations or when circumstances change, our ability to choose a frame provides us with new options. Second, and at least as important, in situations that are familiar, frames focus our mind, thereby reducing our cognitive load. It's an incredibly efficient way for us to reach a suitable decision. To do that successfully,

three elements are in play: causal thinking, our aptitude to create counterfactuals, and our capacity to constrain and shape our imagination toward a particular goal. Consider each element in turn.

Chapter 3 looks at causality. Humans see the world through the lens of cause and effect. This makes the world understandable. We can predict in advance what a given action will provoke and repeat it to our advantage. Causal reasoning is the foundation of our cognition. Children learn to think causally as they grow up, and thinking in cause and effect has enabled human societies to evolve over time. We are causal-inference engines.

Our causal inferences are often wrong, because the world is complex. It's hard to suss out all the intricacies with the three pounds of spongy fat and protein in our skulls. We now depend on scientific methodology to prevent us from jumping to causal conclusions too quickly—for example, that performing a special dance can cause rain to fall. But our tendency to see causes everywhere has a valuable consequence: it gives us a tool to understand the world and to place it more or less under our control.

The second element of our mental models is counterfactuals, which we examine in chapter 4. These are imagined alternatives to reality; hypotheses of a world in which one or several things are changed. As with causality, we think in counterfactuals all the time. They come naturally to us. Counterfactuals let us escape the cognitive here and now: we are not locked into the reality before us—in our mind's eye we can invent a new one.

Counterfactual thinking is crucial to progress. People can envision what does not exist, as a way to understand the actual world and conceive of how it might be different. We do this by asking what-if questions. Our imaginings need not be meaningless daydreams; they are an essential precursor for action, an element of our preparation to make decisions. Often when we imagine and visual-

ize, we are conjuring and assessing counterfactuals. This is what children do when they engage in pretend play and what scientists do when they devise experiments.

Though we can improve the world through counterfactual thinking, it might be foolish to stray too far into the realm of the unrealistic. Chapter 5 introduces the third element of framing, the role of constraints. The right constraints help us to adjust our imagination so that our counterfactuals remain actionable, showing us actions that are actually possible. Framing isn't an exercise of following flights of fancy, untethered balloons tossed by the wind upward to nowhere. It's dreaming within limits. Our counterfactuals are useful only when they are bounded.

Constraints act as the glue that holds the mental model together, so that we can think *What if?* in a structured and deliberate way. If we have a flat tire and have never changed a tire before, we don't look to Starfleet antigravity to help us; we look at the tools we find in the trunk (say, a car jack and a wrench) and visualize how these could work for us.

These three dimensions of cognition—causality, counterfactuals, and constraints—together form the basis of framing. They are our tools for seeing beyond the obvious and thinking forward.

However, sometimes we need to change frames, especially if the context of a problem has shifted. Chapter 6 looks at our ability to sift through our collection of frames to find one that is better suited to a situation. Or we can repurpose an existing frame from one domain and use it in another—deliberately expanding our range by amending that frame to fit a new context and goal.

As we gain experience, we develop a repertoire of frames and can trot out a suitable one when we need to, much as good pianists have a repertoire of styles that they use. Accomplished musicians can instantly find the underlying key, tempo, rhythm, and harmony in a

wide variety of pieces they've never heard before. This is the essence of improvisation. Each genre is very different, with its own set of rules. It is hard for a pianist to switch from lovelorn Chopin to Lady Gaga. But reframing is as possible on the ebony and ivory as it is in life.

Choosing among different frames does have its limits, however. In some cases, the right frame may not exist. What's needed then is an act of invention, to devise a completely new one. We laud those who create entirely new frames, since, when successful, they end up changing the world.

How we can become better framers is explained in chapter 7. Learning this skill depends on embracing a diversity of frames. We offer three strategies to achieve this: expanding our repertoire, developing our curiosity through a process we call "cognitive foraging," and if all else fails, having the courage to leap into the cognitive unknown.

Chapter 8 considers the importance of framing for communities, countries, and civilization writ large. The goal is pluralism: fostering and celebrating differences, instead of aspiring to a homogenous way of seeing the world. We need to promote the "colorful," not believe that we achieve some form of valor by deliberately making ourselves "color-blind."

What we cannot allow are frames that deny the existence of other frames (the sole exception to the idea that there are no bad frames). Uniformity of mental models is what crushes human progress. It makes people no wiser than automata that perpetuate the past because they cannot see beyond the present. If our frames cannot coexist, how can we?

The true heroes of human framing do not live in palaces, manage big companies, or teach at fancy universities. They are all around us—indeed, they are us. As individuals, our framing matters because

it shapes our lives. But for humanity, what matters is the richness of the mental models we collectively possess. Pundits often suggest that people must come together and converge on a perspective to meet pressing challenges. But it is quite the opposite: our power lies in the difference of human frames and in our ability to see the world from a myriad of angles. Only if we tap the breadth of human frames can we devise the original solutions we will need to survive as a species.

The Machine and the Mob

Belief in the value and power of framing is under threat. Around the world, people are losing confidence in humanity's cognitive abilities, and turning to solutions that deny the role of mental models. On one side are those who place their faith in the machine, and on the other are those who accept the rough justice and easy answers of the mob.

The hyper-rationalists, who embrace facts and value reason, represent the first side. They believe we can rely exclusively on data and algorithms to solve our myriad problems, and are tempted to bypass human framing for artificial intelligence. They are not just a handful of geeks waiting for the "singularity," the moment when computers are smarter than people. A growing cohort looks to technology to deliver exactly the kind of detached, objective superrational decision-making that people seem to be incapable of. For them, there is no doubt that humans will remain in the seat of power but that everyday judgments will be handled by machines.

Need to drive to Delhi? Fight crime in Berlin? Deliver groceries in Wuhan? There's an algorithm for that. As the technology evolves, many look to AI to remedy social ills that people have shown themselves to be unable to address. This, they hope, will lead humanity

out of our current irrational darkness and ensure the victory of reason. Supporters laud AI for its potential to take away decision-making from humans and place it in the computer.

Equally vocal are the emotionalists, who argue the opposite: that humanity has been suffering from too much rationality, too much reliance on data, and cold, merciless, analytical reasoning. They believe that humanity's core problem is not too much passion but too little; that we are suffering because we do not rely enough on our gut and our instincts. They long for collective bonding in communities of similar-minded folks, and for clear distinctions and boundaries with "others" who do not belong. The appeal to emotional roots is a call to accept irrationality as a core quality of humanness.

We see this phenomenon on both the right and the left, in industrialized democracies as well as developing countries. It is associated with right-wing populists, who prefer decisive action over long-standing processes to weigh evidence. Governance is a feeling; leadership is an emotion; decisions are made from an innate sense of the rightness of one's beliefs. Yet it also crops up in the social sphere on the left, when activists wish to silence critics of their worldview or delegitimize others with whom they disagree.

Though the increasing power of AI is new, we have been dealing with this basic struggle for centuries. The tension between reason and emotion, artifice and nature, deliberation and gut instinct has shaped how we order our lives and govern our societies. In the 1600s the French philosopher and mathematician René Descartes advocated for a life of rationality, order, and evidence. Parisian parks that are laid out in perfect symmetry remind us of his influence.

A century later Jean-Jacques Rousseau urged a different approach, one of trust in feelings and intuition, and of looking to the inner self for answers: "All the evil I ever did in my life was the result of reflection," he wrote. "And the little good I have been able to do

was the result of impulse." It is a world of inklings, passions, and appetites—wild fits of fury perfectly excusable as expressions of one's humanity. When British and American landscapers make urban parks resemble rambling nature, it is an unwitting nod to Rousseau's outlook.

The same dichotomy played out in twentieth-century business. Frederick Taylor's influential theory of scientific management aimed to quantify every aspect of a company's operations. Managers armed with stopwatches and clipboards paraded the factory floor to ensure productivity. Yet the century ended with the celebrated success of GE's voluble chief executive Jack Welch, whose business autobiography was aptly subtitled *Straight from the Gut*.

There is a certain feeling of authenticity in the rejection of linear, fact-filled rationality for the joyous, human embrace of emotion in decision-making. Not everything can be reduced to a number or a formula in logic. But this ethos cannot solve problems—it can only glorify them. It can tear down but not construct. Over the past half century, psychologists and behavioral economists have amassed a mountain of experimental evidence showing the inferiority, in most cases, of decisions driven by the gut. Relying on instinct may give us a warm feeling of doing what feels right. But it fails to deliver a viable strategy to solve the challenges we face.

Meanwhile AI may make better decisions than people and steal our jobs, but computers and algorithms cannot frame. AI is brilliant at answering what it is asked; framers pose questions never before voiced. Computers work only in a world that exists; humans live in ones they imagine through framing.

Consider the computer's shortcomings in the very arena where it is usually feted for its excellence: board games. Even people who are familiar with this story extract the wrong lesson.

In 2018 Google DeepMind unveiled a system called AlphaZero that learned to win at chess, Go, and shogi purely by playing against

itself, with zero human input other than the rules. After just nine hours, during which it played itself in forty-four million games of chess, it was beating the world's best chess program, Stockfish. When grandmasters played against it, they were amazed by its alien approach. For over a century chess experts had a settled consensus on basic concepts and strategies, such as the value of pieces or board position. AlphaZero made radical moves, privileging mobility over position, and feeling no reluctance to make sacrifices. AlphaZero seemed to have conceived of an entirely new strategy for the game.

Only, it hadn't.

An AI system cannot conceive of anything. It cannot concoct mental models. It can neither generalize nor explain. AlphaZero is a black box—for us and for itself. It was people, not AI, who could look at the moves and develop the concepts of "board position" or "sacrifice." Humans frame AlphaZero's actions, making them explainable and applying them generally. Humans become smarter because we can abstract AI's accomplishments. Appreciating and applying the lessons is something that AI cannot do on its own.

Both the rationalists and the emotionalists correctly identify something unique about human cognition. But both lead to dead ends. Neither offers a suitable answer to our civilizational challenges. Nor can we expect much from a synthesis of the two. Amalgamating two approaches built on unsound foundations, at best, will simply maintain a fragile tension without hope of real progress.

The crucial insight is that our choice isn't limited to these two options. We do not have to decide between a dehumanizing singularity or a tsunami of populist terror—nor try to meld them into a suboptimal mix. We have at our disposal another strategy, a different human capacity that until now has been overlooked: framing. Our ability to apply, hone, and reinvent mental models provides us with the means to solve our problems without deferring to the machine or accepting the mob.

That brings us back to Regina Barzilay. We find ourselves at a crossroads. Huge challenges loom. As with antibiotics, many of our vulnerabilities are homegrown, the consequences of decisions we made, alternatives we failed to identify, actions we did not take. We got ourselves into our current troubles. The good news is that we can get ourselves out. But it requires a new mind-set.

What Is Without Starts from Within

There is a project called "Our World in Data" run by a team at the University of Oxford. It takes all aspects of life and, as its name suggests, views them through the lens of information. Infant mortality stats? They're on it. Global GDP? They've got you covered. The material is beloved by Bill Gates, who occasionally retweets its charts and whose foundation supports its work. And to judge from the rainbow-hued lines and bar charts they pump out, we've never had it so good.

It is true that by almost every metric, the world has been inexorably improving. There are fewer wars, less disease, more literacy, cleaner water, richer countries, happier people, longer life spans. Covid-19 will make a dent in some of these trend lines, but only a temporary one—raise our eyes to the more distant future and the dip will surely be smoothed out as time and progress march on.

The evolution of human thinking has played a crucial role in all these improvements. Before there is a change on the ground, there is a transformation in the mind. All that is without starts from within. We frame and reframe our world, and civilization advances.

But this sunny optimism may be misplaced. Its purveyors extrapolate to how things will evolve into the future. But their analyses disguise troubles. There is a pathology of human progress, that the very fruits of our creation risk being the sources of our destruction.

2

framing

mental models infuse everything we do, even if we're unaware of them

On Sunday, October 15, 2017, Alyssa Milano was sitting in bed at home in LA reading the news. The internet was abuzz with revelations about the movie producer Harvey Weinstein. As a child star of 1980s sitcoms now making it as a fortysomething actor, she knew all the names in the news. Sexual harassment is hardly unheard of in Hollywood; the proverbial "casting couch" is an ugly institution. But this felt different. These were not unwanted advances but physical assaults, dozens of them, which were not merely ignored but actively covered up, stretching back decades.

A message popped up on her phone from a friend suggesting that if women spoke out on Twitter, the world would see the extent of the problem. Milano liked the idea. She had always had a deeply moral streak. As a celebrity at fifteen, she kissed a boy who was HIV positive on a television talk show, to make the point that casual contact with AIDS victims is safe. In 2013 she "leaked" online what seemed to be a sex video of herself and her husband, which, rather than show steamy stuff, focused the camera on a two-minute news segment

explaining the conflict in Syria—clickbait in the service of humanitarianism.

Turning to Twitter made sense. "This is an amazing way to get some idea of the magnitude," Milano recalls thinking, and "a way to get the focus off these horrible men and to put the focus back on the victims and survivors." Milano herself had been assaulted on a film set almost twenty-five years earlier—but she had never mentioned it publicly. She opened a message window on Twitter and typed: "If you've been sexually harassed or assaulted write 'me too' as a reply to this tweet." She turned off the device, looked at her sleeping three-year-old daughter, and went to bed.

When she awoke, she was stunned to see the tweet had thirty-five thousand replies and growing. It spread across the world in a flash. By the end of the day, the #MeToo hashtag was used in more than twelve million posts. Reporters started calling. It became a global phenomenon.

The MeToo movement is many things, but perhaps most powerfully, it is a frame. It transformed how sexual assault is perceived, not as something to be kept private but something that could be made public. The declarations on Twitter became a source of empowerment and liberation. MeToo reversed the stigma: women need not be ashamed, and could bring shame upon the men who assaulted them.

Before MeToo, a woman speaking about an assault might be seen not as a victim but as complacent, complicit, or culpable. (*Why did you go to his apartment? Why did you wear that provocative dress?*) With this new frame, women could bear witness knowing they had strength in numbers, with a ready, global support group.

The new frame didn't merely provide an alternative way of thinking about the issue: it opened up new possibilities for decisions and actions.

Mapping the World

Whether it is the way women respond to sexual harassment or how scientists conceive of the molecular structures in antibiotics, frames make the complexity of the world intelligible. Our minds are filled with them. That's the way we think. Frames can be simple or sophisticated, accurate or imprecise, beautiful or evil. But they all capture some aspect of reality. In so doing, they help us to explain, focus, and decide.

Democracy is a frame, as is monarchism. In business, lean manufacturing is a frame and so is OKR (objectives and key results, the management system popularized by Intel and later Google). Religion is a frame, as is secular humanism (that is, morality without a god). The rule of law is a frame, as is the notion that might makes right. Racial equality is a frame, as is racism.

Frames are as foundational to our reasoning as they are versatile. In recent decades researchers in disciplines as broad as philosophy and neuroscience have studied human framing (though the terms they use to describe it vary, including *templates*, *abstractions*, *representations*, and *schemas*).

Today the idea that humans think by way of mental models is widely accepted in the hard sciences and the social sciences. Yet it is a comparatively recent phenomenon. In the early twentieth century, pondering human thinking was largely left to philosophers. Sigmund Freud and his interest in the mind's mysterious ways was an exception, not the rule. Between the wars, philosophers such as Ernst Cassirer and Ludwig Wittgenstein understood the mind to be based on symbols and the words it manipulates. It was a step toward a more rational view of cognition, but it was all theory, no empirics.

After the Second World War, the empirical sciences turned to the

human mind. Research shifted from philosophers to psychologists, especially as the latter began pondering cognitive processes inside the brain. Initially, they likened these to strict logical operations, but empirical studies failed to back up that view. Around the 1970s, the idea of "mental models" gained traction—along with the concept that human reasoning isn't an operation of formal logic but works more like a *simulation of reality*: we assess options for action by imagining what might happen.

This view has now been empirically confirmed in numerous experiments by psychologists and cognitive scientists. Neuroscience has chipped in as well in recent years, thanks to functional MRI scanners that can visualize brain activity of test subjects in real time. Studies have shown, for example, that when we plan for the future we activate brain areas that are associated with spatial cognition and our ability to think in three dimensions. We quite literally engage in a kind of deliberate and purposeful dreaming.

This work has led to a quiet transformation in understanding how people think. It points to mental models as the fundamental building blocks of human cognition. What we see and know, feel and believe, starts with the way we look at the universe. We can understand the world in relation to how we believe it works: why things happen, how they may unfold in the future, and what will happen if we act. Frames are not "imagination" or "creativity" but they enable it.

Most people probably haven't thought much about their mental state when they make decisions. This is because most decisions we make are of low consequence: what shirt to wear, what toppings to put on salad, and so on. But when people make more substantial decisions, their work has been transformed by the concept of mental models. Many of them take great strides in being aware of and deliberate about the frames they hold.

intent. Once we pick a frame and apply it, this opens up choices. Without applying a frame, we might endlessly debate but never act. Choosing and applying a frame—becoming "framers"—lays the groundwork for decision and action.

Americans are familiar with the term *framers* from their history lessons. It describes the men (and back then, only men) who drafted the Constitution. They were called framers at the time, because they were designing the "frame of government." The word is well chosen because the US Constitution is a frame that defines and demarcates the institutions and processes of government. It was the result of intense debates between two key camps over different models of government that lasted for months during the summer of 1787.

The Federalists advocated a strong, centralized model of government with a powerful chief executive, firm rule of law, and limited states' rights. Their frame focused attention on what would be needed to build a strong national state that could rise to be a great power. It was "federal" in that power devolved down from the top, not a "confederation" in which power was derived from the component pieces. In contrast, the Anti-Federalists wanted a weak center, a decentralized form of governance, guarantees for individual rights, and more direct democracy. Their frame prioritized building strong local democracies, which could join forces to defend each other against external threats.

As with maps, neither of these two frames is inferior in itself; each has its pros and cons, and either might be appropriate in a particular set of circumstances. To this day, these two distinct mental models have remained central to debates about how to govern democratic republics. Well over two centuries later and across the Atlantic, European nations continue to use similar frames in their debate about whether to envision the European Union as a *demos* (a united

people with a strong center), or a *demoi* (a bundle of peoples desiring a more decentralized form of governance).

Multiple, competing frames can lead to useful debates and elicit a variety of diverse options. But when there is more than one possible frame that could apply (and this happens frequently), choosing the right one for the situation is difficult. It requires a careful understanding of the objectives and the context for which the frame is being applied. And much rides on it.

Misframing and Misfortune

The wrong framing can be catastrophic. We can appreciate the importance of choosing the right frame by looking at how experts and decision-makers handled two different pandemics.

When the Ebola virus broke out in West Africa in the spring of 2014, experts were called in to study it and staunch it. The two main organizations working on a response were the World Health Organization (WHO), a UN agency, and Médecins Sans Frontières (MSF), an international aid group. Experts at both organizations knew that their first weapon in the battle was information. But while they had the same data, they drew opposite conclusions. It wasn't that their analyses were flawed. It was that they used different frames to assess the situation, each based on a distinct view of the outbreak's context and future spread.

The WHO's model was based on a historical frame. Looking at the relatively low number of Ebola cases, they reasoned that the 2014 outbreak was very similar to previous ones in the region, all of which had been contained locally. WHO forecast a limited outbreak and advised against drastic international measures. In contrast MSF took a spatial view of the outbreaks. The virus had proliferated across different villages that were far apart from one another and

which dotted the borders of three countries. Because of this, MSF concluded that it must have scattered farther than the data indicated. The group pushed for immediate, draconian action.

A different way of conceptualizing the crisis—whether the outbreak was concentrated or dispersed—was at the heart of the tension. The plague risked blowing up into a global catastrophe. Hundreds of people had already died, but potentially hundreds of millions of lives were at stake. Initially, WHO won the argument and only local measures were adopted. But Ebola's fast spread validated MSF's alarmist view. A global panic ensued, dubbed "Fearbola." (Donald Trump, then a property developer turned reality-TV star, called President Barack Obama "psycho" for not canceling flights from West Africa—though direct flights didn't exist—and tweeted "KEEP THEM OUT OF HERE!") Only extraordinary actions from governments were able to control the situation, and in the end, the crisis ebbed.

Now, fast-forward to 2020. When the novel coronavirus blipped onto public health authorities' radar early that year, it wasn't clear what sort of disease the world was dealing with. Seven coronaviruses were then known to affect humans, with a wide range of infection rates and lethality. Some cause the common cold. Others, like SARS (in Asia in 2002 to 2004) and MERS (in the Middle East in 2012), proved to have harsher symptoms, longer incubation periods, and case fatality rates of 10 percent and 35 percent, respectively. Yet the world had endured coronavirus outbreaks before, and they were squelched, just as Ebola had been.

Perhaps for this reason, it wasn't clear to countries how seriously to take the discovery of SARS-CoV-2 and the illness it caused, Covid-19. China closed the city of Wuhan, an unprecedented step that seemed like something only an authoritarian regime could or would do. In Italy, cases mushroomed before they knew what had hit them. Lombardy's hospitals were so overrun that for a period,

weeping doctors were forced to give sedatives to the elderly so they could die in less pain, to save limited medical resources for younger sufferers.

All countries were working off the same data, as WHO and MSF had been in 2014. And as in the case of Ebola, the way countries initially framed Covid-19 affected the options they envisioned, the actions they took, and how they fared at the outset of the crisis. The responses of Britain and New Zealand in particular show how different frames lead to different outcomes.

New Zealand framed Covid as being like SARS, and took an "elimination" approach. Though they had not been hit by SARS, the country's officials regularly mingled with their counterparts from places in the region such as Taiwan and South Korea, which had suffered, and had developed robust disease-monitoring systems and policies. Hence, at the very start of the Covid-19 outbreak, health officials in New Zealand went into disaster mode. Prime Minister Jacinda Ardern decided it would be better to overreact than underreact. "We currently have 102 cases—but so did Italy once," she told the nation in March. The country went into lockdown, closed its borders, and committed to contact tracing every case.

Britain, meanwhile, framed Covid as being more like the seasonal flu, and went for a "mitigation" strategy. Health officials assumed that the virus would inevitably spread through the population, eventually creating herd immunity. The government gave up on testing and contact tracing early in the crisis, and was later than its European counterparts to take actions such as banning large public events and closing schools. Officials opted for a national lockdown only after epidemiological models showed that the virus would swamp the National Health Service. In early June, Prime Minister Ardern declared her country Covid-free—while Britain