Part 2

WE JUST CAN'T PREDICT

When I ask people to name three recently implemented technologies that most impact our world today, they usually propose the computer, the Internet, and the laser. All three were unplanned, unpredicted, and unappreciated upon their discovery, and remained unappreciated well after their initial use. They were consequential. They were Black Swans. Of course, we have this retrospective illusion of their partaking in some master plan. You can create your own lists with similar results, whether you use political events, wars, or intellectual epidemics.

You would expect our record of prediction to be horrible: the world is far, far more complicated than we think, which is not a problem, except when most of us don't know it. We tend to "tunnel" while looking into the future, making it business as usual, Black Swan–free, when in fact there is nothing usual about the future. It is not a Platonic category!

We have seen how good we are at narrating backward, at inventing stories that convince us that we understand the past. For many people, knowledge has the remarkable power of producing confidence instead of measurable aptitude. Another problem: the focus on the (inconsequential) regular, the Platonification that makes the forecasting "inside the box."

I find it scandalous that in spite of the empirical record we continue to project into the future as if we were good at it, using tools and methods that exclude rare events. Prediction is firmly institutionalized in our world. We are suckers for those who help us navigate uncertainty, whether the fortune-teller or the "well-published" (dull) academics or civil servants using phony mathematics.

From Yogi Berra to Henri Poincaré

The great baseball coach Yogi Berra has a saying, "It is tough to make predictions, especially about the future." While he did not produce the writings that would allow him to be considered a philosopher, in spite of his wisdom and intellectual abilities, Berra can claim to know something about randomness. He was a practitioner of uncertainty, and, as a baseball player and coach, regularly faced random outcomes, and had to face their results deep into his bones.

In fact, Yogi Berra is not the only thinker who thought about how much of the future lies beyond our abilities. Many less popular, less pithy, but not less competent thinkers than he have examined our inherent limitations in this regard, from the philosophers Jacques Hadamard and Henri Poincaré (commonly described as mathematicians), to the philosopher Friedrich von Hayek (commonly described, alas, as an economist), to the philosopher Karl Popper (commonly known as a philosopher). We can safely call this the Berra-Hadamard-Poincaré-Hayek-Popper conjecture, which puts structural, built-in limits to the enterprise of predicting.

"The future ain't what it used to be," Berra later said.* He seems to have been right: the gains in our ability to model (and predict) the world may be dwarfed by the increases in its complexity—implying a greater and greater role for the unpredicted. The larger the role of the Black Swan, the harder it will be for us to predict. Sorry.

Before going into the limits of prediction, we will discuss our track record in forecasting and the relation between gains in knowledge and the offsetting gains in confidence.

* Note that these sayings attributed to Yogi Berra might be apocryphal—it was the physicist Niels Bohr who came up with the first one, and plenty of others came up with the second. These sayings remain, however, quintessential Berraisms.

Chapter Ten

THE SCANDAL OF PREDICTION

Welcome to Sydney—How many lovers did she have?—How to be an economist, wear a nice suit, and make friends—Not right, just "almost" right—Shallow rivers can have deep spots

One March evening, a few men and women were standing on the esplanade overlooking the bay outside the Sydney Opera House. It was close to the end of the summer in Sydney, but the men were wearing jackets despite the warm weather. The women were more thermally comfortable than the men, but they had to suffer the impaired mobility of high heels.

They all had come to pay the price of sophistication. Soon they would listen for several hours to a collection of oversize men and women singing endlessly in Russian. Many of the opera-bound people looked like they worked for the local office of J. P. Morgan, or some other financial institution where employees experience differential wealth from the rest of the local population, with concomitant pressures on them to live by a sophisticated script (wine and opera). But I was not there to take a peek at the neosophisticates. I had come to look at the Sydney Opera House, a building that adorns every Australian tourist brochure. Indeed, it is striking, though it looks like the sort of building architects create in order to impress other architects.

That evening walk in the very pleasant part of Sydney called the Rocks was a pilgrimage. While Australians were under the illusion that they had built a monument to distinguish their skyline, what they had really done was to construct a monument to our failure to predict, to plan, and to come to grips with our *unknowledge* of the future—our systematic underestimation of what the future has in store.

The Australians had actually built a symbol of the epistemic arrogance of the human race. The story is as follows. The Sydney Opera House was supposed to open in early 1963 at a cost of AU$ 7 million. It finally opened its doors more than ten years later, and, although it was a less ambitious version than initially envisioned, it ended up costing around AU$ 104 million. While there are far worse cases of planning failures (namely the Soviet Union), or failures to forecast (all important historical events), the Sydney Opera House provides an aesthetic (at least in principle) illustration of the difficulties. This opera-house story is the mildest of all the distortions we will discuss in this section (it was only money, and it did not cause the spilling of innocent blood). But it is nevertheless emblematic.

This chapter has two topics. First, we are demonstrably arrogant about what we think we know. We certainly know a lot, but we have a built-in tendency to think that we know a little bit more than we actually do, enough of *that little bit* to occasionally get into serious trouble. We shall see how you can verify, even measure, such arrogance in your own living room.

Second, we will look at the implications of this arrogance for all the activities involving prediction.

Why on earth do we predict so much? Worse, even, and more interesting: Why don't we talk about our record in predicting? Why don't we see how we (almost) always miss the big events? I call this the scandal of prediction.

ON THE VAGUENESS OF CATHERINE'S LOVER COUNT

Let us examine what I call *epistemic arrogance,* literally, our hubris concerning the limits of our knowledge. *Epistēmē* is a Greek word that refers to knowledge; giving a Greek name to an abstract concept makes it sound important. True, our knowledge does grow, but it is threatened by greater increases in confidence, which make our increase in knowledge at the same time an increase in confusion, ignorance, and conceit.

Take a room full of people. Randomly pick a number. The number could correspond to anything: the proportion of psychopathic stockbrokers in western Ukraine, the sales of this book during the months with *r* in them, the average IQ of business-book editors (or business writers), the number of lovers of Catherine II of Russia, et cetera. Ask each person in the room to independently estimate a range of possible values for that number set in such a way that they believe that they have a 98 percent chance of being right, and less than 2 percent chance of being wrong. In other words, whatever they are guessing has about a 2 percent chance to fall outside their range. For example:

"I am 98 percent confident that the population of Rajastan is between 15 and 23 million."

"I am 98 percent confident that Catherine II of Russia had between 34 and 63 lovers."

You can make inferences about human nature by counting how many people in your sample guessed wrong; it is not expected to be too much higher than two out of a hundred participants. Note that the subjects (your victims) are free to set their range as wide as they want: you are not trying to gauge their knowledge but rather *their evaluation of their own knowledge.*

Now, the results. Like many things in life, the discovery was unplanned, serendipitous, surprising, and took a while to digest. Legend has it that Albert and Raiffa, the researchers who noticed it, were actually looking for something quite different, and more boring: how humans figure out probabilities in their decision making when uncertainty is involved (what the learned call *calibrating*). The researchers came out befuddled. The 2 percent error rate turned out to be close to 45 percent in the population being tested! It is quite telling that the first sample consisted of Harvard Business School students, a breed not particularly renowned for their humility or introspective orientation. MBAs are particularly nasty in this regard, which might explain their business success. Later studies document more humility, or rather a smaller degree of arrogance, in other populations. Janitors and cabdrivers are rather humble. Politicians and corporate executives, alas . . . I'll leave them for later.

Are we twenty-two times too comfortable with what we know? It seems so.

This experiment has been replicated dozens of times, across populations, professions, and cultures, and just about every empirical psychologist and decision theorist has tried it on his class to show his students the big problem of humankind: we are simply not wise enough to be trusted with knowledge. The intended 2 percent error rate usually turns out to be between 15 percent and 30 percent, depending on the population and the subject matter.

I have tested myself and, sure enough, failed, even while consciously trying to be humble by carefully setting a wide range—and yet such underestimation happens to be, as we will see, the core of my professional activities. This bias seems present in all cultures, even those that favor humility—there may be no consequential difference between downtown Kuala Lumpur and the ancient settlement of Amioun, (currently) Lebanon. Yesterday afternoon, I gave a workshop in London, and had been mentally writing on my way to the venue because the cabdriver had an above-average ability to "find traffic." I decided to make a quick experiment during my talk.

I asked the participants to take a stab at a range for the number of books in Umberto Eco's library, which, as we know from the introduction to Part One, contains 30,000 volumes. Of the sixty attendees, not a single one made the range wide enough to include the actual number (the 2 percent error rate became 100 percent). This case may be an aberration, but the distortion is exacerbated with quantities that are out of the ordinary. Interestingly, the crowd erred on the very high and the very low sides: some set their ranges at 2,000 to 4,000; others at 300,000 to 600,000.

True, someone warned about the nature of the test can play it safe and set the range between zero and infinity; but this would no longer be "calibrating"—that person would not be conveying any information, and could not produce an informed decision in such a manner. In this case it is more honorable to just say, "I don't want to play the game; I have no clue."

It is not uncommon to find counterexamples, people who overshoot in the opposite direction and actually overestimate their error rate: you may have a cousin particularly careful in what he says, or you may remember that college biology professor who exhibited pathological humility; the tendency that I am discussing here applies to the average of the population, not to every single individual. There are sufficient variations around the average to warrant occasional counterexamples. Such people are in the minority—and, sadly, since they do not easily achieve prominence, they do not seem to play too influential a role in society.

Epistemic arrogance bears a double effect: we overestimate what we know, and underestimate uncertainty, by compressing the range of possible uncertain states (i.e., by reducing the space of the unknown).

The applications of this distortion extend beyond the mere pursuit of knowledge: just look into the lives of the people around you. Literally any decision pertaining to the future is likely to be infected by it. Our human race is affected by a chronic underestimation of the possibility of the future straying from the course initially envisioned (in addition to other biases that sometimes exert a compounding effect). To take an obvious example, think about how many people divorce. Almost all of them are acquainted with the statistic that between one-third and one-half of all marriages fail, something the parties involved did not forecast while tying the knot. Of course, "not us," because "we get along so well" (as if others tying the knot got along poorly).

I remind the reader that I am not testing how much people know, but assessing *the difference between what people actually know and how much they think they know.* I am reminded of a measure my mother concocted, as a joke, when I decided to become a businessman. Being ironic about my (perceived) confidence, though not necessarily unconvinced of my abilities, she found a way for me to make a killing. How? Someone who could figure out how to buy me at the price I am truly worth and sell me at what I think I am worth would be able to pocket a huge difference. Though I keep trying to convince her of my internal humility and insecurity concealed under a confident exterior; though I keep telling her that I am an introspector—she remains skeptical. Introspector shmintrospector, she still jokes at the time of this writing that I am a little ahead of myself.

BLACK SWAN BLINDNESS REDUX

The simple test above suggests the presence of an ingrained tendency in humans to underestimate outliers—or Black Swans. Left to our own devices, we tend to think that what happens every decade in fact only happens once every century, and, furthermore, that we know what's going on.

This miscalculation problem is a little more subtle. In truth, outliers are not as sensitive to underestimation since they are fragile to estimation errors, which can go in both directions. As we saw in Chapter 6, there are conditions under which people overestimate the unusual or some specific unusual event (say when sensational images come to their minds)—which, we have seen, is how insurance companies thrive. So my general point is that these events are very fragile to *miscalculation,* with a general severe underestimation mixed with an occasional severe overestimation.

The errors get worse with the degree of remoteness to the event. So far, we have only considered a 2 percent error rate in the game we saw earlier, but if you look at, say, situations where the odds are one in a hundred, one in a thousand, or one in a million, then the errors become monstrous. The longer the odds, the larger the epistemic arrogance.

Note here one particularity of our intuitive judgment: even if we lived in Mediocristan, in which large events are rare, we would still underestimate extremes—we would think that they are even rarer. We underestimate our error rate even with Gaussian variables. Our intuitions are sub-Mediocristani. But we do not live in Mediocristan. The numbers we are likely to estimate on a daily basis belong largely in Extremistan, i.e., they are run by concentration and subjected to Black Swans.

Guessing and Predicting

There is no effective difference between my guessing a variable that is not random, but for which my information is partial or deficient, such as the number of lovers who transited through the bed of Catherine II of Russia, and predicting a random one, like tomorrow's unemployment rate or next year's stock market. In this sense, guessing (what I don't know, but what someone else may know) and predicting (what has not taken place yet) are the same thing.

To further appreciate the connection between guessing and predicting, assume that instead of trying to gauge the number of lovers of Catherine of Russia, you are estimating the less interesting but, for some, more important question of the population growth for the next century, the stock-market returns, the social-security deficit, the price of oil, the results of your great-uncle's estate sale, or the environmental conditions of Brazil two decades from now. Or, if you are the publisher of Yevgenia Krasnova's book, you may need to produce an estimate of the possible future sales. We are now getting into dangerous waters: just consider that most professionals who make forecasts are also afflicted with the mental impediment discussed above. Furthermore, people who make forecasts professionally are often *more* affected by such impediments than those who don't.

INFORMATION IS BAD FOR KNOWLEDGE

You may wonder how learning, education, and experience affect epistemic arrogance—how educated people might score on the above test, as compared with the rest of the population (using Mikhail the cabdriver as a benchmark). You will be surprised by the answer: it depends on the profession. I will first look at the advantages of the "informed" over the rest of us in the humbling business of prediction.

I recall visiting a friend at a New York investment bank and seeing a frenetic hotshot "master of the universe" type walking around with a set of wireless headphones wrapped around his ears and a microphone jutting out of the right side that prevented me from focusing on his lips during my twenty-second conversation with him. I asked my friend the purpose of that contraption. "He likes to keep in touch with London," I was told. When you are employed, hence dependent on other people's judgment, looking busy can help you claim responsibility for the results in a random environment. The appearance of busyness reinforces the perception of causality, of the link between results and one's role in them. This of course applies even more to the CEOs of large companies who need to trumpet a link between their "presence" and "leadership" and the results of the company. I am not aware of any studies that probe the usefulness of their time being invested in conversations and the absorption of small-time information—nor have too many writers had the guts to question how large the CEO's role is in a corporation's success.

Let us discuss one main effect of information: impediment to knowledge.

Aristotle Onassis, perhaps the first mediatized tycoon, was principally famous for being rich—and for exhibiting it. An ethnic Greek refugee from southern Turkey, he went to Argentina, made a lump of cash by importing Turkish tobacco, then became a shipping magnate. He was reviled when he married Jacqueline Kennedy, the widow of the American president John F. Kennedy, which drove the heartbroken opera singer Maria Callas to immure herself in a Paris apartment to await death.

If you study Onassis's life, which I spent part of my early adulthood doing, you would notice an interesting regularity: "work," in the conventional sense, was not his thing. He did not even bother to have a desk, let alone an office. He was not just a dealmaker, which does not necessitate having an office, but he also ran a shipping empire, which requires day-to-day monitoring. Yet his main tool was a notebook, which contained all the information he needed. Onassis spent his life trying to socialize with the rich and famous, and to pursue (and collect) women. He generally woke up at noon. If he needed legal advice, he would summon his lawyers to some nightclub in Paris at two A.M. He was said to have an irresistible charm, which helped him take advantage of people.

Let us go beyond the anecdote. There may be a "fooled by randomness" effect here, of making a causal link between Onassis's success and his modus operandi. I may never know if Onassis was skilled or lucky, though I am convinced that his charm opened doors for him, but I can subject his modus to a rigorous examination by looking at empirical research on the link between information and understanding. So this statement, *additional knowledge of the minutiae of daily business can be useless, even actually toxic,* is indirectly but quite effectively testable.

Show two groups of people a blurry image of a fire hydrant, blurry enough for them not to recognize what it is. For one group, increase the resolution slowly, in ten steps. For the second, do it faster, in five steps. Stop at a point where both groups have been presented an identical image and ask each of them to identify what they see. The members of the group that saw fewer intermediate steps are likely to recognize the hydrant much faster. Moral? The more information you give someone, the more hypotheses they will formulate along the way, and the worse off they will be. They see more random noise and mistake it for information.

The problem is that our ideas are sticky: once we produce a theory, we are not likely to change our minds—so those who delay developing their theories are better off. When you develop your opinions on the basis of weak evidence, you will have difficulty interpreting subsequent information that contradicts these opinions, even if this new information is obviously more accurate. Two mechanisms are at play here: the confirmation bias that we saw in Chapter 5, and belief perseverance, the tendency not to reverse opinions you already have. Remember that we treat ideas like possessions, and it will be hard for us to part with them.

The fire hydrant experiment was first done in the sixties, and replicated several times since. I have also studied this effect using the mathematics of information: the more detailed knowledge one gets of empirical reality, the more one will see the noise (i.e., the anecdote) and mistake it for actual information. Remember that we are swayed by the sensational. Listening to the news on the radio every hour is far worse for you than reading a weekly magazine, because the longer interval allows information to be filtered a bit.

In 1965, Stuart Oskamp supplied clinical psychologists with successive files, each containing an increasing amount of information about patients; the psychologists' diagnostic abilities did not grow with the additional supply of information. They just got more confident in their original diagnosis. Granted, one may not expect too much of psychologists of the 1965 variety, but these findings seem to hold across disciplines.

Finally, in another telling experiment, the psychologist Paul Slovic asked bookmakers to select from eighty-eight variables in past horse races those that they found useful in computing the odds. These variables included all manner of statistical information about past performances. The bookmakers were given the ten most useful variables, then asked to predict the outcome of races. Then they were given ten more and asked to predict again. The increase in the information set did not lead to an increase in their accuracy; their confidence in their choices, on the other hand, went up markedly. Information proved to be toxic. I've struggled much of my life with the common middlebrow belief that "more is better"—more is sometimes, but not always, better. This toxicity of knowledge will show in our investigation of the so-called expert.

THE EXPERT PROBLEM, OR THE TRAGEDY OF THE EMPTY SUIT

So far we have not questioned the authority of the professionals involved but rather their ability to gauge the boundaries of their own knowledge. Epistemic arrogance does not preclude skills. A plumber will almost always know more about plumbing than a stubborn essayist and mathematical trader. A hernia surgeon will rarely know less about hernias than a belly dancer. But their probabilities, on the other hand, will be off—and, this is the disturbing point, you may know much more on that score than the expert. No matter what anyone tells you, it is a good idea to question *the error rate* of an expert's procedure. Do not question his procedure, only his confidence. (As someone who was burned by the medical establishment, I learned to be cautious, and I urge everyone to be: if you walk into a doctor's office with a symptom, do not listen to his odds of its *not* being cancer.)

I will separate the two cases as follows. The mild case: *arrogance in the presence of (some) competence,* and the severe case: *arrogance mixed with incompetence (the empty suit).* There are some professions in which you know more than the experts, who are, alas, people for whose opinions you are paying—instead of them paying you to listen to them. Which ones?

What Moves and What Does Not Move

There is a very rich literature on the so-called expert problem, running empirical testing on experts to verify their record. But it seems to be confusing at first. On one hand, we are shown by a class of expert-busting researchers such as Paul Meehl and Robyn Dawes that the "expert" is the closest thing to a fraud, performing no better than a computer using a single metric, their intuition getting in the way and blinding them. (As an example of a computer using a single metric, the ratio of liquid assets to debt fares better than the majority of credit analysts.) On the other hand, there is abundant literature showing that many people can beat computers thanks to their intuition. Which one is correct?

There must be some disciplines with true experts. Let us ask the following questions: Would you rather have your upcoming brain surgery performed by a newspaper's science reporter or by a certified brain surgeon? On the other hand, would you prefer to listen to an economic forecast by someone with a PhD in finance from some "prominent" institution such as the Wharton School, or by a newspaper's business writer? While the answer to the first question is empirically obvious, the answer to the second one isn't at all. We can already see the difference between "know-how" and "know-what." The Greeks made a distinction between *technē* and *epistēmē*. The empirical school of medicine of Menodotus of Nicomedia and Heraclites of Tarentum wanted its practitioners to stay closest to *technē* (i.e., "craft"), and away from *epistēmē* (i.e., "knowledge," "science").

The psychologist James Shanteau undertook the task of finding out which disciplines have experts and which have none. Note the confirmation problem here: if you want to prove that there are no experts, then you will be able to find *a* profession in which experts are useless. And you can prove the opposite just as well. But there is a regularity: there are professions where experts play a role, and others where there is no evidence of skills. Which are which?

Experts who tend to be experts: livestock judges, astronomers, test pilots, soil judges, chess masters, physicists, mathematicians (when they deal with mathematical problems, not empirical ones), accountants, grain inspectors, photo interpreters, insurance analysts (dealing with bell curve-style statistics).

Experts who tend to be . . . not experts: stockbrokers, clinical psychologists, psychiatrists, college admissions officers, court judges, councilors, personnel selectors, intelligence analysts (the CIA's record, in spite of its costs, is pitiful). I would add these results from my own examination of the literature: economists, financial forecasters, finance professors, political scientists, "risk experts," Bank for International Settlements staff, august members of the International Association of Financial Engineers, and personal financial advisers.

Simply, *things that move,* and therefore require knowledge, do not usually have experts, while things that don't move seem to have some experts. In other words, professions that deal with the future and base their studies on the nonrepeatable past have an expert problem (with the exception of the weather and businesses involving short-term physical processes, not socioeconomic ones). I am not saying that no one who deals with the future provides any valuable information (as I pointed out earlier, newspapers can predict theater opening hours rather well), but rather that those who provide no tangible added value are generally dealing with the future.

Another way to see it is that things that move are often Black Swan–prone. Experts are narrowly focused persons who need to "tunnel." In situations where tunneling is safe, because Black Swans are not consequential, the expert will do well.

Robert Trivers, an evolutionary psychologist and a man of supernormal insights, has another answer (he became one of the most influential evolutionary thinkers since Darwin with ideas he developed while trying to go to law school). He links it to self-deception. In fields where we have ancestral traditions, such as pillaging, we are very good at predicting outcomes by gauging the balance of power. Humans and chimps can immediately sense which side has the upper hand, and make a cost-benefit analysis about whether to attack and take the goods and the mates. Once you start raiding, you put yourself into a delusional mind-set that makes you ignore additional information—it is best to avoid wavering during battle. On the other hand, unlike raids, large-scale wars are not something present in human heritage—we are new to them—so we tend to misestimate their duration and overestimate our relative power. Recall the underestimation of the duration of the Lebanese war. Those who fought in the Great War thought it would be a mere cakewalk. So it was with the Vietnam conflict, so it is with the Iraq war, and just about every modern conflict.

You cannot ignore self-delusion. The problem with experts is that they do not know what they do not know. Lack of knowledge and delusion about the quality of your knowledge come together—the same process that makes you know less also makes you satisfied with your knowledge.

Next, instead of the range of forecasts, we will concern ourselves with the accuracy of forecasts, i.e., the ability to predict the number itself.

How to Have the Last Laugh

We can also learn about prediction errors from trading activities. We quants have ample data about economic and financial forecasts—from general data about large economic variables to the forecasts and market calls of the television "experts" or "authorities." The abundance of such data and the ability to process it on a computer make the subject invaluable for an empiricist. If I had been a journalist, or, God forbid, a historian, I would have had a far more difficult time testing the predictive effectiveness of these verbal discussions. You cannot process verbal commentaries with a computer—at least not so easily. Furthermore, many economists naïvely make the mistake of producing a lot of forecasts concerning many variables, giving us a database of economists and variables, which enables us to see whether some economists are better than others (there is no consequential difference) or if there are certain variables for which they are more competent (alas, none that are meaningful).

I was in a seat to observe from very close our ability to predict. In my full-time trader days, a couple of times a week, at 8:30 A.M., my screen would flash some economic number released by the Department of Commerce, or Treasury, or Trade, or some such honorable institution. I never had a clue about what these numbers meant and never saw any need to invest energy in finding out. So I would not have cared the least about them except that people got all excited and talked quite a bit about what these figures were going to mean, pouring verbal sauce around the forecasts. Among such numbers you have the Consumer Price Index (CPI), Nonfarm Payrolls (changes in the number of employed individuals), the Index of Leading Economic Indicators, Sales of Durable Goods (dubbed "doable girls" by traders), the Gross Domestic Product (the most important one), and many more that generate different levels of excitement depending on their presence in the discourse.

The data vendors allow you to take a peek at forecasts by "leading economists," people (in suits) who work for the venerable institutions, such as J. P. Morgan Chase or Morgan Stanley. You can watch these economists talk, theorizing eloquently and convincingly. Most of them earn seven figures and they rank as stars, with teams of researchers crunching numbers and projections. But the stars are foolish enough to publish their projected numbers, right there, for posterity to observe and assess their degree of competence.

Worse yet, many financial institutions produce booklets every year-end called "Outlook for 200X," reading into the following year. Of course they do not check how their previous forecasts fared *after* they were formulated. The public might have been even more foolish in buying the arguments without requiring the following simple tests—easy though they are, very few of them have been done. One elementary empirical test is to compare these star economists to a hypothetical cabdriver (the equivalent of Mikhail from Chapter 1): you create a synthetic agent, someone who takes the most recent number as the best predictor of the next, while assuming that he does not know anything. Then all you have to do is compare the error rates of the hotshot economists and your synthetic agent. The problem is that when you are swayed by stories you forget about the necessity of such testing.

Events Are Outlandish

The problem with prediction is a little more subtle. It comes mainly from the fact that we are living in Extremistan, not Mediocristan. Our predictors may be good at predicting the ordinary, but not the irregular, and this is where they ultimately fail. All you need to do is miss one interest-rates move, from 6 percent to 1 percent in a longer-term projection (what happened between 2000 and 2001) to have all your subsequent forecasts rendered completely ineffectual in correcting your cumulative track record. What matters is not how often you are right, but how large your cumulative errors are.

And these cumulative errors depend largely on the big surprises, the big opportunities. Not only do economic, financial, and political predictors miss them, but they are quite ashamed to say anything outlandish to their clients—and yet *events, it turns out, are almost always outlandish*. Furthermore, as we will see in the next section, economic forecasters tend to fall closer to one another than to the resulting outcome. Nobody wants to be off the wall.

Since my testing has been informal, for commercial and entertainment purposes, for my own consumption and not formatted for publishing, I will use the more formal results of other researchers who did the dog work of dealing with the tedium of the publishing process. I am surprised that so little introspection has been done to check on the usefulness of these professions. There are a few—but not many—formal tests in three domains: security analysis, political science, and economics. We will no doubt have more in a few years. Or perhaps not—the authors of such papers might become stigmatized by his colleagues. Out of close to a million papers published in politics, finance, and economics, there have been only a small number of checks on the predictive quality of such knowledge.

Herding Like Cattle

A few researchers have examined the work and attitude of security analysts, with amazing results, particularly when one considers the epistemic arrogance of these operators. In a study comparing them with weather forecasters, Tadeusz Tyszka and Piotr Zielonka document that the analysts are worse at predicting, while having a greater faith in their own skills. Somehow, the analysts' self-evaluation did not decrease their error margin after their failures to forecast.

Last June I bemoaned the dearth of such published studies to Jean-Philippe Bouchaud, whom I was visiting in Paris. He is a boyish man who looks half my age though he is only slightly younger than I, a matter that I half jokingly attribute to the beauty of physics. Actually he is not exactly a physicist but one of those quantitative scientists who apply methods of statistical physics to economic variables, a field that was started by Benoît Mandelbrot in the late 1950s. This community does not use Mediocristan mathematics, so they seem to care about the truth. They are completely outside the economics and business-school finance establishment, and survive in physics and mathematics departments or, very often, in trading houses (traders rarely hire economists for their own consumption, but rather to provide stories for their less sophisticated clients). Some of them also operate in sociology with the same hostility on the part of the "natives." Unlike economists who wear suits and spin theories, they use empirical methods to observe the data and do not use the bell curve.

He surprised me with a research paper that a summer intern had just finished under his supervision and that had just been accepted for publication; it scrutinized two thousand predictions by security analysts. What it showed was that these brokerage-house analysts predicted *nothing*—a naïve forecast made by someone who takes the figures from one period as predictors of the next would not do markedly worse. Yet analysts are informed about companies' orders, forthcoming contracts, and planned expenditures, so this advanced knowledge *should* help them do considerably better than a naïve forecaster looking at the past data without further information. Worse yet, the forecasters' errors were significantly larger than the average difference between individual forecasts, which indicates herding. Normally, forecasts should be as far from one another as they are from the predicted number. But to understand how they manage to stay in business, and why they don't develop severe nervous breakdowns (with weight loss, erratic behavior, or acute alcoholism), we must look at the work of the psychologist Philip Tetlock.

I Was "Almost" Right

Tetlock studied the business of political and economic "experts." He asked various specialists to judge the likelihood of a number of political, economic, and military events occurring within a specified time frame (about five years ahead). The outcomes represented a total number of around twenty-seven thousand predictions, involving close to three hundred specialists. Economists represented about a quarter of his sample. The study revealed that experts' error rates were clearly many times what they had estimated. His study exposed an expert problem: there was no difference in results whether one had a PhD or an undergraduate degree. Well-published professors had no advantage over journalists. The only regularity Tetlock found was the negative effect of reputation on prediction: those who had a big reputation were worse predictors than those who had none.

But Tetlock's focus was not so much to show the real competence of experts (although the study was quite convincing with respect to that) as to investigate why the experts did not realize that they were not so good at their own business, in other words, how they spun their stories. There seemed to be a logic to such incompetence, mostly in the form of belief defense, or the protection of self-esteem. He therefore dug further into the mechanisms by which his subjects generated ex post explanations.

I will leave aside how one's ideological commitments influence one's perception and address the more general aspects of this blind spot toward one's own predictions.

You tell yourself that you were playing a different game. Let's say you failed to predict the weakening and precipitous fall of the Soviet Union (which no social scientist saw coming). It is easy to claim that you were excellent at understanding the political workings of the Soviet Union, but that these Russians, being exceedingly Russian, were skilled at hiding from you crucial economic elements. Had you been in possession of such economic intelligence, you would certainly have been able to predict the demise of the Soviet regime. It is not your skills that are to blame. The same might apply to you if you had forecast the landslide victory for Al Gore over George W. Bush. You were not aware that the economy was in such dire straits; indeed, this fact seemed to be concealed from everyone. Hey, you are not an economist, and the game turned out to be about economics.

You invoke the outlier. Something happened that was outside the system, outside the scope of your science. Given that it was not predictable, you are not to blame. It was a Black Swan and you are not supposed to predict Black Swans. Black Swans, NNT tells us, are fundamentally unpredictable (but then I think that NNT would ask you, Why rely on predictions?). Such events are "exogenous," coming from outside your science. Or maybe it was an event of very, very low probability, a thousand-year flood, and we were unlucky to be exposed to it. But next time, it will not happen. This focus on the narrow game and linking one's performance to a given script is how the nerds explain the failures of mathematical methods in society. The model was right, it worked well, but the game turned out to be a different one than anticipated.

The "almost right" defense. Retrospectively, with the benefit of a revision of values and an informational framework, it is easy to feel that it was a close call. Tetlock writes, "Observers of the former Soviet Union who, in 1988, thought the Communist Party could not be driven from power by 1993 or 1998 were especially likely to believe that Kremlin hardliners almost overthrew Gorbachev in the 1991 coup attempt, and they would have if the conspirators had been more resolute and less inebriated, or if key military officers had obeyed orders to kill civilians challenging martial law or if Yeltsin had not acted so bravely."

I will go now into more general defects uncovered by this example. These "experts" were lopsided: on the occasions when they were right, they attributed it to their own depth of understanding and expertise; when wrong, it was either the situation that was to blame, since it was unusual, or, worse, they did not recognize that they were wrong and spun stories around it. They found it difficult to accept that their grasp was a little short. But this attribute is universal to all our activities: there is something in us designed to protect our self-esteem.

We humans are the victims of an asymmetry in the perception of random events. We attribute our successes to our skills, and our failures to external events outside our control, namely to randomness. We feel responsible for the good stuff, but not for the bad. This causes us to think that we are better than others at whatever we do for a living. Ninety-four percent of Swedes believe that their driving skills put them in the top 50 percent of Swedish drivers; 84 percent of Frenchmen feel that their lovemaking abilities put them in the top half of French lovers.

The other effect of this asymmetry is that we feel a little unique, unlike others, for whom we do not perceive such an asymmetry. I have mentioned the unrealistic expectations about the future on the part of people in the process of tying the knot. Also consider the number of families who tunnel on their future, locking themselves into hard-to-flip real estate thinking they are going to live there permanently, not realizing that the general track record for sedentary living is dire. Don't they see those well-dressed real-estate agents driving around in fancy two-door German cars? We are very nomadic, far more than we plan to be, and forcibly so. Consider how many people who have abruptly lost their job deemed it likely to occur, even a few days before. Or consider how many drug addicts entered the game willing to stay in it so long.

There is another lesson from Tetlock's experiment. He found what I mentioned earlier, that many university stars, or "contributors to top journals," are no better than the average *New York Times* reader or journalist in detecting changes in the world around them. These sometimes overspecialized experts failed tests in their own specialties.

The hedgehog and the fox. Tetlock distinguishes between two types of predictors, the hedgehog and the fox, according to a distinction promoted by the essayist Isaiah Berlin. As in Aesop's fable, the hedgehog knows one thing, the fox knows many things—these are the adaptable types you need in daily life. Many of the prediction failures come from hedgehogs who are mentally married to a single big Black Swan event, a big bet that is not likely to play out. The hedgehog is someone focusing on a single, improbable, and consequential event, falling for the narrative fallacy that makes us so blinded by one single outcome that we cannot imagine others.

Hedgehogs, because of the narrative fallacy, are easier for us to understand—their ideas work in sound bites. Their category is overrepresented among famous people; ergo famous people are on average worse at forecasting than the rest of the predictors.

I have avoided the press for a long time because whenever journalists hear my Black Swan story, they ask me to give them a list of future impacting events. They want me to be *predictive* of these Black Swans. Strangely, my book *Fooled by Randomness*, published a week before September 11, 2001, had a discussion of the possibility of a plane crashing into my office building. So I was naturally asked to show "how I predicted the event." I didn't predict it—it was a chance occurrence. I am not playing oracle! I even recently got an e-mail asking me to list the next ten Black Swans. Most fail to get my point about the error of specificity, the narrative fallacy, and the idea of prediction. Contrary to what people might expect, I am not recommending that anyone become a hedgehog—rather, be a fox with an open mind. I know that history is going to be dominated by an improbable event, I just don't know what that event will be.

Reality? What For?

I found no formal, Tetlock-like comprehensive study in economics journals. But, suspiciously, I found no paper trumpeting economists' ability to produce reliable projections. So I reviewed what articles and working papers in economics I could find. They collectively show no convincing evidence that economists as a community have an ability to predict, and, if they have some ability, their predictions are at best just *slightly* better than random ones—not good enough to help with serious decisions.

The most interesting test of how academic methods fare in the real world was run by Spyros Makridakis, who spent part of his career managing competitions between forecasters who practice a "scientific method" called econometrics—an approach that combines economic theory with statistical measurements. Simply put, he made people forecast *in real life* and then he judged their accuracy. This led to the series of "M-Competitions" he ran, with assistance from Michele Hibon, of which M3 was the third and most recent one, completed in 1999. Makridakis and Hibon reached the sad conclusion that "statistically sophisticated or complex methods do not necessarily provide more accurate forecasts than simpler ones."

I had an identical experience in my quant days—the foreign scientist with the throaty accent spending his nights on a computer doing complicated mathematics rarely fares better than a cabdriver using the simplest methods within his reach. The problem is that we focus on the rare occasion when these methods work and almost never on their far more numerous failures. I kept begging anyone who would listen to me: "Hey, I am an uncomplicated, no-nonsense fellow from Amioun, Lebanon, and have trouble understanding why something is considered valuable if it requires running computers overnight but does not enable me to predict better than any other guy from Amioun." The only reactions I got from these colleagues were related to the geography and history of Amioun rather than a no-nonsense explanation of their business. Here again, you see the narrative fallacy at work, except that in place of journalistic stories you have the more dire situation of the "scientists" with a Russian accent looking in the rearview mirror, narrating with equations, and refusing to look ahead because he may get too dizzy. The econometrician Robert Engel, an otherwise charming gentleman, invented a very complicated statistical method called GARCH and got a Nobel for it. No one tested it to see if it has any validity in real life. Simpler, less sexy methods fare exceedingly better, but they do not take you to Stockholm. You have an expert problem in Stockholm, and I will discuss it in Chapter 17.

This unfitness of complicated methods seems to apply to all methods. Another study effectively tested practitioners of something called game theory, in which the most notorious player is John Nash, the schizophrenic mathematician made famous by the film *A Beautiful Mind*. Sadly, for all the intellectual appeal of these methods and all the media attention, its practitioners are no better at predicting than university students.

There is another problem, and it is a little more worrisome. Makridakis and Hibon were to find out that the strong empirical evidence of their studies has been ignored by theoretical statisticians. Furthermore, they encountered shocking hostility toward their empirical verifications. "Instead [statisticians] have concentrated their efforts in building more sophisticated models without regard to the ability of such models to more accurately predict real-life data," Makridakis and Hibon write.

Someone may counter with the following argument: Perhaps economists' forecasts create feedback that cancels their effect (this is called the Lucas critique, after the economist Robert Lucas). Let's say economists predict inflation; in response to these expectations the Federal Reserve acts and lowers inflation. So you cannot judge the forecast accuracy in economics as you would with other events. I agree with this point, but I do not believe that it is the cause of the economists' failure to predict. The world is far too complicated for their discipline.

When an economist fails to predict outliers he often invokes the issue of earthquakes or revolutions, claiming that he is not into geodesics, atmospheric sciences, or political science, instead of incorporating these fields into his studies and accepting that his field does not exist in isolation. Economics is the most insular of fields; it is the one that quotes least from outside itself! Economics is perhaps the subject that currently has the highest number of philistine scholars—scholarship without erudition and natural curiosity can close your mind and lead to the fragmentation of disciplines.

"OTHER THAN THAT," IT WAS OKAY

We have used the story of the Sydney Opera House as a springboard for our discussion of prediction. We will now address another constant in human nature: a systematic error made by project planners, coming from a mixture of human nature, the complexity of the world, or the structure of organizations. In order to survive, institutions may need to give themselves and others the appearance of having a "vision."

Plans fail because of what we have called tunneling, the neglect of sources of uncertainty outside the plan itself.

The typical scenario is as follows. Joe, a nonfiction writer, gets a book contract with a set final date for delivery two years from now. The topic is relatively easy: the authorized biography of the writer Salman Rushdie, for which Joe has compiled ample data. He has even tracked down Rushdie's former girlfriends and is thrilled at the prospect of pleasant interviews. Two years later, minus, say, three months, he calls to explain to the publisher that he will be *a little* delayed. The publisher has seen this coming; he is used to authors being late. The publishing house now has cold feet because the subject has *unexpectedly* faded from public attention—the firm projected that interest in Rushdie would remain high, but attention has faded, seemingly because the Iranians, for some reason, lost interest in killing him.

Let's look at the source of the biographer's underestimation of the time for completion. He projected his own schedule, but he tunneled, as he did not forecast that some "external" events would emerge to slow him down. Among these external events were the disasters on September 11, 2001, which set him back several months; trips to Minnesota to assist his ailing mother (who eventually recovered); and many more, like a broken engagement (though not with Rushdie's ex-girlfriend). "Other than that," it was all within his plan; his own work did not stray the least from schedule. He does not feel responsible for his failure.*

The unexpected has a one-sided effect with projects. Consider the

* The book you have in your hands is approximately and "unexpectedly" fifteen months late.

track records of builders, paper writers, and contractors. The unexpected almost always pushes in a single direction: higher costs and a longer time to completion. On very rare occasions, as with the Empire State Building, you get the opposite: shorter completion and lower costs—these occasions are truly exceptional.

We can run experiments and test for repeatability to verify if such errors in projection are part of human nature. Researchers have tested how students estimate the time needed to complete their projects. In one representative test, they broke a group into two varieties, optimistic and pessimistic. Optimistic students promised twenty-six days; the pessimistic ones forty-seven days. The average actual time to completion turned out to be fifty-six days.

The example of Joe the writer is not acute. I selected it because it concerns a repeatable, routine task—for such tasks our planning errors are milder. With projects of great novelty, such as a military invasion, an all-out war, or something entirely new, errors explode upward. In fact, the more routine the task, the better you learn to forecast. But there is always something nonroutine in our modern environment.

There may be incentives for people to promise shorter completion dates—in order to win the book contract or in order for the builder to get your down payment and use it for his upcoming trip to Antigua. But the planning problem exists even where there is no incentive to underestimate the duration (or the costs) of the task. As I said earlier, we are too narrow-minded a species to consider the possibility of events straying from our mental projections, but furthermore, we are too focused on matters internal to the project to take into account external uncertainty, the "unknown unknown," so to speak, the contents of the unread books.

There is also the nerd effect, which stems from the mental elimination of off-model risks, or *focusing* on what you know. You view the world from *within* a model. Consider that most delays and cost overruns arise from unexpected elements that did not enter into the plan—that is, they lay outside the model at hand—such as strikes, electricity shortages, accidents, bad weather, or rumors of Martian invasions. These small Black Swans that threaten to hamper our projects do not seem to be taken into account. They are too abstract—we don't know how they look and cannot talk about them intelligently.

We cannot truly plan, because we do not understand the future—but this is not necessarily bad news. We could plan *while bearing in mind such limitations*. It just takes guts.

The Beauty of Technology: Excel Spreadsheets

In the not too distant past, say the precomputer days, projections remained vague and qualitative, one had to make a mental effort to keep track of them, and it was a strain to push scenarios into the future. It took pencils, erasers, reams of paper, and huge wastebaskets to engage in the activity. Add to that an accountant's love for tedious, slow work. The activity of projecting, in short, was effortful, undesirable, and marred with self-doubt.

But things changed with the intrusion of the spreadsheet. When you put an Excel spreadsheet into computer-literate hands you get a "sales projection" effortlessly extending ad infinitum! Once on a page or on a computer screen, or, worse, in a PowerPoint presentation, the projection takes on a life of its own, losing its vagueness and abstraction and becoming what philosophers call reified, invested with concreteness; it takes on a new life as a tangible object.

My friend Brian Hinchcliffe suggested the following idea when we were both sweating at the local gym. Perhaps the ease with which one can project into the future by dragging cells in these spreadsheet programs is responsible for the armies of forecasters confidently producing longer-term forecasts (all the while tunneling on their assumptions). We have become worse planners than the Soviet Russians thanks to these potent computer programs given to those who are incapable of handling their knowledge. Like most commodity traders, Brian is a man of incisive and sometimes brutally painful realism.

A classical mental mechanism, called anchoring, seems to be at work here. You lower your anxiety about uncertainty by producing a number, then you "anchor" on it, like an object to hold on to in the middle of a vacuum. This anchoring mechanism was discovered by the fathers of the psychology of uncertainty, Danny Kahneman and Amos Tversky, early in their heuristics and biases project. It operates as follows. Kahneman and Tversky had their subjects spin a wheel of fortune. The subjects first looked at the number on the wheel, *which they knew was random,* then they were asked to estimate the number of African countries in the United Nations. Those who had a low number on the wheel estimated a low number of African nations; those with a high number produced a higher estimate.

Similarly, ask someone to provide you with the last four digits of his social security number. Then ask him to estimate the number of dentists in

Manhattan. You will find that by making him aware of the four-digit number, you elicit an estimate that is correlated with it.

We use reference points in our heads, say sales projections, and start building beliefs around them because less mental effort is needed to compare an idea to a reference point than to evaluate it in the absolute (*System 1* at work!). We cannot work without a point of reference.

So the introduction of a reference point in the forecaster's mind will work wonders. This is no different from a starting point in a bargaining episode: you open with high number ("I want a million for this house"); the bidder will answer "only eight-fifty"—the discussion will be determined by that initial level.

The Character of Prediction Errors

Like many biological variables, life expectancy is from Mediocristan, that is, it is subjected to mild randomness. It is not scalable, since the older we get, the less likely we are to live. In a developed country a newborn female is expected to die at around 79, according to insurance tables. When she reaches her 79th birthday, her life expectancy, assuming that she is in typical health, is another 10 years. At the age of 90, she should have another 4.7 years to go. At the age of 100, 2.5 years. At the age of 119, if she miraculously lives that long, she should have about nine months left. As she lives beyond the expected date of death, the number of additional years to go decreases. This illustrates the major property of random variables related to the bell curve. The conditional expectation of additional life drops as a person gets older.

With human projects and ventures we have another story. These are often scalable, as I said in Chapter 3. With scalable variables, the ones from Extremistan, you will witness the exact opposite effect. Let's say a project is expected to terminate in 79 days, the same expectation in days as the newborn female has in years. On the 79th day, if the project is not finished, it will be expected to take another 25 days to complete. But on the 90th day, if the project is still not completed, it should have about 58 days to go. On the 100th, it should have 89 days to go. On the 119th, it should have an extra 149 days. On day 600, if the project is not done, you will be expected to need an extra 1,590 days. As you see, *the longer you wait, the longer you will be expected to wait.*

Let's say you are a refugee waiting for the return to your homeland. Each day that passes you are getting farther from, not closer to, the day of triumphal return. The same applies to the completion date of your next opera house. If it was expected to take two years, and three years later you are asking questions, do not expect the project to be completed any time soon. If wars last on average six months, and your conflict has been going on for two years, expect another few years of problems. The Arab-Israeli conflict is sixty years old, and counting—yet it was considered "a simple problem" sixty years ago. (Always remember that, in a modern environment, wars last longer and kill more people than is typically planned.) Another example: Say that you send your favorite author a letter, knowing that he is busy and has a two-week turnaround. If three weeks later your mailbox is still empty, do not expect the letter to come tomorrow—it will take on average another three weeks. If three months later you still have nothing, you will have to expect to wait another year. Each day will bring you closer to your death but further from the receipt of the letter.

This subtle but extremely consequential property of scalable randomness is unusually counterintuitive. We misunderstand the logic of large deviations from the norm.

I will get deeper into these properties of scalable randomness in Part Three. But let us say for now that they are central to our misunderstanding of the business of prediction.

DON'T CROSS A RIVER IF IT IS (ON AVERAGE) FOUR FEET DEEP

Corporate and government projections have an additional easy-to-spot flaw: they do not attach a *possible error rate* to their scenarios. Even in the absence of Black Swans this omission would be a mistake.

I once gave a talk to policy wonks at the Woodrow Wilson Center in Washington, D.C., challenging them to be aware of our weaknesses in seeing ahead.

The attendees were tame and silent. What I was telling them was against everything they believed and stood for; I had gotten carried away with my aggressive message, but they looked thoughtful, compared to the testosterone-charged characters one encounters in business. I felt guilty for my aggressive stance. Few asked questions. The person who organized the talk and invited me must have been pulling a joke on his colleagues. I was like an aggressive atheist making his case in front of a synod of cardinals, while dispensing with the usual formulaic euphemisms.

Yet some members of the audience were sympathetic to the message. One anonymous person (he is employed by a governmental agency) explained to me privately after the talk that in January 2004 his department was forecasting the price of oil for twenty-five years later at $27 a barrel, slightly higher than what it was at the time. Six months later, around June 2004, after oil doubled in price, they had to revise their estimate to $54 (the price of oil is currently, as I am writing these lines, close to $79 a barrel). It did not dawn on them that it was ludicrous to forecast a second time given that their forecast was off so early and so markedly, that this business of forecasting had to be somehow questioned. And they were looking *twenty-five years* ahead! Nor did it hit them that there was something called an error rate to take into account.*

Forecasting without incorporating an error rate uncovers three fallacies, all arising from the same misconception about the nature of uncertainty.

The first fallacy: *variability matters.* The first error lies in taking a projection too seriously, without heeding its accuracy. Yet, for planning purposes, the accuracy in your forecast matters far more the forecast itself. I will explain it as follows.

Don't cross a river if it is four feet deep on average. You would take a different set of clothes on your trip to some remote destination if I told you that the temperature was expected to be seventy degrees Fahrenheit, with an expected error rate of forty degrees than if I told you that my margin of error was only five degrees. The policies we need to make decisions on should depend far more on the range of possible outcomes than on the expected final number. I have seen, while working for a bank, how people project cash flows for companies without wrapping them in the thinnest layer of uncertainty. Go to the stockbroker and check on what method they use to forecast sales ten years ahead to "calibrate" their valuation models. Go find out how analysts forecast government deficits. Go to a bank or security-analysis training program and see how they teach

* While forecast errors have always been entertaining, commodity prices have been a great trap for suckers. Consider this 1970 forecast by U.S. officials (signed by the U.S. Secretaries of the Treasury, State, Interior, and Defense): "the standard price of foreign crude oil by 1980 may well decline and will in any event not experience a substantial increase." Oil prices went up tenfold by 1980. I just wonder if current forecasters lack in intellectual curiosity or if they are intentionally ignoring forecast errors.

Also note this additional aberration: since high oil prices are marking up their inventories, oil companies are making record bucks and oil executives are getting huge bonuses because "they did a good job"—as if they brought profits by *causing* the rise of oil prices.