The Illusion of Understanding

The trader-philosopher-statistician Nassim Taleb could also be considered a psychologist. In *The Black Swan*, Taleb introduced the notion of a *narrative fallacy* to describe how flawed stories of the past shape our views of the world and our expectations for the future. Narrative fallacies arise inevitably from our continuous attempt to make sense of the world. The explanatory stories that people find compelling are simple; are concrete rather than abstract; assign a larger role to talent, stupidity, and intentions than to luck; and focus on a few striking events that happened rather than on the countless events that failed to happen. Any recent salient event is a candidate to become the kernel of a causal narrative. Taleb suggests that we humans constantly fool ourselves by constructing flimsy accounts of the past and believing they are true.

Good stories provide a simple and coherent account >

A compelling narrative fosters an illusion of inevitability. Consider the story of how Google turned into a giant of the technology industry. Two creative graduate students in the computer science department at Stanford University come up with a superior way of searching information on the Internet. They seek and obtain funding to start a company and make a series of decisions that work out well. Within a few years, the company they started is one of the most valuable stocks in America, and the two former graduate students are among the richest people on the planet. On one memorable occasion, they were lucky, which makes the story even more compelling: a year after founding Google, they were willing to sell their company for less than $1 million, but the buyer said the price was too high. Mentioning the single lucky incident actually makes it easier to underestimate the multitude of ways in which luck affected the outcome.

A detailed history would specify the decisions of Google's founders, but for our purposes it suffices to say that almost every choice they made had a good outcome. A more complete narrative would describe the actions of the firms that Google defeated. The hapless competitors would appear to be blind, slow, and altogether inadequate in dealing with the threat that eventually overwhelmed them.

I intentionally told this tale blandly, but you get the idea: there is a very good story here. Fleshed out in more detail, the story could give

you the sense that you understand what made Google succeed; it would also make you feel that you have learned a valuable general lesson about what makes businesses succeed. Unfortunately, there is good reason to believe that your sense of understanding and learning from the Google story is largely illusory. The ultimate test of an explanation is whether it would have made the event predictable in advance. No story of Google's unlikely success will meet that test, because no story can include the myriad of events that would have caused a different outcome. The human mind does not deal well with nonevents. The fact that many of the important events that did occur involve choices further tempts you to exaggerate the role of skill and underestimate the part that luck played in the outcome. Because every critical decision turned out well, the record suggests almost flawless prescience—but bad luck could have disrupted any one of the successful steps. The halo effect adds the final touches, lending an aura of invincibility to the heroes of the story.

Like watching a skilled rafter avoiding one potential calamity after another as he goes down the rapids, the unfolding of the Google story is thrilling because of the constant risk of disaster. However, there is fop an instructive difference between the two cases. The skilled rafter has gone down rapids hundreds of times. He has learned to read the roiling water in front of him and to anticipate obstacles. He has learned to make the tiny adjustments of posture that keep him upright. There are fewer opportunities for young men to learn how to create a giant company, and fewer chances to avoid hidden rocks—such as a brilliant innovation by a competing firm. Of course there was a great deal of skill in the Google story, but luck played a more important role in the actual event than it does in the telling of it. And the more luck was involved, the less there is to be learned.

At work here is that powerful WY SIATI rule. You cannot help dealing with the limited information you have as if it were all there is to know. You build the best possible story from the information available to you, and if it is a good story, you believe it. Paradoxically, it is easier to construct a coherent story when you know little, when there are fewer pieces to fit into the puzzle. Our comforting conviction that the world makes sense rests on a secure foundation: our almost unlimited ability to ignore our ignorance.

I have heard of too many people who "knew well before it happened that the 2008 financial crisis was inevitable." This sentence contains a

highly objectionable word, which should be removed from our vocabulary in discussions of major events. The word is, of course, *knew*. Some people thought well in advance that there would be a crisis, but they did not know it. They now say they knew it because the crisis did in fact happen. This is a misuse of an important concept. In everyday language, we apply the word *know* only when what was known is true and can be shown to be true. We can know something only if it is both true and knowable. But the people who thought there would be a crisis (and there are fewer of them than now remember thinking it) could not conclusively show it at the time. Many intelligent and well-informed people were keenly interested in the future of the economy and did not believe a catastrophe was imminent; I infer from this fact that the crisis was not knowable. What is perverse about the use of *know* in this context is not that some individuals get credit for prescience that they do not deserve. It is that the language implies that the world is more knowable than it is. It helps perpetuate a pernicious illusion.

The core of the illusion is that we believe we understand the past, which implies that the future also should be knowable, but in fact we understand the past less than we believe we do. *Know* is not the only word that fosters this illusion. In common usage, the words *intuition* and *premonition* also are reserved for past thoughts that turned out to be true. The statement “I had a premonition that the marriage would not last, but I was wrong” sounds odd, as does any sentence about an intuition that turned out to be false. To think clearly about the future, we need to clean up the language that we use in labeling the beliefs we had in the past.

The Social Costs of Hindsight

The mind that makes up narratives about the past is a sense-making organ. When an unpredicted event occurs, we immediately adjust our view of the world to accommodate the surprise. Imagine yourself before a football game between two teams that have the same record of wins and losses. Now the game is over, and one team trashed the other. In your revised model of the world, the winning team is much stronger than the loser, and your view of the past as well as of the future has been altered be fpy that new perception. Learning from surprises is a reasonable thing to do, but it can have some dangerous consequences.

A general limitation of the human mind is its imperfect ability to reconstruct past states of knowledge, or beliefs that have changed. Once you adopt a new view of the world (or of any part of it), you immediately lose much of your ability to recall what you used to believe before your mind changed.

Many psychologists have studied what happens when people change their minds. Choosing a topic on which minds are not completely made up—say, the death penalty—the experimenter carefully measures people's attitudes. Next, the participants see or hear a persuasive pro or con message. Then the experimenter measures people's attitudes again; they usually are closer to the persuasive message they were exposed to. Finally, the participants report the opinion they held beforehand. This task turns out to be surprisingly difficult. Asked to reconstruct their former beliefs, people retrieve their current ones instead—an instance of substitution—and many cannot believe that they ever felt differently.

Your inability to reconstruct past beliefs will inevitably cause you to underestimate the extent to which you were surprised by past events. Baruch Fischh off first demonstrated this "I-knew-it-all-along" effect, or *hindsight bias*, when he was a student in Jerusalem. Together with Ruth Beyth (another of our students), Fischh off conducted a survey before President Richard Nixon visited China and Russia in 1972. The respondents assigned probabilities to fifteen possible outcomes of Nixon's diplomatic initiatives. Would Mao Zedong agree to meet with Nixon? Might the United States grant diplomatic recognition to China? After decades of enmity, could the United States and the Soviet Union agree on anything significant?

After Nixon's return from his travels, Fischh off and Beyth asked the same people to recall the probability that they had originally assigned to each of the fifteen possible outcomes. The results were clear. If an event had actually occurred, people exaggerated the probability that they had assigned to it earlier. If the possible event had not come to pass, the participants erroneously recalled that they had always considered it unlikely. Further experiments showed that people were driven to overstate the accuracy not only of their original predictions but also of those made by others. Similar results have been found for other events that gripped public attention, such as the O. J. Simpson murder trial and the impeachment of President Bill Clinton. The tendency to

revise the history of one's beliefs in light of what actually happened produces a robust cognitive illusion.

Hindsight bias has pernicious effects on the evaluations of decision makers. It leads observers to assess the quality of a decision not by whether the process was sound but by whether its outcome was good or bad. Consider a low-risk surgical intervention in which an unpredictable accident occurred that caused the patient's death. The jury will be prone to believe, after the fact, that the operation was actually risky and that the doctor who ordered it should have known better. This outcome bias makes it almost impossible to evaluate a decision properly—in terms of the beliefs that were reasonable when the decision was made.

Hindsight is especially unkind to decision makers who act as agents for others—physicians, financial advisers, third-base coaches, CEOs, social workers, diplomats, politicians. We are prone to blame decision makers for good decisions that worked out badly and to give them too little credit for successful movesecap that appear obvious only after the fact. There is a clear *outcome bias*. When the outcomes are bad, the clients often blame their agents for not seeing the handwriting on the wall—forgetting that it was written in invisible ink that became legible only afterward. Actions that seemed prudent in foresight can look irresponsibly negligent in hindsight. Based on an actual legal case, students in California were asked whether the city of Duluth, Minnesota, should have shouldered the considerable cost of hiring a full-time bridge monitor to protect against the risk that debris might get caught and block the free flow of water. One group was shown only the evidence available at the time of the city's decision; 24% of these people felt that Duluth should take on the expense of hiring a flood monitor. The second group was informed that debris had blocked the river, causing major flood damage; 56% of these people said the city should have hired the monitor, although they had been explicitly instructed not to let hindsight distort their judgment.

The worse the consequence, the greater the hindsight bias. In the case of a catastrophe, such as 9/11, we are especially ready to believe that the officials who failed to anticipate it were negligent or blind. On July 10, 2001, the Central Intelligence Agency obtained information that al-Qaeda might be planning a major attack against the United States. George Tenet, director of the CIA, brought the information not to President George W. Bush but to National Security Adviser

Condoleezza Rice. When the facts later emerged, Ben Bradlee, the legendary executive editor of *The Washington Post*, declared, “It seems to me elementary that if you’ve got the story that’s going to dominate history you might as well go right to the president.” But on July 10, no one knew—or could have known—that this tidbit of intelligence would turn out to dominate history.

Because adherence to standard operating procedures is difficult to second-guess, decision makers who expect to have their decisions scrutinized with hindsight are driven to bureaucratic solutions—and to an extreme reluctance to take risks. As malpractice litigation became more common, physicians changed their procedures in multiple ways: ordered more tests, referred more cases to specialists, applied conventional treatments even when they were unlikely to help. These actions protected the physicians more than they benefited the patients, creating the potential for conflicts of interest. Increased accountability is a mixed blessing.

Although hindsight and the outcome bias generally foster risk aversion, they also bring undeserved rewards to irresponsible risk seekers, such as a general or an entrepreneur who took a crazy gamble and won. Leaders who have been lucky are never punished for having taken too much risk. Instead, they are believed to have had the flair and foresight to anticipate success, and the sensible people who doubted them are seen in hindsight as mediocre, timid, and weak. A few lucky gambles can crown a reckless leader with a halo of prescience and boldness.

Recipes for Success

The sense-making machinery of System 1 makes us see the world as more tidy, simple, predictable, and coherent than it really is. The illusion that one has understood the past feeds the further illusion that one can predict and control the future. These illusions are comforting. They reduce the anxiety that we would experience if we allowed ourselves to fully acknowledge the uncertainties of existence. We all have a need for the reassuring message that actions have appropriate consequences, and that success will reward wisdom and courage. Many bdecpusiness books are tailor-made to satisfy this need.

Do leaders and management practices influence the outcomes of firms in the market? Of course they do, and the effects have been

confirmed by systematic research that objectively assessed the characteristics of CEOs and their decisions, and related them to subsequent outcomes of the firm. In one study, the CEOs were characterized by the strategy of the companies they had led before their current appointment, as well as by management rules and procedures adopted after their appointment. CEOs do influence performance, but the effects are much smaller than a reading of the business press suggests.

Researchers measure the strength of relationships by a correlation coefficient, which varies between 0 and 1. The coefficient was defined earlier (in relation to regression to the mean) by the extent to which two measures are determined by shared factors. A very generous estimate of the correlation between the success of the firm and the quality of its CEO might be as high as .30, indicating 30% overlap. To appreciate the significance of this number, consider the following question:

> Suppose you consider many pairs of firms. The two firms in each pair are generally similar, but the CEO of one of them is better than the other. How often will you find that the firm with the stronger CEO is the more successful of the two?

In a well-ordered and predictable world, the correlation would be perfect (1), and the stronger CEO would be found to lead the more successful firm in 100% of the pairs. If the relative success of similar firms was determined entirely by factors that the CEO does not control (call them luck, if you wish), you would find the more successful firm led by the weaker CEO 50% of the time. A correlation of .30 implies that you would find the stronger CEO leading the stronger firm in about 60% of the pairs—an improvement of a mere 10 percentage points over random guessing, hardly grist for the hero worship of CEOs we so often witness.

If you expected this value to be higher—and most of us do—then you should take that as an indication that you are prone to overestimate the predictability of the world you live in. Make no mistake: improving the odds of success from 1:1 to 3:2 is a very significant advantage, both at the racetrack and in business. From the perspective of most business writers, however, a CEO who has so little control over performance would not be particularly impressive even if her firm did well. It is difficult to imagine people lining up at airport bookstores to buy a book that

enthusiastically describes the practices of business leaders who, on average, do somewhat better than chance. Consumers have a hunger for a clear message about the determinants of success and failure in business, and they need stories that offer a sense of understanding, however illusory.

In his penetrating book *The Halo Effect*, Philip Rosenzweig, a business school professor based in Switzerland, shows how the demand for illusory certainty is met in two popular genres of business writing: histories of the rise (usually) and fall (occasionally) of particular individuals and companies, and analyses of differences between successful and less successful firms. He concludes that stories of success and failure consistently exaggerate the impact of leadership style and management practices on firm outcomes, and thus their message is rarely useful.

To appreciate what is going on, imagine that business experts, such as other CEOs, are asked to comment on the reputation of the chief executive of a company. They popare keenly aware of whether the company has recently been thriving or failing. As we saw earlier in the case of Google, this knowledge generates a halo. The CEO of a successful company is likely to be called flexible, methodical, and decisive. Imagine that a year has passed and things have gone sour. The same executive is now described as confused, rigid, and authoritarian. Both descriptions sound right at the time: it seems almost absurd to call a successful leader rigid and confused, or a struggling leader flexible and methodical.

Indeed, the halo effect is so powerful that you probably find yourself resisting the idea that the same person and the same behaviors appear methodical when things are going well and rigid when things are going poorly. Because of the halo effect, we get the causal relationship backward: we are prone to believe that the firm fails because its CEO is rigid, when the truth is that the CEO appears to be rigid because the firm is failing. This is how illusions of understanding are born.

The halo effect and outcome bias combine to explain the extraordinary appeal of books that seek to draw operational morals from systematic examination of successful businesses. One of the best-known examples of this genre is Jim Collins and Jerry I. Porras's *Built to Last*. The book contains a thorough analysis of eighteen pairs of competing companies, in which one was more successful than the other. The data for these comparisons are ratings of various aspects of

corporate culture, strategy, and management practices. “We believe every CEO, manager, and entrepreneur in the world should read this book,” the authors proclaim. “You can build a visionary company.”

The basic message of *Built to Last* and other similar books is that good managerial practices can be identified and that good practices will be rewarded by good results. Both messages are overstated. The comparison of firms that have been more or less successful is to a significant extent a comparison between firms that have been more or less lucky. Knowing the importance of luck, you should be particularly suspicious when highly consistent patterns emerge from the comparison of successful and less successful firms. In the presence of randomness, regular patterns can only be mirages.

Because luck plays a large role, the quality of leadership and management practices cannot be inferred reliably from observations of success. And even if you had perfect foreknowledge that a CEO has brilliant vision and extraordinary competence, you still would be unable to predict how the company will perform with much better accuracy than the flip of a coin. On average, the gap in corporate profitability and stock returns between the outstanding firms and the less successful firms studied in *Built to Last* shrank to almost nothing in the period following the study. The average profitability of the companies identified in the famous *In Search of Excellence* dropped sharply as well within a short time. A study of *Fortune*’s “Most Admired Companies” finds that over a twenty-year period, the firms with the worst ratings went on to earn much higher stock returns than the most admired firms.

You are probably tempted to think of causal explanations for these observations: perhaps the successful firms became complacent, the less successful firms tried harder. But this is the wrong way to think about what happened. The average gap must shrink, because the original gap was due in good part to luck, which contributed both to the success of the top firms and to the lagging performance of the rest. We have already encountered this statistical fact of life: regression to the mean.

Stories of how businesses rise and fall strike a chord with readers by offering what the human mind needs: a simple message of triumph and failure that identifies clear causes and ignores the determinative power of luck and the inevitability of regression. These stories induce and maintain an illusion of understanding, imparting lessons of little enduring value to readers who are all too eager to believe them.

Speaking of Hindsight

“The mistake appears obvious, but it is just hindsight. You could not have known in advance.”

“He’s learning too much from this success story, which is too tidy. He has fallen for a narrative fallacy.”

“She has no evidence for saying that the firm is badly managed. All she knows is that its stock has gone down. This is an outcome bias, part hindsight and part halo effect.”

“Let’s not fall for the outcome bias. This was a stupid decision even though it worked out well.”

The Illusion of Validity

System 1 is designed to jump to conclusions from little evidence—and it is not designed to know the size of its jumps. Because of WYSIATI, only the evidence at hand counts. Because of confidence by coherence, the subjective confidence we have in our opinions reflects the coherence of the story that System 1 and System 2 have constructed. The amount of evidence and its quality do not count for much, because poor evidence can make a very good story. For some of our most important beliefs we have no evidence at all, except that people we love and trust hold these beliefs. Considering how little we know, the confidence we have in our beliefs is preposterous—and it is also essential.

The Illusion of Validity

Many decades ago I spent what seemed like a great deal of time under a scorching sun, watching groups of sweaty soldiers as they solved a problem. I was doing my national service in the Israeli Army at the time. I had completed an undergraduate degree in psychology, and after a year as an infantry officer was assigned to the army's Psychology Branch, where one of my occasional duties was to help evaluate candidates for officer training. We used methods that had been developed by the British Army in World War II.

One test, called the "leaderless group challenge," was conducted on an obstacle field. Eight candidates, strangers to each other, with all insignia of rank removed and only numbered tags to identify them, were instructed to lift a long log from the ground and haul it to a wall about six feet high. The entire group had to get to the other side of the wall without the log touching either the ground or the wall, and without anyone touching the wall. If any of these things happened, they had to declare itsigpЛЬ T and start again.

There was more than one way to solve the problem. A common solution was for the team to send several men to the other side by crawling over the pole as it was held at an angle, like a giant fishing rod, by other members of the group. Or else some soldiers would climb onto someone's shoulders and jump across. The last man would then have to jump up at the pole, held up at an angle by the rest of the group,

shinny his way along its length as the others kept him and the pole suspended in the air, and leap safely to the other side. Failure was common at this point, which required them to start all over again.

As a colleague and I monitored the exercise, we made note of who took charge, who tried to lead but was rebuffed, how cooperative each soldier was in contributing to the group effort. We saw who seemed to be stubborn, submissive, arrogant, patient, hot-tempered, persistent, or a quitter. We sometimes saw competitive spite when someone whose idea had been rejected by the group no longer worked very hard. And we saw reactions to crisis: who berated a comrade whose mistake had caused the whole group to fail, who stepped forward to lead when the exhausted team had to start over. Under the stress of the event, we felt, each man's true nature revealed itself. Our impression of each candidate's character was as direct and compelling as the color of the sky.

After watching the candidates make several attempts, we had to summarize our impressions of soldiers' leadership abilities and determine, with a numerical score, who should be eligible for officer training. We spent some time discussing each case and reviewing our impressions. The task was not difficult, because we felt we had already seen each soldier's leadership skills. Some of the men had looked like strong leaders, others had seemed like wimps or arrogant fools, others mediocre but not hopeless. Quite a few looked so weak that we ruled them out as candidates for officer rank. When our multiple observations of each candidate converged on a coherent story, we were completely confident in our evaluations and felt that what we had seen pointed directly to the future. The soldier who took over when the group was in trouble and led the team over the wall was a leader at that moment. The obvious best guess about how he would do in training, or in combat, was that he would be as effective then as he had been at the wall. Any other prediction seemed inconsistent with the evidence before our eyes.

Because our impressions of how well each soldier had performed were generally coherent and clear, our formal predictions were just as definite. A single score usually came to mind and we rarely experienced doubts or formed conflicting impressions. We were quite willing to declare, "This one will never make it," "That fellow is mediocre, but he should do okay," or "He will be a star." We felt no need to question our forecasts, moderate them, or equivocate. If challenged, however, we were prepared to admit, "But of course anything could happen." We

were willing to make that admission because, despite our definite impressions about individual candidates, we knew with certainty that our forecasts were largely useless.

The evidence that we could not forecast success accurately was overwhelming. Every few months we had a feedback session in which we learned how the cadets were doing at the officer-training school and could compare our assessments against the opinions of commanders who had been monitoring them for some time. The story was always the same: our ability to predict performance at the school was negligible. Our forecasts were better than blind guesses, but not by much.

We weed [?]re downcast for a while after receiving the discouraging news. But this was the army. Useful or not, there was a routine to be followed and orders to be obeyed. Another batch of candidates arrived the next day. We took them to the obstacle field, we faced them with the wall, they lifted the log, and within a few minutes we saw their true natures revealed, as clearly as before. The dismal truth about the quality of our predictions had no effect whatsoever on how we evaluated candidates and very little effect on the confidence we felt in our judgments and predictions about individuals.

What happened was remarkable. The global evidence of our previous failure should have shaken our confidence in our judgments of the candidates, but it did not. It should also have caused us to moderate our predictions, but it did not. We knew as a general fact that our predictions were little better than random guesses, but we continued to feel and act as if each of our specific predictions was valid. I was reminded of the Müller-Lyer illusion, in which we know the lines are of equal length yet still see them as being different. I was so struck by the analogy that I coined a term for our experience: the *illusion of validity*.

I had discovered my first cognitive illusion.

Decades later, I can see many of the central themes of my thinking—and of this book—in that old story. Our expectations for the soldiers' future performance were a clear instance of substitution, and of the representativeness heuristic in particular. Having observed one hour of a soldier's behavior in an artificial situation, we felt we knew how well he would face the challenges of officer training and of leadership in

combat. Our predictions were completely nonregressive—we had no reservations about predicting failure or outstanding success from weak evidence. This was a clear instance of WYSIATI. We had compelling impressions of the behavior we observed and no good way to represent our ignorance of the factors that would eventually determine how well the candidate would perform as an officer.

Looking back, the most striking part of the story is that our knowledge of the general rule—that we could not predict—had no effect on our confidence in individual cases. I can see now that our reaction was similar to that of Nisbett and Borgida's students when they were told that most people did not help a stranger suffering a seizure. They certainly believed the statistics they were shown, but the base rates did not influence their judgment of whether an individual they saw on the video would or would not help a stranger. Just as Nisbett and Borgida showed, people are often reluctant to infer the particular from the general.

Subjective confidence in a judgment is not a reasoned evaluation of the probability that this judgment is correct. Confidence is a feeling, which reflects the coherence of the information and the cognitive ease of processing it. It is wise to take admissions of uncertainty seriously, but declarations of high confidence mainly tell you that an individual has constructed a coherent story in his mind, not necessarily that the story is true.

The Illusion of Stock-Picking Skill

In 1984, Amos and I and our friend Richard Thaler visited a Wall Street firm. Our host, a senior investment manager, had invited us to discuss the role of judgment biases in investing. I knew so little about finance that I did not even know what to ask him, but I remember one exchange. "When you sell a stock," d n I asked, "who buys it?" He answered with a wave in the vague direction of the window, indicating that he expected the buyer to be someone else very much like him. That was odd: What made one person buy and the other sell? What did the sellers think they knew that the buyers did not?

Since then, my questions about the stock market have hardened into a larger puzzle: a major industry appears to be built largely on an *illusion of skill*. Billions of shares are traded every day, with many people buying each stock and others selling it to them. It is not unusual

for more than 100 million shares of a single stock to change hands in one day. Most of the buyers and sellers know that they have the same information; they exchange the stocks primarily because they have different opinions. The buyers think the price is too low and likely to rise, while the sellers think the price is high and likely to drop. The puzzle is why buyers and sellers alike think that the current price is wrong. What makes them believe they know more about what the price should be than the market does? For most of them, that belief is an illusion.

In its broad outlines, the standard theory of how the stock market works is accepted by all the participants in the industry. Everybody in the investment business has read Burton Malkiel's wonderful book *A Random Walk Down Wall Street*. Malkiel's central idea is that a stock's price incorporates all the available knowledge about the value of the company and the best predictions about the future of the stock. If some people believe that the price of a stock will be higher tomorrow, they will buy more of it today. This, in turn, will cause its price to rise. If all assets in a market are correctly priced, no one can expect either to gain or to lose by trading. Perfect prices leave no scope for cleverness, but they also protect fools from their own folly. We now know, however, that the theory is not quite right. Many individual investors lose consistently by trading, an achievement that a dart-throwing chimp could not match. The first demonstration of this startling conclusion was collected by Terry Odean, a finance professor at UC Berkeley who was once my student.

Odean began by studying the trading records of 10,000 brokerage accounts of individual investors spanning a seven-year period. He was able to analyze every transaction the investors executed through that firm, nearly 163,000 trades. This rich set of data allowed Odean to identify all instances in which an investor sold some of his holdings in one stock and soon afterward bought another stock. By these actions the investor revealed that he (most of the investors were men) had a definite idea about the future of the two stocks: he expected the stock that he chose to buy to do better than the stock he chose to sell.

To determine whether those ideas were well founded, Odean compared the returns of the stock the investor had sold and the stock he had bought in its place, over the course of one year after the transaction. The results were unequivocally bad. On average, the shares that individual traders sold did better than those they bought, by

a very substantial margin: 3.2 percentage points per year, above and beyond the significant costs of executing the two trades.

It is important to remember that this is a statement about averages: some individuals did much better, others did much worse. However, it is clear that for the large majority of individual investors, taking a shower and doing nothing would have been a better policy than implementing the ideas that came to their minds. Later research by Odean and his colleague Brad Barber supported this conclusion. In a paper titled “Trading Is Hazardous to Yourt-t? Wealth,” they showed that, on average, the most active traders had the poorest results, while the investors who traded the least earned the highest returns. In another paper, titled “Boys Will Be Boys,” they showed that men acted on their useless ideas significantly more often than women, and that as a result women achieved better investment results than men.

Of course, there is always someone on the other side of each transaction; in general, these are financial institutions and professional investors, who are ready to take advantage of the mistakes that individual traders make in choosing a stock to sell and another stock to buy. Further research by Barber and Odean has shed light on these mistakes. Individual investors like to lock in their gains by selling “winners,” stocks that have appreciated since they were purchased, and they hang on to their losers. Unfortunately for them, recent winners tend to do better than recent losers in the short run, so individuals sell the wrong stocks. They also buy the wrong stocks. Individual investors predictably flock to companies that draw their attention because they are in the news. Professional investors are more selective in responding to news. These findings provide some justification for the label of “smart money” that finance professionals apply to themselves.

Although professionals are able to extract a considerable amount of wealth from amateurs, few stock pickers, if any, have the skill needed to beat the market consistently, year after year. Professional investors, including fund managers, fail a basic test of skill: persistent achievement. The diagnostic for the existence of any skill is the consistency of individual differences in achievement. The logic is simple: if individual differences in any one year are due entirely to luck, the ranking of investors and funds will vary erratically and the year-to-year correlation will be zero. Where there is skill, however, the rankings will be more stable. The persistence of individual differences is the

measure by which we confirm the existence of skill among golfers, car salespeople, orthodontists, or speedy toll collectors on the turnpike.

Mutual funds are run by highly experienced and hardworking professionals who buy and sell stocks to achieve the best possible results for their clients. Nevertheless, the evidence from more than fifty years of research is conclusive: for a large majority of fund managers, the selection of stocks is more like rolling dice than like playing poker. Typically at least two out of every three mutual funds underperform the overall market in any given year.

More important, the year-to-year correlation between the outcomes of mutual funds is very small, barely higher than zero. The successful funds in any given year are mostly lucky; they have a good roll of the dice. There is general agreement among researchers that nearly all stock pickers, whether they know it or not—and few of them do—are playing a game of chance. The subjective experience of traders is that they are making sensible educated guesses in a situation of great uncertainty. In highly efficient markets, however, educated guesses are no more accurate than blind guesses.

Some years ago I had an unusual opportunity to examine the illusion of financial skill up close. I had been invited to speak to a group of investment advisers in a firm that provided financial advice and other services to very wealthy clients. I asked for some data to prepare my presentation and was granted a small treasure: a spreadsheet summarizing the investment outcomes of some twenty-five anonymous wealth advisers, for each of eight consecutive years. Each adviser's scoof [?]re for each year was his (most of them were men) main determinant of his year-end bonus. It was a simple matter to rank the advisers by their performance in each year and to determine whether there were persistent differences in skill among them and whether the same advisers consistently achieved better returns for their clients year after year.

To answer the question, I computed correlation coefficients between the rankings in each pair of years: year 1 with year 2, year 1 with year 3, and so on up through year 7 with year 8. That yielded 28 correlation coefficients, one for each pair of years. I knew the theory and was prepared to find weak evidence of persistence of skill. Still, I was surprised to find that the average of the 28 correlations was .01. In

other words, zero. The consistent correlations that would indicate differences in skill were not to be found. The results resembled what you would expect from a dice-rolling contest, not a game of skill.

No one in the firm seemed to be aware of the nature of the game that its stock pickers were playing. The advisers themselves felt they were competent professionals doing a serious job, and their superiors agreed. On the evening before the seminar, Richard Thaler and I had dinner with some of the top executives of the firm, the people who decide on the size of bonuses. We asked them to guess the year-to-year correlation in the rankings of individual advisers. They thought they knew what was coming and smiled as they said “not very high” or “performance certainly fluctuates.” It quickly became clear, however, that no one expected the average correlation to be zero.

Our message to the executives was that, at least when it came to building portfolios, the firm was rewarding luck as if it were skill. This should have been shocking news to them, but it was not. There was no sign that they disbelieved us. How could they? After all, we had analyzed their own results, and they were sophisticated enough to see the implications, which we politely refrained from spelling out. We all went on calmly with our dinner, and I have no doubt that both our findings and their implications were quickly swept under the rug and that life in the firm went on just as before. The illusion of skill is not only an individual aberration; it is deeply ingrained in the culture of the industry. Facts that challenge such basic assumptions—and thereby threaten people’s livelihood and self-esteem—are simply not absorbed. The mind does not digest them. This is particularly true of statistical studies of performance, which provide base-rate information that people generally ignore when it clashes with their personal impressions from experience.

The next morning, we reported the findings to the advisers, and their response was equally bland. Their own experience of exercising careful judgment on complex problems was far more compelling to them than an obscure statistical fact. When we were done, one of the executives I had dined with the previous evening drove me to the airport. He told me, with a trace of defensiveness, “I have done very well for the firm and no one can take that away from me.” I smiled and said nothing. But I thought, “Well, I took it away from you this morning. If your success was due mostly to chance, how much credit are you entitled to take for it?”

What Supports the Illusions of Skill and Validity?

Cognitive illusions can be more stubborn than visual illusions. What you learned about the Müller-Lyer illusion did not change the way you see the lines, but it changed your behavior. You now know that you cannot trust your impression of the lenglli?th of lines that have fins appended to them, and you also know that in the standard Müller-Lyer display you cannot trust what you see. When asked about the length of the lines, you will report your informed belief, not the illusion that you continue to see. In contrast, when my colleagues and I in the army learned that our leadership assessment tests had low validity, we accepted that fact intellectually, but it had no impact on either our feelings or our subsequent actions. The response we encountered in the financial firm was even more extreme. I am convinced that the message that Thaler and I delivered to both the executives and the portfolio managers was instantly put away in a dark corner of memory where it would cause no damage.

Why do investors, both amateur and professional, stubbornly believe that they can do better than the market, contrary to an economic theory that most of them accept, and contrary to what they could learn from a dispassionate evaluation of their personal experience? Many of the themes of previous chapters come up again in the explanation of the prevalence and persistence of an illusion of skill in the financial world.

The most potent psychological cause of the illusion is certainly that the people who pick stocks are exercising high-level skills. They consult economic data and forecasts, they examine income statements and balance sheets, they evaluate the quality of top management, and they assess the competition. All this is serious work that requires extensive training, and the people who do it have the immediate (and valid) experience of using these skills. Unfortunately, skill in evaluating the business prospects of a firm is not sufficient for successful stock trading, where the key question is whether the information about the firm is already incorporated in the price of its stock. Traders apparently lack the skill to answer this crucial question, but they appear to be ignorant of their ignorance. As I had discovered from watching cadets on the obstacle field, subjective confidence of traders is a feeling, not a

judgment. Our understanding of cognitive ease and associative coherence locates subjective confidence firmly in System 1.

Finally, the illusions of validity and skill are supported by a powerful professional culture. We know that people can maintain an unshakable faith in any proposition, however absurd, when they are sustained by a community of like-minded believers. Given the professional culture of the financial community, it is not surprising that large numbers of individuals in that world believe themselves to be among the chosen few who can do what they believe others cannot.

The Illusions of Pundits

The idea that the future is unpredictable is undermined every day by the ease with which the past is explained. As Nassim Taleb pointed out in *The Black Swan*, our tendency to construct and believe coherent narratives of the past makes it difficult for us to accept the limits of our forecasting ability. Everything makes sense in hindsight, a fact that financial pundits exploit every evening as they offer convincing accounts of the day's events. And we cannot suppress the powerful intuition that what makes sense in hindsight today was predictable yesterday. The illusion that we understand the past fosters overconfidence in our ability to predict the future.

The often-used image of the "march of history" implies order and direction. Marches, unlike strolls or walks, are not random. We think that we should be able to explain the past by focusing on either large social movements and cultural and technological developments or the intentions and abilities of a few g co?reat men. The idea that large historical events are determined by luck is profoundly shocking, although it is demonstrably true. It is hard to think of the history of the twentieth century, including its large social movements, without bringing in the role of Hitler, Stalin, and Mao Zedong. But there was a moment in time, just before an egg was fertilized, when there was a fifty-fifty chance that the embryo that became Hitler could have been a female. Compounding the three events, there was a probability of one-eighth of a twentieth century without any of the three great villains and it is impossible to argue that history would have been roughly the same in their absence. The fertilization of these three eggs had momentous consequences, and it makes a joke of the idea that long-term developments are predictable.

Yet the illusion of valid prediction remains intact, a fact that is exploited by people whose business is prediction—not only financial experts but pundits in business and politics, too. Television and radio stations and newspapers have their panels of experts whose job it is to comment on the recent past and foretell the future. Viewers and readers have the impression that they are receiving information that is somehow privileged, or at least extremely insightful. And there is no doubt that the pundits and their promoters genuinely believe they are offering such information. Philip Tetlock, a psychologist at the University of Pennsylvania, explained these so-called expert predictions in a landmark twenty-year study, which he published in his 2005 book *Expert Political Judgment: How Good Is It? How Can We Know?* Tetlock has set the terms for any future discussion of this topic.

Tetlock interviewed 284 people who made their living "commenting or offering advice on political and economic trends." He asked them to assess the probabilities that certain events would occur in the not too distant future, both in areas of the world in which they specialized and in regions about which they had less knowledge. Would Gorbachev be ousted in a coup? Would the United States go to war in the Persian Gulf? Which country would become the next big emerging market? In all, Tetlock gathered more than 80,000 predictions. He also asked the experts how they reached their conclusions, how they reacted when proved wrong, and how they evaluated evidence that did not support their positions. Respondents were asked to rate the probabilities of three alternative outcomes in every case: the persistence of the status quo, more of something such as political freedom or economic growth, or less of that thing.

The results were devastating. The experts performed worse than they would have if they had simply assigned equal probabilities to each of the three potential outcomes. In other words, people who spend their time, and earn their living, studying a particular topic produce poorer predictions than dart-throwing monkeys who would have distributed their choices evenly over the options. Even in the region they knew best, experts were not significantly better than nonspecialists.

Those who know more forecast very slightly better than those who know less. But those with the most knowledge are often less reliable. The reason is that the person who acquires more knowledge develops an enhanced illusion of her skill and becomes unrealistically overconfident. "We reach the point of diminishing marginal predictive

returns for knowledge disconcertingly quickly," Tetlock writes. "In this age of academic hyperspecialization, there is no reason for supposing that contributors to top journals—distinguished political scientists, area study specialists, economists, and so on—are any better than journalists or attentive readers of *The New York Times* in 'reading&#oul?8217; emerging situations." The more famous the forecaster, Tetlock discovered, the more flamboyant the forecasts. "Experts in demand," he writes, "were more overconfident than their colleagues who eked out existences far from the limelight."

Tetlock also found that experts resisted admitting that they had been wrong, and when they were compelled to admit error, they had a large collection of excuses: they had been wrong only in their timing, an unforeseeable event had intervened, or they had been wrong but for the right reasons. Experts are just human in the end. They are dazzled by their own brilliance and hate to be wrong. Experts are led astray not by what they believe, but by how they think, says Tetlock. He uses the terminology from Isaiah Berlin's essay on Tolstoy, "The Hedgehog and the Fox." Hedgehogs "know one big thing" and have a theory about the world; they account for particular events within a coherent framework, bristle with impatience toward those who don't see things their way, and are confident in their forecasts. They are also especially reluctant to admit error. For hedgehogs, a failed prediction is almost always "off only on timing" or "very nearly right." They are opinionated and clear, which is exactly what television producers love to see on programs. Two hedgehogs on different sides of an issue, each attacking the idiotic ideas of the adversary, make for a good show.

Foxes, by contrast, are complex thinkers. They don't believe that one big thing drives the march of history (for example, they are unlikely to accept the view that Ronald Reagan single-handedly ended the cold war by standing tall against the Soviet Union). Instead the foxes recognize that reality emerges from the interactions of many different agents and forces, including blind luck, often producing large and unpredictable outcomes. It was the foxes who scored best in Tetlock's study, although their performance was still very poor. They are less likely than hedgehogs to be invited to participate in television debates.

It is Not the Experts' Fault—The World is Difficult

The main point of this chapter is not that people who attempt to predict the future make many errors; that goes without saying. The first lesson is that errors of prediction are inevitable because the world is unpredictable. The second is that high subjective confidence is not to be trusted as an indicator of accuracy (low confidence could be more informative).

Short-term trends can be forecast, and behavior and achievements can be predicted with fair accuracy from previous behaviors and achievements. But we should not expect performance in officer training and in combat to be predictable from behavior on an obstacle field—behavior both on the test and in the real world is determined by many factors that are specific to the particular situation. Remove one highly assertive member from a group of eight candidates and everyone else's personalities will appear to change. Let a sniper's bullet move by a few centimeters and the performance of an officer will be transformed. I do not deny the validity of all tests—if a test predicts an important outcome with a validity of .20 or .30, the test should be used. But you should not expect more. You should expect little or nothing from Wall Street stock pickers who hope to be more accurate than the market in predicting the future of prices. And you should not expect much from pundits making long-term forecasts—although they may have valuable insights into the near future. The line that separates the possibly predictable future from the unpredictable distant future is in? yet to be drawn.

Speaking of Illusory Skill

> "He knows that the record indicates that the development of this illness is mostly unpredictable. How can he be so confident in this case? Sounds like an illusion of validity."

> "She has a coherent story that explains all she knows, and the coherence makes her feel good."

“What makes him believe that he is smarter than the market? Is this an illusion of skill?”

“She is a hedgehog. She has a theory that explains everything, and it gives her the illusion that she understands the world.”

“The question is not whether these experts are well trained. It is whether their world is predictable.”