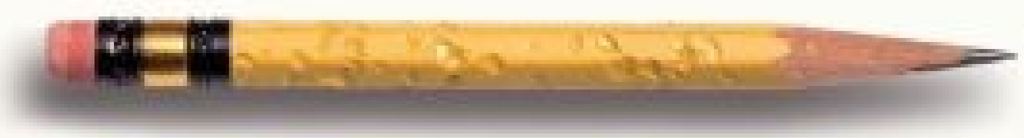
THINKING, FAST AND SLOW



DANIEL KAHNEMAN

WINNER OF THE NOBEL PRIZE IN ECONOMICS

THINKING, FAST AND SLOW

DANIEL KAHNEMAN

FARRAR, STRAUS AND GIROUX / NEW YORK

P

In memory of Amos Tversky

P

Contents

P

Introduction

Every author, I suppose, has in mind a setting in which readers of his or her work could benefit from having read it. Mine is the proverbial office watercooler, where opinions are shared and gossip is exchanged. I hope to enrich the vocabulary that people use when they talk about the judgments and choices of others, the company's new policies, or a colleague's investment decisions. Why be concerned with gossip? Because it is much easier, as well as far more enjoyable, to identify and label the mistakes of others than to recognize our own. Questioning what we believe and want is difficult at the best of times, and especially difficult when we most need to do it, but we can benefit from the informed opinions of others. Many of us spontaneously anticipate how friends and colleagues will evaluate our choices; the quality and content of these anticipated judgments therefore matters. The expectation of intelligent gossip is a powerful motive for serious self-criticism, more powerful than New Year resolutions to improve one's decision making at work and at home.

To be a good diagnostician, a physician needs to acquire a large set of labels for diseases, each of which binds an idea of the illness and its symptoms, possible antecedents and causes, possible developments and consequences, and possible interventions to cure or mitigate the illness. Learning medicine consists in part of learning the language of medicine. A deeper understanding of judgments and choices also requires a richer vocabulary than is available in everyday language. The hope for informed gossip is that there are distinctive patterns in the errors people make. Systematic errors are known as biases, and they recur predictably in particular circumstances. When the handsome and confident speaker bounds onto the stage, for example, you can anticipate that the audience will judge his comments more favorably than he deserves. The availability of a diagnostic label for this bias—the halo effect—makes it easier to anticipate, recognize, and understand.

When you are asked what you are thinking about, you can normally answer. You believe you know what goes on in your mind, which often consists of one conscious thought leading in an orderly way to another. But that is not the only way the mind works, nor indeed is that the typical way. Most impressions and thoughts arise in your conscious experience without your knowing how they got there. You cannot tracryd>e how you came to the belief that there is a lamp on the desk in front of you, or how you detected a hint of irritation in your spouse's voice on the telephone, or how you managed to avoid a threat on the road before you became consciously aware of it. The mental work that produces impressions, intuitions, and many decisions goes on in silence in our mind.

Much of the discussion in this book is about biases of intuition. However, the focus on error does not denigrate human intelligence, any more than the attention to diseases in medical texts denies good health. Most of us are healthy most of the time, and most of our judgments and actions are appropriate most of the time. As we navigate our lives, we normally allow ourselves to be guided by impressions and feelings, and the confidence we have in our intuitive beliefs and preferences is usually justified. But not always. We are often confident even when we are wrong, and an objective observer is more likely to detect our errors than we are.

So this is my aim for watercooler conversations: improve the ability to identify and understand errors of judgment and choice, in others and eventually in ourselves, by providing a richer and more precise language to discuss them. In at least some cases, an accurate diagnosis may suggest an intervention to limit the damage that bad judgments and choices often cause.

Origins

This book presents my current understanding of judgment and decision making, which has been shaped by psychological discoveries of recent decades. However, I trace the central ideas to the lucky day in 1969 when I asked a colleague to speak as a guest to a seminar I was teaching in the Department of Psychology at the Hebrew University of Jerusalem. Amos Tversky was considered a rising star in the field of decision research—indeed, in anything he did—so I knew we would have an interesting time. Many people who knew Amos thought he was the most intelligent person they had ever met. He was brilliant, voluble, and charismatic. He was also blessed with a perfect memory for jokes and an exceptional ability to use them to make a point. There was never a dull moment when Amos was around. He was then thirty-two; I was thirty-five.

Amos told the class about an ongoing program of research at the University of Michigan that sought to answer this question: Are people good intuitive statisticians? We already knew that people are good intuitive grammarians: at age four a child effortlessly conforms to the rules of grammar as she speaks, although she has no idea that such rules exist. Do people have a similar intuitive feel for the basic principles of statistics? Amos reported that the answer was a qualified yes. We had a lively debate in the seminar and ultimately concluded that a qualified no was a better answer.

Amos and I enjoyed the exchange and concluded that intuitive statistics was an interesting topic and that it would be fun to explore it together. That Friday we met for lunch at Café Rimon, the favorite hangout of bohemians and professors in Jerusalem, and planned a study of the statistical intuitions of sophisticated researchers. We had concluded in the seminar that our own intuitions were deficient. In spite of years of teaching and using statistics, we had not developed an intuitive sense of the reliability of statistical results observed in small samples. Our subjective judgments were biased: we were far too willing to believe research findings based on inadequate evidence and prone to collect too few observations in our own research. The goal of our study was to examine whether other researchers suffered from the same affliction.

We prepared a survey that included realistic scenarios of statistical issues that arise in research. Amos collected the responses of a group of expert participants in a meeting of the Society of Mathematical Psychology, including the authors of two statistical textbooks. As expected, we found that our expert colleagues, like us, greatly exaggerated the likelihood that the original result of an experiment would be successfully replicated even with a small sample. They also gave very poor advice to a fictitious graduate student about the number of observations she needed to collect. Even statisticians were not good intuitive statisticians.

While writing the article that reported these findings, Amos and I discovered that we enjoyed working together. Amos was always very funny, and in his presence I became funny as well, so we spent hours of solid work in continuous amusement. The pleasure we found in working together made us exceptionally patient; it is much easier to strive for perfection when you are never bored. Perhaps most important, we checked our critical weapons at the door. Both Amos and I were critical and argumentative, he even more than I, but during the years of our collaboration neither of us ever rejected out of hand anything the other said. Indeed, one of the great joys I found in the collaboration was that Amos frequently saw the point of my vague ideas much more clearly than I did. Amos was the more logical thinker, with an orientation to theory and an unfailing sense of direction. I was more intuitive and rooted in the psychology of perception, from which we borrowed many ideas. We were sufficiently similar to understand each other easily, and sufficiently different to surprise each other. We developed a routine in which we spent much of our working days together, often on long walks. For the next fourteen years our collaboration was the focus of our lives, and the work we did together during those years was the best either of us ever did.

We quickly adopted a practice that we maintained for many years. Our research was a conversation, in which we invented questions and jointly examined our intuitive answers. Each question was a small experiment, and we carried out many experiments in a single day. We were not seriously looking for the correct answer to the statistical questions we posed. Our aim was to identify and analyze the intuitive answer, the first one that came to mind, the one we were tempted to make even when we knew it to be wrong. We believed —correctly, as it happened—that any intuition that the two of us shared would be shared by many other people as well, and that it would be easy to demonstrate its effects on judgments.

We once discovered with great delight that we had identical silly ideas about the future professions of several toddlers we both knew. We could identify the argumentative three-year-old lawyer, the nerdy professor, the empathetic and mildly intrusive psychotherapist. Of course these predictions were absurd, but we still found them appealing. It was also clear that our intuitions were governed by the resemblance of each child to the cultural stereotype of a profession. The amusing exercise helped us develop a theory that was emerging in our minds at the time, about the role of resemblance in predictions. We went on to test and elaborate that theory in dozens of experiments, as in the following example.

As you consider the next question, please assume that Steve was selected at random from a representative sample:

An individual has been described by a neighbor as follows: “Steve is very shy and

> withdrawn, invariably helpful but with little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structurut and stre, and a passion for detail." Is Steve more likely to be a librarian or a farmer?

The resemblance of Steve's personality to that of a stereotypical librarian strikes everyone immediately, but equally relevant statistical considerations are almost always ignored. Did it occur to you that there are more than 20 male farmers for each male librarian in the United States? Because there are so many more farmers, it is almost certain that more "meek and tidy" souls will be found on tractors than at library information desks. However, we found that participants in our experiments ignored the relevant statistical facts and relied exclusively on resemblance. We proposed that they used resemblance as a simplifying heuristic (roughly, a rule of thumb) to make a difficult judgment. The reliance on the heuristic caused predictable biases (systematic errors) in their predictions.

On another occasion, Amos and I wondered about the rate of divorce among professors in our university. We noticed that the question triggered a search of memory for divorced professors we knew or knew about, and that we judged the size of categories by the ease with which instances came to mind. We called this reliance on the ease of memory search the availability heuristic. In one of our studies, we asked participants to answer a simple question about words in a typical English text:

> Consider the letter *K*.
>
> Is *K* more likely to appear as the first letter in a word OR as the third letter?

As any Scrabble player knows, it is much easier to come up with words that begin with a particular letter than to find words that have the same letter in the third position. This is true for every letter of the alphabet. We therefore expected respondents to exaggerate the frequency of letters appearing in the first position—even those letters (such as *K*, *L*, *N*, *R*, *V*) which in fact occur more frequently in the third position. Here again, the reliance on a heuristic produces a predictable bias in judgments. For example, I recently came to doubt my long-held impression that adultery is more common among politicians than among physicians or lawyers. I had even come up with explanations for that "fact," including the aphrodisiac effect of power and the temptations of life away from home. I eventually realized that the transgressions of politicians are much more likely to be reported than the transgressions of lawyers and doctors. My intuitive impression could be due entirely to journalists' choices of topics and to my reliance on the availability heuristic.

Amos and I spent several years studying and documenting biases of intuitive thinking in various tasks—assigning probabilities to events, forecasting the future, assessing hypotheses, and estimating frequencies. In the fifth year of our collaboration, we presented our main findings in *Science* magazine, a publication read by scholars in many disciplines. The article (which is reproduced in full at the end of this book) was titled "Judgment Under Uncertainty: Heuristics and Biases." It described the simplifying shortcuts of intuitive thinking and explained some 20 biases as manifestations of these heuristics—and also as demonstrations of the role of heuristics in judgment.

Historians of science have often noted that at any given time scholars in a particular field tend to share basic re share assumptions about their subject. Social scientists are no exception; they rely on a view of human nature that provides the background of most discussions of specific behaviors but is rarely questioned. Social scientists in the 1970s broadly accepted two ideas about human nature. First, people are generally rational, and their thinking is normally sound. Second, emotions such as fear, affection, and hatred explain most of the occasions on which people depart from rationality. Our article challenged both assumptions without discussing them directly. We documented systematic errors in the thinking of normal people, and we traced these errors to the design of the machinery of cognition rather than to the corruption of thought by emotion.

Our article attracted much more attention than we had expected, and it remains one of the most highly cited works in social science (more than three hundred scholarly articles referred to it in 2010). Scholars in other disciplines found it useful, and the ideas of heuristics and biases have been used productively in many fields, including medical diagnosis, legal judgment, intelligence analysis, philosophy, finance, statistics, and military strategy.

For example, students of policy have noted that the availability heuristic helps explain why some issues are highly salient in the public's mind while others are neglected. People tend to assess the relative importance of issues by the ease with which they are retrieved from memory—and this is largely determined by the extent of coverage in the media. Frequently mentioned topics populate the mind even as others slip away from awareness. In turn, what the media choose to report corresponds to their view of what is currently on the public's mind. It is no accident that authoritarian regimes exert substantial pressure on independent media. Because public interest is most easily aroused by dramatic events and by celebrities, media feeding frenzies are common. For several weeks after Michael Jackson's death, for example, it was virtually impossible to find a television channel reporting on another topic. In contrast, there is little coverage of critical but unexciting issues that provide less drama, such as declining educational standards or overinvestment of medical resources in the last year of life. (As I write this, I notice that my choice of "little-covered" examples was guided by availability. The topics I chose as examples are mentioned often; equally important issues that are less available did not come to my mind.)

We did not fully realize it at the time, but a key reason for the broad appeal of "heuristics and biases" outside psychology was an incidental feature of our work: we almost always included in our articles the full text of the questions we had asked ourselves and our respondents. These questions served as demonstrations for the reader, allowing him to recognize how his own thinking was tripped up by cognitive biases. I hope you had such an experience as you read the question about Steve the librarian, which was intended to help you appreciate the power of resemblance as a cue to probability and to see how easy it is to ignore relevant statistical facts.

The use of demonstrations provided scholars from diverse disciplines—notably philosophers and economists—an unusual opportunity to observe possible flaws in their own thinking. Having seen themselves fail, they became more likely to question the dogmatic assumption, prevalent at the time, that the human mind is rational and logical. The choice of method was crucial: if we had reported results of only conventional experiments, the article would have been less noteworthy and less memorable. Furthermore, skeptical readers would have distanced themselves from the results by attributing judgment errors to the familiar l the famifecklessness of undergraduates, the typical participants in psychological studies. Of course, we did not choose demonstrations over standard experiments because we wanted to influence philosophers and economists. We preferred demonstrations because they were more fun, and we were lucky in our choice of method as well as in many other ways. A recurrent theme of this book is that luck plays a large role in every story of success; it is almost always easy to identify a small change in the story that would have turned a remarkable achievement into a mediocre outcome. Our story was no exception.

The reaction to our work was not uniformly positive. In particular, our focus on biases was criticized as suggesting an unfairly negative view of the mind. As expected in normal science, some investigators refined our ideas and others offered plausible alternatives. By and large, though, the idea that our minds are susceptible to systematic errors is now generally accepted. Our research on judgment had far more effect on social science than we thought possible when we were working on it.

Immediately after completing our review of judgment, we switched our attention to decision making under uncertainty. Our goal was to develop a psychological theory of how people make decisions about simple gambles. For example: Would you accept a bet on the toss of a coin where you win $130 if the coin shows heads and lose $100 if it shows tails? These elementary choices had long been used to examine broad questions about decision making, such as the relative weight that people assign to sure things and to uncertain outcomes. Our method did not change: we spent many days making up choice problems and examining whether our intuitive preferences conformed to the logic of choice. Here again, as in judgment, we observed systematic biases in our own decisions, intuitive preferences that consistently violated the rules of rational choice. Five years after the *Science* article, we published “Prospect Theory: An Analysis of Decision Under Risk,” a theory of choice that is by some counts more influential than our work on judgment, and is one of the foundations of behavioral economics.

Until geographical separation made it too difficult to go on, Amos and I enjoyed the extraordinary good fortune of a shared mind that was superior to our individual minds and of a relationship that made our work fun as well as productive. Our collaboration on judgment and decision making was the reason for the Nobel Prize that I received in 2002, which Amos would have shared had he not died, aged fifty-nine, in 1996.

Where we are now

This book is not intended as an exposition of the early research that Amos and I conducted together, a task that has been ably carried out by many authors over the years. My main aim here is to present a view of how the mind works that draws on recent developments in cognitive and social psychology. One of the more important developments is that we now understand the marvels as well as the flaws of intuitive thought.

Amos and I did not address accurate intuitions beyond the casual statement that judgment heuristics “are quite useful, but sometimes lead to severe and systematic errors.” We focused on biases, both because we found them interesting in their own right and because they provided evidence for the heuristics of judgment. We did not ask ourselves whether all intuitive judgments under uncertainty are produced by the heuristics we studied; it is now clear that they are not. In particular, the accurate intuitions of experts are better explained by the effects of prolonged practice than by heuristics. We can now draw a richer andigha riche more balanced picture, in which skill and heuristics are alternative sources of intuitive judgments and choices.

The psychologist Gary Klein tells the story of a team of firefighters that entered a house in which the kitchen was on fire. Soon after they started hosing down the kitchen, the commander heard himself shout, “Let’s get out of here!” without realizing why. The floor collapsed almost immediately after the firefighters escaped. Only after the fact did the commander realize that the fire had been unusually quiet and that his ears had been unusually hot. Together, these impressions prompted what he called a “sixth sense of danger.” He had no idea what was wrong, but he knew something was wrong. It turned out that the heart of the fire had not been in the kitchen but in the basement beneath where the men had stood.

We have all heard such stories of expert intuition: the chess master who walks past a street game and announces “White mates in three” without stopping, or the physician who makes a complex diagnosis after a single glance at a patient. Expert intuition strikes us as magical, but it is not. Indeed, each of us performs feats of intuitive expertise many times each day. Most of us are pitch-perfect in detecting anger in the first word of a telephone call, recognize as we enter a room that we were the subject of the conversation, and quickly react to subtle signs that the driver of the car in the next lane is dangerous. Our everyday intuitive abilities are no less marvelous than the striking insights of an experienced firefighter or physician—only more common.

The psychology of accurate intuition involves no magic. Perhaps the best short statement of it is by the great Herbert Simon, who studied chess masters and showed that after thousands of hours of practice they come to see the pieces on the board differently from the rest of us. You can feel Simon’s impatience with the mythologizing of expert intuition when he writes: “The situation has provided a cue; this cue has given the expert access to information stored in memory, and the information provides the answer. Intuition is nothing more and nothing less than recognition.”

We are not surprised when a two-year-old looks at a dog and says “doggie!” because we are used to the miracle of children learning to recognize and name things. Simon’s point is that the miracles of expert intuition have the same character. Valid intuitions develop when experts have learned to recognize familiar elements in a new situation and to act in a manner that is appropriate to it. Good intuitive judgments come to mind with the same immediacy as “doggie!”

Unfortunately, professionals’ intuitions do not all arise from true expertise. Many years ago I visited the chief investment officer of a large financial firm, who told me that he had just invested some tens of millions of dollars in the stock of Ford Motor Company. When I asked how he had made that decision, he replied that he had recently attended an automobile show and had been impressed. “Boy, do they know how to make a car!” was his explanation. He made it very clear that he trusted his gut feeling and was satisfied with himself and with his decision. I found it remarkable that he had apparently not considered the one question that an economist would call relevant: Is Ford stock currently underpriced? Instead, he had listened to his intuition; he liked the cars, he liked the company, and he liked the idea of owning its stock. From what we know about the accuracy of stock picking, it is reasonable to believe that he did not know what he was doing.

The specific heuristics that Amos and I studied proviheitudied de little help in understanding how the executive came to invest in Ford stock, but a broader conception of heuristics now exists, which offers a good account. An important advance is that emotion now looms much larger in our understanding of intuitive judgments and choices than it did in the past. The executive’s decision would today be described as an example of the affect heuristic, where judgments and decisions are guided directly by feelings of liking and disliking, with little deliberation or reasoning.

When confronted with a problem—choosing a chess move or deciding whether to invest in a stock—the machinery of intuitive thought does the best it can. If the individual has relevant expertise, she will recognize the situation, and the intuitive solution that comes to her mind is likely to be correct. This is what happens when a chess master looks at a complex position: the few moves that immediately occur to him are all strong. When the question is difficult and a skilled solution is not available, intuition still has a shot: an answer may come to mind quickly—but it is not an answer to the original question. The question that the executive faced (should I invest in Ford stock?) was difficult, but the answer to an easier and related question (do I like Ford cars?) came readily to his mind and determined his choice. This is the essence of intuitive heuristics: when faced with a difficult question, we often answer an easier one instead, usually without noticing the substitution.

The spontaneous search for an intuitive solution sometimes fails—neither an expert solution nor a heuristic answer comes to mind. In such cases we often find ourselves switching to a slower, more deliberate and effortful form of thinking. This is the slow thinking of the title. Fast thinking includes both variants of intuitive thought—the expert and the heuristic—as well as the entirely automatic mental activities of perception and memory, the operations that enable you to know there is a lamp on your desk or retrieve the name of the capital of Russia.

The distinction between fast and slow thinking has been explored by many psychologists over the last twenty-five years. For reasons that I explain more fully in the next chapter, I describe mental life by the metaphor of two agents, called System 1 and System 2, which respectively produce fast and slow thinking. I speak of the features of intuitive and deliberate thought as if they were traits and dispositions of two characters in your mind. In the picture that emerges from recent research, the intuitive System 1 is more influential than your experience tells you, and it is the secret author of many of the choices and judgments you make. Most of this book is about the workings of System 1 and the mutual influences between it and System 2.

What Comes Next

The book is divided into five parts. Part 1 presents the basic elements of a two-systems approach to judgment and choice. It elaborates the distinction between the automatic operations of System 1 and the controlled operations of System 2, and shows how associative memory, the core of System 1, continually constructs a coherent interpretation of what is going on in our world at any instant. I attempt to give a sense of the complexity and richness of the automatic and often unconscious processes that underlie intuitive thinking, and of how these automatic processes explain the heuristics of judgment. A goal is to introduce a language for thinking and talking about the mind.

Part 2 updates the study of judgment heuristics and explores a major puzzle: Why is it so difficult for us to think statistically? We easily think associativelm 1assocciay, we think metaphorically, we think causally, but statistics requires thinking about many things at once, which is something that System 1 is not designed to do.

The difficulties of statistical thinking contribute to the main theme of Part 3, which describes a puzzling limitation of our mind: our excessive confidence in what we believe we know, and our apparent inability to acknowledge the full extent of our ignorance and the uncertainty of the world we live in. We are prone to overestimate how much we understand about the world and to underestimate the role of chance in events. Overconfidence is fed by the illusory certainty of hindsight. My views on this topic have been influenced by Nassim Taleb, the author of *The Black Swan*. I hope for watercooler conversations that intelligently explore the lessons that can be learned from the past while resisting the lure of hindsight and the illusion of certainty.

The focus of part 4 is a conversation with the discipline of economics on the nature of decision making and on the assumption that economic agents are rational. This section of the book provides a current view, informed by the two-system model, of the key concepts of prospect theory, the model of choice that Amos and I published in 1979. Subsequent chapters address several ways human choices deviate from the rules of rationality. I deal with the unfortunate tendency to treat problems in isolation, and with framing effects, where decisions are shaped by inconsequential features of choice problems. These observations, which are readily explained by the features of System 1, present a deep challenge to the rationality assumption favored in standard economics.

Part 5 describes recent research that has introduced a distinction between two selves, the experiencing self and the remembering self, which do not have the same interests. For example, we can expose people to two painful experiences. One of these experiences is strictly worse than the other, because it is longer. But the automatic formation of memories —a feature of System 1—has its rules, which we can exploit so that the worse episode leaves a better memory. When people later choose which episode to repeat, they are, naturally, guided by their remembering self and expose themselves (their experiencing self) to unnecessary pain. The distinction between two selves is applied to the measurement of well-being, where we find again that what makes the experiencing self happy is not quite the same as what satisfies the remembering self. How two selves within a single body can pursue happiness raises some difficult questions, both for individuals and for societies that view the well-being of the population as a policy objective.

A concluding chapter explores, in reverse order, the implications of three distinctions drawn in the book: between the experiencing and the remembering selves, between the conception of agents in classical economics and in behavioral economics (which borrows from psychology), and between the automatic System 1 and the effortful System 2. I return to the virtues of educating gossip and to what organizations might do to improve the quality of judgments and decisions that are made on their behalf.

Two articles I wrote with Amos are reproduced as appendixes to the book. The first is the review of judgment under uncertainty that I described earlier. The second, published in 1984, summarizes prospect theory as well as our studies of framing effects. The articles present the contributions that were cited by the Nobel committee—and you may be surprised by how simple they are. Reading them will give you a sense of how much we knew a long time ago, and also of how much we have learned in recent decades.

P

Part 1

P

Two Systems

P

The Characters of the Story

To observe your mind in automatic mode, glance at the image below.

Figure 1

Your experience as you look at the woman's face seamlessly combines what we normally call seeing and intuitive thinking. As surely and quickly as you saw that the young woman's hair is dark, you knew she is angry. Furthermore, what you saw extended into the future. You sensed that this woman is about to say some very unkind words, probably in a loud and strident voice. A premonition of what she was going to do next came to mind automatically and effortlessly. You did not intend to assess her mood or to anticipate what she might do, and your reaction to the picture did not have the feel of something you did. It just happened to you. It was an instance of fast thinking.

Now look at the following problem:

$$17 \times 24$$

You knew immediately that this is a multiplication problem, and probably knew that you could solve it, with paper and pencil, if not without. You also had some vague intuitive knowledge of the range of possible results. You would be quick to recognize that both 12,609 and 123 are implausible. Without spending some time on the problem, however, you would not be certain that the answer is not 568. A precise solution did not come to mind, and you felt that you could choose whether or not to engage in the computation. If you have not done so yet, you should attempt the multiplication problem now, completing at least part of it.

You experienced slow thinking as you proceeded through a sequence of steps. You first retrieved from memory the cognitive program for multiplication that you learned in school, then you implemented it. Carrying out the computation was a strain. You felt the burden of holding much material in memory, as you needed to keep track of where you were and of where you were going, while holding on to the intermediate result. The process was mental work: deliberate, effortful, and orderly—a prototype of slow thinking. The computation was not only an event in your mind; your body was also involved. Your muscles tensed up, your blood pressure rose, and your heart rate increased. Someone looking closely at your eyes while you tackled this problem would have seen your pupils dilate. Your pupils contracted back to normal size as soon as you ended your work—when you found the answer (which is 408, by the way) or when you gave up.

Two Systems

Psychologists have been intensely interested for several decades in the two modagee fi Pn="cees of thinking evoked by the picture of the angry woman and by the multiplication problem, and have offered many labels for them. I adopt terms originally proposed by the psychologists Keith Stanovich and Richard West, and will refer to two systems in the mind, System 1 and System 2.

- *System 1* operates automatically and quickly, with little or no effort and no sense of voluntary control.
- *System 2* allocates attention to the effortful mental activities that demand it, including complex computations. The operations of System 2 are often associated with the subjective experience of agency, choice, and concentration.

The labels of System 1 and System 2 are widely used in psychology, but I go further than most in this book, which you can read as a psychodrama with two characters.

When we think of ourselves, we identify with System 2, the conscious, reasoning self that has beliefs, makes choices, and decides what to think about and what to do. Although System 2 believes itself to be where the action is, the automatic System 1 is the hero of the book. I describe System 1 as effortlessly originating impressions and feelings that are the main sources of the explicit beliefs and deliberate choices of System 2. The automatic operations of System 1 generate surprisingly complex patterns of ideas, but only the slower System 2 can construct thoughts in an orderly series of steps. I also describe circumstances in which System 2 takes over, overruling the freewheeling impulses and associations of System 1. You will be invited to think of the two systems as agents with their individual abilities, limitations, and functions.

In rough order of complexity, here are some examples of the automatic activities that are attributed to System 1:

- Detect that one object is more distant than another.
- Orient to the source of a sudden sound.
- Complete the phrase “bread and…”
- Make a “disgust face” when shown a horrible picture.
- Detect hostility in a voice.
- Answer to 2 + 2 = ?
- Read words on large billboards.
- Drive a car on an empty road.
- Find a strong move in chess (if you are a chess master).
- Understand simple sentences.
- Recognize that a “meek and tidy soul with a passion for detail” resembles an occupational stereotype.

All these mental events belong with the angry woman—they occur automatically and require little or no effort. The capabilities of System 1 include innate skills that we share with other animals. We are born prepared to perceive the world around us, recognize objects, orient attention, avoid losses, and fear spiders. Other mental activities become fast and automatic through prolonged practice. System 1 has learned associations between ideas (the capital of France?); it has also learned skills such as reading and understanding nuances of social situations. Some skills, such as finding strong chess moves, are acquired only by specialized experts. Others are widely shared. Detecting the similarity of a personality sketch to an occupatiohein occupatnal stereotype requires broad knowledge of the language and the culture, which most of us possess. The knowledge is stored in memory and accessed without intention and without effort.

Several of the mental actions in the list are completely involuntary. You cannot refrain from understanding simple sentences in your own language or from orienting to a loud unexpected sound, nor can you prevent yourself from knowing that 2 + 2 = 4 or from thinking of Paris when the capital of France is mentioned. Other activities, such as chewing, are susceptible to voluntary control but normally run on automatic pilot. The control of attention is shared by the two systems. Orienting to a loud sound is normally an involuntary operation of System 1, which immediately mobilizes the voluntary attention of System 2. You may be able to resist turning toward the source of a loud and offensive comment at a crowded party, but even if your head does not move, your attention is initially directed to it, at least for a while. However, attention can be moved away from an unwanted focus, primarily by focusing intently on another target.

The highly diverse operations of System 2 have one feature in common: they require attention and are disrupted when attention is drawn away. Here are some examples:

- Brace for the starter gun in a race.
- Focus attention on the clowns in the circus.
- Focus on the voice of a particular person in a crowded and noisy room.
- Look for a woman with white hair.
- Search memory to identify a surprising sound.
- Maintain a faster walking speed than is natural for you.
- Monitor the appropriateness of your behavior in a social situation.
- Count the occurrences of the letter *a* in a page of text.
- Tell someone your phone number.
- Park in a narrow space (for most people except garage attendants).
- Compare two washing machines for overall value.
- Fill out a tax form.
- Check the validity of a complex logical argument.

In all these situations you must pay attention, and you will perform less well, or not at all, if you are not ready or if your attention is directed inappropriately. System 2 has some ability to change the way System 1 works, by programming the normally automatic functions of attention and memory. When waiting for a relative at a busy train station, for example, you can set yourself at will to look for a white-haired woman or a bearded man, and thereby increase the likelihood of detecting your relative from a distance. You can set your memory to search for capital cities that start with *N* or for French existentialist novels. And when you rent a car at London’s Heathrow Airport, the attendant will probably remind you that “we drive on the left side of the road over here.” In all these cases, you are asked to do something that does not come naturally, and you will find that the consistent maintenance of a set requires continuous exertion of at least some effort.

The often-used phrase “pay attention” is apt: you dispose of a limited budget of attention that you can allocate to activities, and if you try to i>Cyou try tgo beyond your budget, you will fail. It is the mark of effortful activities that they interfere with each other, which is why it is difficult or impossible to conduct several at once. You could not compute the product of 17 × 24 while making a left turn into dense traffic, and you certainly should not try. You can do several things at once, but only if they are easy and undemanding. You are probably safe carrying on a conversation with a passenger while driving on an empty highway, and many parents have discovered, perhaps with some guilt, that they can read a story to a child while thinking of something else.

Everyone has some awareness of the limited capacity of attention, and our social behavior makes allowances for these limitations. When the driver of a car is overtaking a truck on a narrow road, for example, adult passengers quite sensibly stop talking. They know that distracting the driver is not a good idea, and they also suspect that he is temporarily deaf and will not hear what they say.

Intense focusing on a task can make people effectively blind, even to stimuli that normally attract attention. The most dramatic demonstration was offered by Christopher Chabris and Daniel Simons in their book *The Invisible Gorilla*. They constructed a short film of two teams passing basketballs, one team wearing white shirts, the other wearing black. The viewers of the film are instructed to count the number of passes made by the white team, ignoring the black players. This task is difficult and completely absorbing. Halfway through the video, a woman wearing a gorilla suit appears, crosses the court, thumps her chest, and moves on. The gorilla is in view for 9 seconds. Many thousands of people have seen the video, and about half of them do not notice anything unusual. It is the counting task—and especially the instruction to ignore one of the teams—that causes the blindness. No one who watches the video without that task would miss the gorilla. Seeing and orienting are automatic functions of System 1, but they depend on the allocation of some attention to the relevant stimulus. The authors note that the most remarkable observation of their study is that people find its results very surprising. Indeed, the viewers who fail to see the gorilla are initially sure that it was not there—they cannot imagine missing such a striking event. The gorilla study illustrates two important facts about our minds: we can be blind to the obvious, and we are also blind to our blindness.

Plot Synopsis

The interaction of the two systems is a recurrent theme of the book, and a brief synopsis of the plot is in order. In the story I will tell, Systems 1 and 2 are both active whenever we are awake. System 1 runs automatically and System 2 is normally in a comfortable low-effort mode, in which only a fraction of its capacity is engaged. System 1 continuously generates suggestions for System 2: impressions, intuitions, intentions, and feelings. If endorsed by System 2, impressions and intuitions turn into beliefs, and impulses turn into voluntary actions. When all goes smoothly, which is most of the time, System 2 adopts the suggestions of System 1 with little or no modification. You generally believe your impressions and act on your desires, and that is fine—usually.

When System 1 runs into difficulty, it calls on System 2 to support more detailed and specific processing that may solve the problem of the moment. System 2 is mobilized when a question arises for which System 1 does not offer an answer, as probably happened to you when you encountered the multiplication problem 17 × 24. You can also feel a surge of conscious attention whenever you are surprised. System 2 is activ”>< 2 is actated when an event is detected that violates the model of the world that System 1 maintains. In that world, lamps do not jump, cats do not bark, and gorillas do not cross basketball courts. The gorilla experiment demonstrates that some attention is needed for the surprising stimulus to be detected. Surprise then activates and orients your attention: you will stare, and you will search your memory for a story that makes sense of the surprising event. System 2 is also credited with the continuous monitoring of your own behavior—

the control that keeps you polite when you are angry, and alert when you are driving at night. System 2 is mobilized to increased effort when it detects an error about to be made. Remember a time when you almost blurted out an offensive remark and note how hard you worked to restore control. In summary, most of what you (your System 2) think and do originates in your System 1, but System 2 takes over when things get difficult, and it normally has the last word.

The division of labor between System 1 and System 2 is highly efficient: it minimizes effort and optimizes performance. The arrangement works well most of the time because System 1 is generally very good at what it does: its models of familiar situations are accurate, its short-term predictions are usually accurate as well, and its initial reactions to challenges are swift and generally appropriate. System 1 has biases, however, systematic errors that it is prone to make in specified circumstances. As we shall see, it sometimes answers easier questions than the one it was asked, and it has little understanding of logic and statistics. One further limitation of System 1 is that it cannot be turned off. If you are shown a word on the screen in a language you know, you will read it—unless your attention is totally focused elsewhere.

Conflict

Figure 2 is a variant of a classic experiment that produces a conflict between the two systems. You should try the exercise before reading on.

Your first task is to go down both columns, calling out whether each word is printed in lowercase or in uppercase. When you are done with the first task, go down both columns again, saying whether each word is printed to the left or to the right of center by saying (or whispering to yourself) “LEFT” or “RIGHT.”

LEFT		upper	
	left	lower	
right			LOWER
RIGHT		upper	
	RIGHT	UPPER	
	left		lower
LEFT			LOWER
	right		upper

Figure 2

You were almost certainly successful in saying the correct words in both tasks, and you surely discovered that some parts of each task were much easier than others. When you identified upper- and lowercase, the left-hand column was easy and the right-hand column caused you to slow down and perhaps to stammer or stumble. When you named the position of words, the left-hand column was difficult and the right-hand column was much easier.

These tasks engage System 2, because saying “upper/lower” or “right/left” is not what you routinely do when looking down a column of words. One of the things you did to set yourself for the task was to program your memory so that the relevant words (*upper* and *lower* for the first task) were “on the tip of your tongue.” The prioritizing of the chosen words is effective and the mild temptation to read other words was fairly easy to resist when you went through the first column. But the second column was different, because it contained words for which you were set, and you could not ignore them. You were mostly able to respond correctly, but overcoming the competing response was a strain, and it slowed you down. You experienced a conflict between a task that you intended to carry out and an automatic response that interfered with it.

Conflict between an automatic reaction and an intention to conWhetion to ctrol it is common in our lives. We are all familiar with the experience of trying not to stare at the oddly dressed couple at the neighboring table in a restaurant. We also know what it is like to force our attention on a boring book, when we constantly find ourselves returning to the point at which the reading lost its meaning. Where winters are hard, many drivers have memories of their car skidding out of control on the ice and of the struggle to follow well-rehearsed instructions that negate what they would naturally do: “Steer into the skid, and whatever you do, do not touch the brakes!” And every human being has had the experience of *not* telling someone to go to hell. One of the tasks of System 2 is to overcome the impulses of System 1. In other words, System 2 is in charge of self-control.

Illusions

To appreciate the autonomy of System 1, as well as the distinction between impressions and beliefs, take a good look at figure 3.

This picture is unremarkable: two horizontal lines of different lengths, with fins appended, pointing in different directions. The bottom line is obviously longer than the one above it. That is what we all see, and we naturally believe what we see. If you have already encountered this image, however, you recognize it as the famous Müller-Lyer illusion. As you can easily confirm by measuring them with a ruler, the horizontal lines are in fact identical in length.

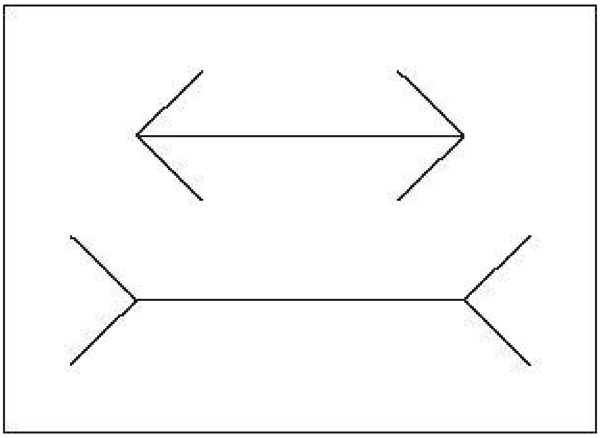


Figure 3

Now that you have measured the lines, you—your System 2, the conscious being you call "I"—have a new belief: you *know* that the lines are equally long. If asked about their length, you will say what you know. But you still *see* the bottom line as longer. You have chosen to believe the measurement, but you cannot prevent System 1 from doing its thing; you cannot decide to see the lines as equal, although you know they are. To resist the illusion, there is only one thing you can do: you must learn to mistrust your impressions of the length of lines when fins are attached to them. To implement that rule, you must be able to recognize the illusory pattern and recall what you know about it. If you can do this, you will never again be fooled by the Müller-Lyer illusion. But you will still see one line as longer than the other.

Not all illusions are visual. There are illusions of thought, which we call *cognitive illusions*. As a graduate student, I attended some courses on the art and science of psychotherapy. During one of these lectures, our teacher imparted a morsel of clinical wisdom. This is what he told us: "You will from time to time meet a patient who shares a disturbing tale of multiple mistakes in his previous treatment. He has been seen by several clinicians, and all failed him. The patient can lucidly describe how his therapists misunderstood him, but he has quickly perceived that you are different. You share the same feeling, are convinced that you understand him, and will be able to help." At this point my teacher raised his voice as he said, "Do not even *think* of taking on this patient! Throw him out of the office! He is most likely a psychopath and you will not be able to help him."

Many years later I learned that the teacher had warned us against psychopathic charm, and the leading authority in the strn y in the udy of psychopathy confirmed that the teacher's advice was sound. The analogy to the Müller-Lyer illusion is close. What we were being taught was not how to feel about that patient. Our teacher took it for granted that the sympathy we would feel for the patient would not be under our control; it would arise from System 1. Furthermore, we were not being taught to be generally suspicious of our feelings about patients. We were told that a strong attraction to a patient with a repeated history of failed treatment is a danger sign—like the fins on the parallel lines. It is an illusion—a cognitive illusion—and I (System 2) was taught how to recognize it and advised not to believe it or act on it.

The question that is most often asked about cognitive illusions is whether they can be overcome. The message of these examples is not encouraging. Because System 1 operates automatically and cannot be turned off at will, errors of intuitive thought are often difficult to prevent. Biases cannot always be avoided, because System 2 may have no clue to the error. Even when cues to likely errors are available, errors can be prevented only by the enhanced monitoring and effortful activity of System 2. As a way to live your life, however, continuous vigilance is not necessarily good, and it is certainly impractical. Constantly questioning our own thinking would be impossibly tedious, and System 2 is much too slow and inefficient to serve as a substitute for System 1 in making routine decisions. The best we can do is a compromise: learn to recognize situations in which mistakes are likely and try harder to avoid significant mistakes when the stakes are high. The premise of this book is that it is easier to recognize other people's mistakes than our own.

Useful Fictions

You have been invited to think of the two systems as agents within the mind, with their individual personalities, abilities, and limitations. I will often use sentences in which the systems are the subjects, such as, "System 2 calculates products."

The use of such language is considered a sin in the professional circles in which I travel, because it seems to explain the thoughts and actions of a person by the thoughts and actions of little people inside the person's head. Grammatically the sentence about System 2 is similar to "The butler steals the petty cash." My colleagues would point out that the butler's action actually explains the disappearance of the cash, and they rightly question whether the sentence about System 2 explains how products are calculated. My answer is that the brief active sentence that attributes calculation to System 2 is intended as a description, not an explanation. It is meaningful only because of what you already know about System 2. It is shorthand for the following: "Mental arithmetic is a voluntary activity that requires effort, should not be performed while making a left turn, and is associated with dilated pupils and an accelerated heart rate."

Similarly, the statement that "highway driving under routine conditions is left to System 1" means that steering the car around a bend is automatic and almost effortless. It also implies that an experienced driver can drive on an empty highway while conducting a conversation. Finally, "System 2 prevented James from reacting foolishly to the insult" means that James would have been more aggressive in his response if his capacity for effortful control had been disrupted (for example, if he had been drunk).

System 1 and System 2 are so central to the story I tell in this book that I must make it absolutely clear that they are217at they a fictitious characters. Systems 1 and 2 are not systems in the standard sense of entities with interacting aspects or parts. And there is no one part of the brain that either of the systems would call home. You may well ask: What is the point of introducing fictitious characters with ugly names into a serious book? The answer is that the characters are useful because of some quirks of our minds, yours and mine. A sentence is understood more easily if it describes what an agent (System 2) does than if it describes what something is, what properties it has. In other words, “System 2” is a better subject for a sentence than “mental arithmetic.” The mind—especially System 1—appears to have a special aptitude for the construction and interpretation of stories about active agents, who have personalities, habits, and abilities. You quickly formed a bad opinion of the thieving butler, you expect more bad behavior from him, and you will remember him for a while. This is also my hope for the language of systems.

Why call them System 1 and System 2 rather than the more descriptive “automatic system” and “effortful system”? The reason is simple: “Automatic system” takes longer to say than “System 1” and therefore takes more space in your working memory. This matters, because anything that occupies your working memory reduces your ability to think. You should treat “System 1” and “System 2” as nicknames, like Bob and Joe, identifying characters that you will get to know over the course of this book. The fictitious systems make it easier for me to think about judgment and choice, and will make it easier for you to understand what I say.

Speaking of System 1 and System 2

> “He had an impression, but some of his impressions are illusions.”

> “This was a pure System 1 response. She reacted to the threat before she recognized it.”

> “This is your System 1 talking. Slow down and let your System 2 take control.”

P

Attention and Effort

In the unlikely event of this book being made into a film, System 2 would be a supporting character who believes herself to be the hero. The defining feature of System 2, in this story, is that its operations are effortful, and one of its main characteristics is laziness, a reluctance to invest more effort than is strictly necessary. As a consequence, the thoughts and actions that System 2 believes it has chosen are often guided by the figure at the center of the story, System 1. However, there are vital tasks that only System 2 can perform because they require effort and acts of self-control in which the intuitions and impulses of System 1 are overcome.

Mental Effort

If you wish to experience your System 2 working at full tilt, the following exercise will do; it should br”0%e ca Tting you to the limits of your cognitive abilities within 5 seconds. To start, make up several strings of 4 digits, all different, and write each string on an index card. Place a blank card on top of the deck. The task that you will perform is called Add-1. Here is how it goes:

> Start beating a steady rhythm (or better yet, set a metronome at 1/sec). Remove the blank card and read the four digits aloud. Wait for two beats, then report a string in which each of the original digits is incremented by 1. If the digits on the card are 5294, the correct response is 6305. Keeping the rhythm is important.

Few people can cope with more than four digits in the Add-1 task, but if you want a harder challenge, please try Add-3.

If you would like to know what your body is doing while your mind is hard at work, set up two piles of books on a sturdy table, place a video camera on one and lean your chin on the other, get the video going, and stare at the camera lens while you work on Add-1 or Add-3 exercises. Later, you will find in the changing size of your pupils a faithful record of how hard you worked.

I have a long personal history with the Add-1 task. Early in my career I spent a year at the University of Michigan, as a visitor in a laboratory that studied hypnosis. Casting about for a useful topic of research, I found an article in *Scientific American* in which the psychologist Eckhard Hess described the pupil of the eye as a window to the soul. I reread it recently and again found it inspiring. It begins with Hess reporting that his wife had noticed his pupils widening as he watched beautiful nature pictures, and it ends with two striking pictures of the same good-looking woman, who somehow appears much more attractive in one than in the other. There is only one difference: the pupils of the eyes appear dilated in the attractive picture and constricted in the other. Hess also wrote of belladonna, a pupil-dilating substance that was used as a cosmetic, and of bazaar shoppers who wear dark glasses in order to hide their level of interest from merchants.

One of Hess's findings especially captured my attention. He had noticed that the pupils are sensitive indicators of mental effort—they dilate substantially when people multiply two-digit numbers, and they dilate more if the problems are hard than if they are easy. His observations indicated that the response to mental effort is distinct from emotional arousal. Hess's work did not have much to do with hypnosis, but I concluded that the idea of a visible indication of mental effort had promise as a research topic. A graduate student in the lab, Jackson Beatty, shared my enthusiasm and we got to work.

Beatty and I developed a setup similar to an optician's examination room, in which the experimental participant leaned her head on a chin-and-forehead rest and stared at a camera while listening to prerecorded information and answering questions on the recorded beats of a metronome. The beats triggered an infrared flash every second, causing a picture to be taken. At the end of each experimental session, we would rush to have the film developed, project the images of the pupil on a screen, and go to work with a ruler. The method was a perfect fit for young and impatient researchers: we knew our results almost immediately, and they always told a clear story.

Beatty and I focused on paced tasks, such as Add-1, in which we knew precisely what was on the subject's mind at any time. We recorded strings of digits on beats of the metronome and instructed the subject to repeat or transform the digits one indigits onby one, maintaining the same rhythm. We soon discovered that the size of the pupil varied second by second, reflecting the changing demands of the task. The shape of the response was an inverted V. As you experienced it if you tried Add-1 or Add-3, effort builds up with every added digit that you hear, reaches an almost intolerable peak as you rush to produce a transformed string during and immediately after the pause, and relaxes gradually as you "unload" your short-term memory. The pupil data corresponded precisely to subjective experience: longer strings reliably caused larger dilations, the transformation task compounded the effort, and the peak of pupil size coincided with maximum effort. Add-1 with four digits caused a larger dilation than the task of holding seven digits for immediate recall. Add-3, which is much more difficult, is the most demanding that I ever observed. In the first 5 seconds, the pupil dilates by about 50% of its original area and heart rate increases by about 7 beats per minute. This is as hard as people can work—they give up if more is asked of them. When we exposed our subjects to more digits than they could remember, their pupils stopped dilating or actually shrank.

We worked for some months in a spacious basement suite in which we had set up a closed-circuit system that projected an image of the subject's pupil on a screen in the corridor; we also could hear what was happening in the laboratory. The diameter of the projected pupil was about a foot; watching it dilate and contract when the participant was at work was a fascinating sight, quite an attraction for visitors in our lab. We amused ourselves and impressed our guests by our ability to divine when the participant gave up on a task. During a mental multiplication, the pupil normally dilated to a large size within a few seconds and stayed large as long as the individual kept working on the problem; it contracted immediately when she found a solution or gave up. As we watched from the corridor, we would sometimes surprise both the owner of the pupil and our guests by asking, “Why did you stop working just now?” The answer from inside the lab was often, “How did you know?” to which we would reply, “We have a window to your soul.”

The casual observations we made from the corridor were sometimes as informative as the formal experiments. I made a significant discovery as I was idly watching a woman’s pupil during a break between two tasks. She had kept her position on the chin rest, so I could see the image of her eye while she engaged in routine conversation with the experimenter. I was surprised to see that the pupil remained small and did not noticeably dilate as she talked and listened. Unlike the tasks that we were studying, the mundane conversation apparently demanded little or no effort—no more than retaining two or three digits. This was a eureka moment: I realized that the tasks we had chosen for study were exceptionally effortful. An image came to mind: mental life—today I would speak of the life of System 2—is normally conducted at the pace of a comfortable walk, sometimes interrupted by episodes of jogging and on rare occasions by a frantic sprint. The Add-1 and Add-3 exercises are sprints, and casual chatting is a stroll.

We found that people, when engaged in a mental sprint, may become effectively blind. The authors of *The Invisible Gorilla* had made the gorilla “invisible” by keeping the observers intensely busy counting passes. We reported a rather less dramatic example of blindness during Add-1. Our subjects were exposed to a series of rapidly flashing letters while they worked. They were told to give the task complete priority, but they were also asked to report, at the end of the digit task, whether the letter *K* had appeared at any rored at antime during the trial. The main finding was that the ability to detect and report the target letter changed in the course of the 10 seconds of the exercise. The observers almost never missed a *K* that was shown at the beginning or near the end of the Add-1 task but they missed the target almost half the time when mental effort was at its peak, although we had pictures of their wide-open eye staring straight at it. Failures of detection followed the same inverted-V pattern as the dilating pupil. The similarity was reassuring: the pupil was a good measure of the physical arousal that accompanies mental effort, and we could go ahead and use it to understand how the mind works.

Much like the electricity meter outside your house or apartment, the pupils offer an index of the current rate at which mental energy is used. The analogy goes deep. Your use of electricity depends on what you choose to do, whether to light a room or toast a piece of bread. When you turn on a bulb or a toaster, it draws the energy it needs but no more. Similarly, we decide what to do, but we have limited control over the effort of doing it. Suppose you are shown four digits, say, 9462, and told that your life depends on holding them in memory for 10 seconds. However much you want to live, you cannot exert as much effort in this task as you would be forced to invest to complete an Add-3 transformation on the same digits.

System 2 and the electrical circuits in your home both have limited capacity, but they respond differently to threatened overload. A breaker trips when the demand for current is excessive, causing all devices on that circuit to lose power at once. In contrast, the response to mental overload is selective and precise: System 2 protects the most important activity, so it receives the attention it needs; “spare capacity” is allocated second by second to other tasks. In our version of the gorilla experiment, we instructed the participants to assign priority to the digit task. We know that they followed that instruction, because the timing of the visual target had no effect on the main task. If the critical letter was presented at a time of high demand, the subjects simply did not see it. When the transformation task was less demanding, detection performance was better.

The sophisticated allocation of attention has been honed by a long evolutionary history. Orienting and responding quickly to the gravest threats or most promising opportunities improved the chance of survival, and this capability is certainly not restricted to humans. Even in modern humans, System 1 takes over in emergencies and assigns total priority to self-protective actions. Imagine yourself at the wheel of a car that unexpectedly skids on a large oil slick. You will find that you have responded to the threat before you became fully conscious of it.

Beatty and I worked together for only a year, but our collaboration had a large effect on our subsequent careers. He eventually became the leading authority on "cognitive pupillometry," and I wrote a book titled *Attention and Effort*, which was based in large part on what we learned together and on follow-up research I did at Harvard the following year. We learned a great deal about the working mind—which I now think of as System 2—from measuring pupils in a wide variety of tasks.

As you become skilled in a task, its demand for energy diminishes. Studies of the brain have shown that the pattern of activity associated with an action changes as skill increases, with fewer brain regions involved. Talent has similar effects. Highly intelligent individuals need less effort to solve the same problems, as indicated by both pupil size and brain activity. A general "law of least effort" appd t" alies to cognitive as well as physical exertion. The law asserts that if there are several ways of achieving the same goal, people will eventually gravitate to the least demanding course of action. In the economy of action, effort is a cost, and the acquisition of skill is driven by the balance of benefits and costs. Laziness is built deep into our nature.

The tasks that we studied varied considerably in their effects on the pupil. At baseline, our subjects were awake, aware, and ready to engage in a task—probably at a higher level of arousal and cognitive readiness than usual. Holding one or two digits in memory or learning to associate a word with a digit (3 = door) produced reliable effects on momentary arousal above that baseline, but the effects were minuscule, only 5% of the increase in pupil diameter associated with Add-3. A task that required discriminating between the pitch of two tones yielded significantly larger dilations. Recent research has shown that inhibiting the tendency to read distracting words (as in figure 2 of the preceding chapter) also induces moderate effort. Tests of short-term memory for six or seven digits were more effortful. As you can experience, the request to retrieve and say aloud your phone number or your spouse's birthday also requires a brief but significant effort, because the entire string must be held in memory as a response is organized. Mental multiplication of two-digit numbers and the Add-3 task are near the limit of what most people can do.

What makes some cognitive operations more demanding and effortful than others? What outcomes must we purchase in the currency of attention? What can System 2 do that System 1 cannot? We now have tentative answers to these questions.

Effort is required to maintain simultaneously in memory several ideas that require separate actions, or that need to be combined according to a rule—rehearsing your shopping list as you enter the supermarket, choosing between the fish and the veal at a restaurant, or combining a surprising result from a survey with the information that the sample was small, for example. System 2 is the only one that can follow rules, compare objects on several attributes, and make deliberate choices between options. The automatic System 1 does not have these capabilities. System 1 detects simple relations ("they are all alike," "the son is much taller than the father") and excels at integrating information about one thing, but it does not deal with multiple distinct topics at once, nor is it adept at using purely statistical information. System 1 will detect that a person described as "a meek and tidy soul, with a need for order and structure, and a passion for detail" resembles a caricature librarian, but combining this intuition with knowledge about the small number of librarians is a task that only System 2 can perform—if System 2 knows how to do so, which is true of few people.

A crucial capability of System 2 is the adoption of "task sets": it can program memory to obey an instruction that overrides habitual responses. Consider the following: Count all occurrences of the letter *f* in this page. This is not a task you have ever performed before and it will not come naturally to you, but your System 2 can take it on. It will be effortful to set yourself up for this exercise, and effortful to carry it out, though you will surely improve with practice. Psychologists speak of "executive control" to describe the adoption and termination of task sets, and neuroscientists have identified the main regions of the brain that serve the executive function. One of these regions is involved whenever a conflict must be resolved. Another is the prefrontal area of the brain, a region that is substantially more developed in humans tht un humans an in other primates, and is involved in operations that we associate with intelligence.

Now suppose that at the end of the page you get another instruction: count all the commas in the next page. This will be harder, because you will have to overcome the newly acquired tendency to focus attention on the letter *f*. One of the significant discoveries of cognitive psychologists in recent decades is that switching from one task to another is effortful, especially under time pressure. The need for rapid switching is one of the reasons that Add-3 and mental multiplication are so difficult. To perform the Add-3 task, you must hold several digits in your working memory at the same time, associating each with a particular operation: some digits are in the queue to be transformed, one is in the process of transformation, and others, already transformed, are retained for reporting. Modern tests of working memory require the individual to switch repeatedly between two demanding tasks, retaining the results of one operation while performing the other. People who do well on these tests tend to do well on tests of general intelligence. However, the ability to control attention is not simply a measure of intelligence; measures of efficiency in the control of attention predict performance of air traffic controllers and of Israeli Air Force pilots beyond the effects of intelligence.

Time pressure is another driver of effort. As you carried out the Add-3 exercise, the rush was imposed in part by the metronome and in part by the load on memory. Like a juggler with several balls in the air, you cannot afford to slow down; the rate at which material decays in memory forces the pace, driving you to refresh and rehearse information before it is lost. Any task that requires you to keep several ideas in mind at the same time has the same hurried character. Unless you have the good fortune of a capacious working memory, you may be forced to work uncomfortably hard. The most effortful forms of slow thinking are those that require you to think fast.

You surely observed as you performed Add-3 how unusual it is for your mind to work so hard. Even if you think for a living, few of the mental tasks in which you engage in the course of a working day are as demanding as Add-3, or even as demanding as storing six digits for immediate recall. We normally avoid mental overload by dividing our tasks into multiple easy steps, committing intermediate results to long-term memory or to paper rather than to an easily overloaded working memory. We cover long distances by taking our time and conduct our mental lives by the law of least effort.

Speaking of Attention and Effort

> “I won’t try to solve this while driving. This is a pupil-dilating task. It requires mental effort!”

> “The law of least effort is operating here. He will think as little as possible.”

> “She did not forget about the meeting. She was completely focused on something else when the meeting was set and she just didn’t hear you.”

> “What came quickly to my mind was an intuition from System 1. I’ll have to start over and search my memory deliberately.”

P

The Lazy Controller

I spend a few months each year in Berkeley, and one of my great pleasures there is a daily four-mile walk on a marked path in the hills, with a fine view of San Francisco Bay. I usually keep track of my time and have learned a fair amount about effort from doing so. I have found a speed, about 17 minutes for a mile, which I experience as a stroll. I certainly exert physical effort and burn more calories at that speed than if I sat in a recliner, but I experience no strain, no conflict, and no need to push myself. I am also able to think and work while walking at that rate. Indeed, I suspect that the mild physical arousal of the walk may spill over into greater mental alertness.

System 2 also has a natural speed. You expend some mental energy in random thoughts and in monitoring what goes on around you even when your mind does nothing in particular, but there is little strain. Unless you are in a situation that makes you unusually wary or self-conscious, monitoring what happens in the environment or inside your head demands little effort. You make many small decisions as you drive your car, absorb some information as you read the newspaper, and conduct routine exchanges of pleasantries with a spouse or a colleague, all with little effort and no strain. Just like a stroll.

It is normally easy and actually quite pleasant to walk and think at the same time, but at the extremes these activities appear to compete for the limited resources of System 2. You can confirm this claim by a simple experiment. While walking comfortably with a friend, ask him to compute 23 × 78 in his head, and to do so immediately. He will almost certainly stop in his tracks. My experience is that I can think while strolling but cannot engage in mental work that imposes a heavy load on short-term memory. If I must construct an intricate argument under time pressure, I would rather be still, and I would prefer sitting to standing. Of course, not all slow thinking requires that form of intense concentration and effortful computation—I did the best thinking of my life on leisurely walks with Amos.

Accelerating beyond my strolling speed completely changes the experience of walking, because the transition to a faster walk brings about a sharp deterioration in my ability to think coherently. As I speed up, my attention is drawn with increasing frequency to the experience of walking and to the deliberate maintenance of the faster pace. My ability to bring a train of thought to a conclusion is impaired accordingly. At the highest speed I can sustain on the hills, about 14 minutes for a mile, I do not even try to think of anything else. In addition to the physical effort of moving my body rapidly along the path, a mental effort of self-control is needed to resist the urge to slow down. Self-control and deliberate thought apparently draw on the same limited budget of effort.

For most of us, most of the time, the maintenance of a coherent train of thought and the occasional engagement in effortful thinking also require self-control. Although I have not conducted a systematic survey, I suspect that frequent switching of tasks and speeded-up mental work are not intrinsically pleasurable, and that people avoid them when possible. This is how the law of least effort comes to be a law. Even in the absence of time pressure, maintaining a coherent train of thought requires discipline. An observer of the number of times I look at e-mail or investigate the refrigerator during an hour of writing could wahene dd reasonably infer an urge to escape and conclude that keeping at it requires more self-control than I can readily muster.

Fortunately, cognitive work is not always aversive, and people sometimes expend considerable effort for long periods of time without having to exert willpower. The psychologist Mihaly Csikszentmihalyi (pronounced six-cent-mihaly) has done more than anyone else to study this state of effortless attending, and the name he proposed for it, *flow*, has become part of the language. People who experience flow describe it as "a state of effortless concentration so deep that they lose their sense of time, of themselves, of their problems," and their descriptions of the joy of that state are so compelling that Csikszentmihalyi has called it an "optimal experience." Many activities can induce a sense of flow, from painting to racing motorcycles—and for some fortunate authors I know, even writing a book is often an optimal experience. Flow neatly separates the two forms of effort: concentration on the task and the deliberate control of attention. Riding a motorcycle at 150 miles an hour and playing a competitive game of chess are certainly very effortful. In a state of flow, however, maintaining focused attention on these absorbing activities requires no exertion of self-control, thereby freeing resources to be directed to the task at hand.

The Busy and Depleted System 2

It is now a well-established proposition that both self-control and cognitive effort are forms of mental work. Several psychological studies have shown that people who are simultaneously challenged by a demanding cognitive task and by a temptation are more likely to yield to the temptation. Imagine that you are asked to retain a list of seven digits for a minute or two. You are told that remembering the digits is your top priority. While your attention is focused on the digits, you are offered a choice between two desserts: a sinful chocolate cake and a virtuous fruit salad. The evidence suggests that you would be more likely to select the tempting chocolate cake when your mind is loaded with digits. System 1 has more influence on behavior when System 2 is busy, and it has a sweet tooth.

People who are *cognitively busy* are also more likely to make selfish choices, use sexist language, and make superficial judgments in social situations. Memorizing and repeating digits loosens the hold of System 2 on behavior, but of course cognitive load is not the only cause of weakened self-control. A few drinks have the same effect, as does a sleepless night. The self-control of morning people is impaired at night; the reverse is true of night people. Too much concern about how well one is doing in a task sometimes disrupts performance by loading short-term memory with pointless anxious thoughts. The conclusion is straightforward: self-control requires attention and effort. Another way of saying this is that controlling thoughts and behaviors is one of the tasks that System 2 performs.

A series of surprising experiments by the psychologist Roy Baumeister and his colleagues has shown conclusively that all variants of voluntary effort—cognitive, emotional, or physical—draw at least partly on a shared pool of mental energy. Their experiments involve successive rather than simultaneous tasks.

Baumeister's group has repeatedly found that an effort of will or self-control is tiring; if you have had to force yourself to do something, you are less willing or less able to exert self-control when the next challenge comes around. The phenomenon has been named *ego depletion*. In a typical demo thypical denstration, participants who are instructed to stifle their emotional reaction to an emotionally charged film will later perform poorly on a test of physical stamina—how long they can maintain a strong grip on a dynamometer in spite of increasing discomfort. The emotional effort in the first phase of the experiment reduces the ability to withstand the pain of sustained muscle contraction, and ego-depleted people therefore succumb more quickly to the urge to quit. In another experiment, people are first depleted by a task in which they eat virtuous foods such as radishes and celery while resisting the temptation to indulge in chocolate and rich cookies. Later, these people will give up earlier than normal when faced with a difficult cognitive task.

The list of situations and tasks that are now known to deplete self-control is long and varied. All involve conflict and the need to suppress a natural tendency. They include:

avoiding the thought of white bears

inhibiting the emotional response to a stirring film

making a series of choices that involve conflict

trying to impress others

responding kindly to a partner's bad behavior

interacting with a person of a different race (for prejudiced individuals)

The list of indications of depletion is also highly diverse:

deviating from one's diet

overspending on impulsive purchases

reacting aggressively to provocation

persisting less time in a handgrip task

performing poorly in cognitive tasks and logical decision making

The evidence is persuasive: activities that impose high demands on System 2 require self-control, and the exertion of self-control is depleting and unpleasant. Unlike cognitive load, ego depletion is at least in part a loss of motivation. After exerting self-control in one task, you do not feel like making an effort in another, although you could do it if you really had to. In several experiments, people were able to resist the effects of ego depletion when given a strong incentive to do so. In contrast, increasing effort is not an option when you must keep six digits in short-term memory while performing a task. Ego depletion is not the same mental state as cognitive busyness.

The most surprising discovery made by Baumeister's group shows, as he puts it, that the idea of mental energy is more than a mere metaphor. The nervous system consumes more glucose than most other parts of the body, and effortful mental activity appears to be especially expensive in the currency of glucose. When you are actively involved in difficult cognitive reasoning or engaged in a task that requires self-control, your blood glucose level drops. The effect is analogous to a runner who draws down glucose stored in her muscles during a sprint. The bold implication of this idea is that the effects of ego depletion could be undone by ingesting glucose, and Baumeister and his colleagues have confirmed this hypothesis n ohypothesiin several experiments.

Volunteers in one of their studies watched a short silent film of a woman being interviewed and were asked to interpret her body language. While they were performing the task, a series of words crossed the screen in slow succession. The participants were specifically instructed to ignore the words, and if they found their attention drawn away they had to refocus their concentration on the woman's behavior. This act of self-control was known to cause ego depletion. All the volunteers drank some lemonade before participating in a second task. The lemonade was sweetened with glucose for half of them and with Splenda for the others. Then all participants were given a task in which they needed to overcome an intuitive response to get the correct answer. Intuitive errors are normally much more frequent among ego-depleted people, and the drinkers of Splenda showed the expected depletion effect. On the other hand, the glucose drinkers were not depleted. Restoring the level of available sugar in the brain had prevented the deterioration of performance. It will take some time and much further research to establish whether the tasks that cause glucose-depletion also cause the momentary arousal that is reflected in increases of pupil size and heart rate.

A disturbing demonstration of depletion effects in judgment was recently reported in the *Proceedings of the National Academy of Sciences*. The unwitting participants in the study were eight parole judges in Israel. They spend entire days reviewing applications for parole. The cases are presented in random order, and the judges spend little time on each one, an average of 6 minutes. (The default decision is denial of parole; only 35% of requests are approved. The exact time of each decision is recorded, and the times of the judges' three food breaks—morning break, lunch, and afternoon break—during the day are recorded as well.) The authors of the study plotted the proportion of approved requests against the time since the last food break. The proportion spikes after each meal, when about 65% of requests are granted. During the two hours or so until the judges' next feeding, the approval rate drops steadily, to about zero just before the meal. As you might expect, this is an unwelcome result and the authors carefully checked many alternative explanations. The best possible account of the data provides bad news: tired and hungry judges tend to fall back on the easier default position of denying requests for parole. Both fatigue and hunger probably play a role.

The Lazy System 2

One of the main functions of System 2 is to monitor and control thoughts and actions "suggested" by System 1, allowing some to be expressed directly in behavior and suppressing or modifying others.

For an example, here is a simple puzzle. Do not try to solve it but listen to your intuition:

A bat and ball cost $1.10.

The bat costs one dollar more than the ball.

How much does the ball cost?

A number came to your mind. The number, of course, is 10: 10¢. The distinctive mark of this easy puzzle is that it evokes an answer that is intuitive, appealing, and wrong. Do the math, and you will see. If the ball costs 10¢, then the total cost will be $1.20 (10¢ for the ball and $1.10 for the bat), not $1.10. The correct answer is 5¢. It%">5¢. is safe to assume that the intuitive answer also came to the mind of those who ended up with the correct number—they somehow managed to resist the intuition.

Shane Frederick and I worked together on a theory of judgment based on two systems, and he used the bat-and-ball puzzle to study a central question: How closely does System 2 monitor the suggestions of System 1? His reasoning was that we know a significant fact about anyone who says that the ball costs 10¢: that person did not actively check whether the answer was correct, and her System 2 endorsed an intuitive answer that it could have rejected with a small investment of effort. Furthermore, we also know that the people who give the intuitive answer have missed an obvious social cue; they should have wondered why anyone would include in a questionnaire a puzzle with such an obvious answer. A failure to check is remarkable because the cost of checking is so low: a few seconds of mental work (the problem is moderately difficult), with slightly tensed muscles and dilated pupils, could avoid an embarrassing mistake. People who say 10¢ appear to be ardent followers of the law of least effort. People who avoid that answer appear to have more active minds.

Many thousands of university students have answered the bat-and-ball puzzle, and the results are shocking. More than 50% of students at Harvard, MIT, and Princeton ton gave the intuitive—incorrect—answer. At less selective universities, the rate of demonstrable failure to check was in excess of 80%. The bat-and-ball problem is our first encounter with an observation that will be a recurrent theme of this book: many people are overconfident, prone to place too much faith in their intuitions. They apparently find cognitive effort at least mildly unpleasant and avoid it as much as possible.

Now I will show you a logical argument—two premises and a conclusion. Try to determine, as quickly as you can, if the argument is logically valid. Does the conclusion follow from the premises?

All roses are flowers.

Some flowers fade quickly.

Therefore some roses fade quickly.

A large majority of college students endorse this syllogism as valid. In fact the argument is flawed, because it is possible that there are no roses among the flowers that fade quickly. Just as in the bat-and-ball problem, a plausible answer comes to mind immediately. Overriding it requires hard work—the insistent idea that "it's true, it's true!" makes it difficult to check the logic, and most people do not take the trouble to think through the problem.

This experiment has discouraging implications for reasoning in everyday life. It suggests that when people believe a conclusion is true, they are also very likely to believe arguments that appear to support it, even when these arguments are unsound. If System 1 is involved, the conclusion comes first and the arguments follow.

Next, consider the following question and answer it quickly before reading on:

How many murders occur in the state of Michigan in one year?

The question, which was also devised by Shane Frederick, is again a challenge to System 2. The "trick" is whether the respondent will remember that Detroit, a high-crime c thigh-crimeity, is in Michigan. College students in the United States know this fact and will correctly identify Detroit as the largest city in Michigan. But knowledge of a fact is not all-or-none. Facts that we know do not always come to mind when we need them. People who remember that Detroit is in Michigan give higher estimates of the murder rate in the state than people who do not, but a majority of Frederick's respondents did not think of the city when questioned about the state. Indeed, the average guess by people who were asked about Michigan is *lower* than the guesses of a similar group who were asked about the murder rate in Detroit.

Blame for a failure to think of Detroit can be laid on both System 1 and System 2. Whether the city comes to mind when the state is mentioned depends in part on the automatic function of memory. People differ in this respect. The representation of the state of Michigan is very detailed in some people's minds: residents of the state are more likely to retrieve many facts about it than people who live elsewhere; geography buffs will retrieve more than others who specialize in baseball statistics; more intelligent individuals are more likely than others to have rich representations of most things. Intelligence is not only the ability to reason; it is also the ability to find relevant material in memory and to deploy attention when needed. Memory function is an attribute of System 1. However, everyone has the option of slowing down to conduct an active search of memory for all possibly relevant facts—just as they could slow down to check the intuitive answer in the bat-and-ball problem. The extent of deliberate checking and search is a characteristic of System 2, which varies among individuals.

The bat-and-ball problem, the flowers syllogism, and the Michigan/Detroit problem have something in common. Failing these minitests appears to be, at least to some extent, a matter of insufficient motivation, not trying hard enough. Anyone who can be admitted to a good university is certainly able to reason through the first two questions and to reflect about Michigan long enough to remember the major city in that state and its crime problem. These students can solve much more difficult problems when they are not tempted to accept a superficially plausible answer that comes readily to mind. The ease with which they are satisfied enough to stop thinking is rather troubling. “Lazy” is a harsh judgment about the self-monitoring of these young people and their System 2, but it does not seem to be unfair. Those who avoid the sin of intellectual sloth could be called “engaged.” They are more alert, more intellectually active, less willing to be satisfied with superficially attractive answers, more skeptical about their intuitions. The psychologist Keith Stanovich would call them more rational.

Intelligence, Control, Rationality

Researchers have applied diverse methods to examine the connection between thinking and self-control. Some have addressed it by asking the correlation question: If people were ranked by their self-control and by their cognitive aptitude, would individuals have similar positions in the two rankings?

In one of the most famous experiments in the history of psychology, Walter Mischel and his students exposed four-year-old children to a cruel dilemma. They were given a choice between a small reward (one Oreo), which they could have at any time, or a larger reward (two cookies) for which they had to wait 15 minutes under difficult conditions. They were to remain alone in a room, facing a desk with two objects: a single cookie and a bell that the child could ring at any time to call in the experimenter and receiven oand recei the one cookie. As the experiment was described: “There were no toys, books, pictures, or other potentially distracting items in the room. The experimenter left the room and did not return until 15 min had passed or the child had rung the bell, eaten the rewards, stood up, or shown any signs of distress.”

The children were watched through a one-way mirror, and the film that shows their behavior during the waiting time always has the audience roaring in laughter. About half the children managed the feat of waiting for 15 minutes, mainly by keeping their attention away from the tempting reward. Ten or fifteen years later, a large gap had opened between those who had resisted temptation and those who had not. The resisters had higher measures of executive control in cognitive tasks, and especially the ability to reallocate their attention effectively. As young adults, they were less likely to take drugs. A significant difference in intellectual aptitude emerged: the children who had shown more self-control as four-year-olds had substantially higher scores on tests of intelligence.

A team of researchers at the University of Oregon explored the link between cognitive control and intelligence in several ways, including an attempt to raise intelligence by improving the control of attention. During five 40-minute sessions, they exposed children aged four to six to various computer games especially designed to demand attention and control. In one of the exercises, the children used a joystick to track a cartoon cat and move it to a grassy area while avoiding a muddy area. The grassy areas gradually shrank and the muddy area expanded, requiring progressively more precise control. The testers found that training attention not only improved executive control; scores on nonverbal tests of intelligence also improved and the improvement was maintained for several months. Other research by the same group identified specific genes that are involved in the control of attention, showed that parenting techniques also affected this ability, and demonstrated a close connection between the children's ability to control their attention and their ability to control their emotions.

Shane Frederick constructed a Cognitive Reflection Test, which consists of the bat-and-ball problem and two other questions, chosen because they also invite an intuitive answer that is both compelling and wrong (the questions are shown here). He went on to study the characteristics of students who score very low on this test—the supervisory function of System 2 is weak in these people—and found that they are prone to answer questions with the first idea that comes to mind and unwilling to invest the effort needed to check their intuitions. Individuals who uncritically follow their intuitions about puzzles are also prone to accept other suggestions from System 1. In particular, they are impulsive, impatient, and keen to receive immediate gratification. For example, 63% of the intuitive respondents say they would prefer to get $3,400 this month rather than $3,800 next month. Only 37% of those who solve all three puzzles correctly have the same shortsighted preference for receiving a smaller amount immediately. When asked how much they will pay to get overnight delivery of a book they have ordered, the low scorers on the Cognitive Reflection Test are willing to pay twice as much as the high scorers. Frederick's findings suggest that the characters of our psychodrama have different "personalities." System 1 is impulsive and intuitive; System 2 is capable of reasoning, and it is cautious, but at least for some people it is also lazy. We recognize related differences among individuals: some people are more like their System 2; others are closer to their System 1. This simple test has emerged as one of the better predictors of laztestors of ly thinking.

Keith Stanovich and his longtime collaborator Richard West originally introduced the terms System 1 and System 2 (they now prefer to speak of Type 1 and Type 2 processes). Stanovich and his colleagues have spent decades studying differences among individuals in the kinds of problems with which this book is concerned. They have asked one basic question in many different ways: What makes some people more susceptible than others to biases of judgment? Stanovich published his conclusions in a book titled *Rationality and the Reflective Mind*, which offers a bold and distinctive approach to the topic of this chapter. He draws a sharp distinction between two parts of System 2—indeed, the distinction is so sharp that he calls them separate "minds." One of these minds (he calls it algorithmic) deals with slow thinking and demanding computation. Some people are better than others in these tasks of brain power—they are the individuals who excel in intelligence tests and are able to switch from one task to another quickly and efficiently. However, Stanovich argues that high intelligence does not make people immune to biases. Another ability is involved, which he labels rationality. Stanovich's concept of a rational person is similar to what I earlier labeled "engaged." The core of his argument is that *rationality* should be distinguished from *intelligence*. In his view, superficial or "lazy" thinking is a flaw in the reflective mind, a failure of rationality. This is an attractive and thought-provoking idea. In support of it, Stanovich and his colleagues have found that the bat-and-ball question and others like it are somewhat better indicators of our susceptibility to cognitive errors than are conventional measures of intelligence, such as IQ tests. Time will tell whether the distinction between intelligence and rationality can lead to new discoveries.

Speaking of Control

"She did not have to struggle to stay on task for hours. She was in a state of *flow*."

"His ego was depleted after a long day of meetings. So he just turned to standard operating procedures instead of thinking through the problem."

"He didn't bother to check whether what he said made sense. Does he usually have a lazy System 2 or was he unusually tired?"

"Unfortunately, she tends to say the first thing that comes into her mind. She probably also has trouble delaying gratification. Weak System 2."

P

The Associative Machine

To begin your exploration of the surprising workings of System 1, look at the following words:

Bananas Vomit

A lot happened to you during the last second or two. You experienced some unpleasant images and memories. Your face twisted slightly in an expression of disgust, and you may have pushed this book imperceptibly farther away. Your heart rate increased, the hair on your arms rose a little, and your sweat glands were activated. In short, you responded to the disgusting word with an attenuated version of how you would react to the actual event. All of this was completely automatic, beyond your control.

There was no particular reason to do so, but your mind automatically assumed a temporal sequence and a causal connection between the words *bananas* and *vomit*, forming a sketchy scenario in which bananas caused the sickness. As a result, you are experiencing a temporary aversion to bananas (don't worry, it will pass). The state of your memory has changed in other ways: you are now unusually ready to recognize and respond to objects and concepts associated with "vomit," such as sick, stink, or nausea, and words associated with "bananas," such as yellow and fruit, and perhaps apple and berries.

Vomiting normally occurs in specific contexts, such as hangovers and indigestion. You would also be unusually ready to recognize words associated with other causes of the same unfortunate outcome. Furthermore, your System 1 noticed the fact that the juxtaposition of the two words is uncommon; you probably never encountered it before. You experienced mild surprise.

This complex constellation of responses occurred quickly, automatically, and effortlessly. You did not will it and you could not stop it. It was an operation of System 1. The events that took place as a result of your seeing the words happened by a process called associative activation: ideas that have been evoked trigger many other ideas, in a spreading cascade of activity in your brain. The essential feature of this complex set of mental events is its coherence. Each element is connected, and each supports and strengthens the others. The word evokes memories, which evoke emotions, which in turn evoke facial expressions and other reactions, such as a general tensing up and an avoidance tendency. The facial expression and the avoidance motion intensify the feelings to which they are linked, and the feelings in turn reinforce compatible ideas. All this happens quickly and all at once, yielding a self-reinforcing pattern of cognitive, emotional, and physical responses that is both diverse and integrated—it has been called *associatively coherent*.

In a second or so you accomplished, automatically and unconsciously, a remarkable feat. Starting from a completely unexpected event, your System 1 made as much sense as possible of the situation—two simple words, oddly juxtaposed—by linking the words in a causal story; it evaluated the possible threat (mild to moderate) and created a context for future developments by preparing you for events that had just become more likely; it also created a context for the current event by evaluating how surprising it was. You ended up as informed about the past and as prepared for the future as you could be.

An odd feature of what happened is that your System 1 treated the mere conjunction of two words as representations of reality. Your body reacted in an attenuated replica of a reaction to the real thing, and the emotional response and physical recoil were part of the interpretation of the event. As cognitive scientists have emphasized in recent years, cognition is embodied; you think with your body, not only with your brain.

The mechanism that causes these mental events has been known for a long time: it is the ass12;velyociation of ideas. We all understand from experience that ideas follow each other in our conscious mind in a fairly orderly way. The British philosophers of the seventeenth and eighteenth centuries searched for the rules that explain such sequences. In *An Enquiry Concerning Human Understanding*, published in 1748, the Scottish philosopher David Hume reduced the principles of association to three: resemblance, contiguity in time and place, and causality. Our concept of association has changed radically since Hume’s days, but his three principles still provide a good start.

I will adopt an expansive view of what an idea is. It can be concrete or abstract, and it can be expressed in many ways: as a verb, as a noun, as an adjective, or as a clenched fist. Psychologists think of ideas as nodes in a vast network, called associative memory, in which each idea is linked to many others. There are different types of links: causes are linked to their effects (virus → cold); things to their properties (lime → green); things to the categories to which they belong (banana → fruit). One way we have advanced beyond Hume is that we no longer think of the mind as going through a sequence of conscious ideas, one at a time. In the current view of how associative memory works, a great deal happens at once. An idea that has been activated does not merely evoke one other idea. It activates many ideas, which in turn activate others. Furthermore, only a few of the activated ideas will register in consciousness; most of the work of associative thinking is silent, hidden from our conscious selves. The notion that we have limited access to the workings of our minds is difficult to accept because, naturally, it is alien to our experience, but it is true: you know far less about yourself than you feel you do.

The Marvels of Priming

As is common in science, the first big breakthrough in our understanding of the mechanism of association was an improvement in a method of measurement. Until a few decades ago, the only way to study associations was to ask many people questions such as, “What is the first word that comes to your mind when you hear the word DAY?” The researchers tallied the frequency of responses, such as “night,” “sunny,” or “long.” In the 1980s, psychologists discovered that exposure to a word causes immediate and measurable changes in the ease with which many related words can be evoked. If you have recently seen or heard the word EAT, you are temporarily more likely to complete the word fragment SO_P as SOUP than as SOAP. The opposite would happen, of course, if you had just seen WASH. We call this a *priming effect* and say that the idea of EAT primes the idea of SOUP, and that WASH primes SOAP.

Priming effects take many forms. If the idea of EAT is currently on your mind (whether or not you are conscious of it), you will be quicker than usual to recognize the word SOUP when it is spoken in a whisper or presented in a blurry font. And of course you are primed not only for the idea of soup but also for a multitude of food-related ideas, including fork, hungry, fat, diet, and cookie. If for your most recent meal you sat at a wobbly restaurant table, you will be primed for wobbly as well. Furthermore, the primed ideas have some ability to prime other ideas, although more weakly. Like ripples on a pond, activation spreads through a small part of the vast network of associated ideas. The mapping of these ripples is now one of the most exciting pursuits in psychological research.

Another major advance in our understanding of memory was the discovery that priming is not restricted to concepts and words. You cannot know this from conscious experience, of course, but you must accept the alien idea that your actions and your emotions can be primed by events of which you are not even aware. In an experiment that became an instant classic, the psychologist John Bargh and his collaborators asked students at New York University—most aged eighteen to twenty-two—to assemble four-word sentences from a set of five words (for example, “finds he it yellow instantly”). For one group of students, half the scrambled sentences contained words associated with the elderly, such as *Florida*, *forgetful*, *bald*, *gray*, or *wrinkle*. When they had completed that task, the young participants were sent out to do another experiment in an office down the hall. That short walk was what the experiment was about. The researchers unobtrusively measured the time it took people to get from one end of the corridor to the other. As Bargh had predicted, the young people who had fashioned a sentence from words with an elderly theme walked down the hallway significantly more slowly than the others.

The “Florida effect” involves two stages of priming. First, the set of words primes thoughts of old age, though the word *old* is never mentioned; second, these thoughts prime a behavior, walking slowly, which is associated with old age. All this happens without any awareness. When they were questioned afterward, none of the students reported noticing that the words had had a common theme, and they all insisted that nothing they did after the first experiment could have been influenced by the words they had encountered. The idea of old age had not come to their conscious awareness, but their actions had changed nevertheless. This remarkable priming phenomenon—the influencing of an action by the idea—is known as the ideomotor effect. Although you surely were not aware of it, reading this paragraph primed you as well. If you had needed to stand up to get a glass of water, you would have been slightly slower than usual to rise from your chair—unless you happen to dislike the elderly, in which case research suggests that you might have been slightly faster than usual!

The ideomotor link also works in reverse. A study conducted in a German university was the mirror image of the early experiment that Bargh and his colleagues had carried out in New York. Students were asked to walk around a room for 5 minutes at a rate of 30 steps per minute, which was about one-third their normal pace. After this brief experience, the participants were much quicker to recognize words related to old age, such as *forgetful, old*, and *lonely*. Reciprocal priming effects tend to produce a coherent reaction: if you were primed to think of old age, you would tend to act old, and acting old would reinforce the thought of old age.

Reciprocal links are common in the associative network. For example, being amused tends to make you smile, and smiling tends to make you feel amused. Go ahead and take a pencil, and hold it between your teeth for a few seconds with the eraser pointing to your right and the point to your left. Now hold the pencil so the point is aimed straight in front of you, by pursing your lips around the eraser end. You were probably unaware that one of these actions forced your face into a frown and the other into a smile. College students were asked to rate the humor of cartoons from Gary Larson's *The Far Side* while holding a pencil in their mouth. Those who were "smiling" (without any awareness of doing so) found the cartoons rri221; (withfunnier than did those who were "frowning." In another experiment, people whose face was shaped into a frown (by squeezing their eyebrows together) reported an enhanced emotional response to upsetting pictures—starving children, people arguing, maimed accident victims.

Simple, common gestures can also unconsciously influence our thoughts and feelings. In one demonstration, people were asked to listen to messages through new headphones. They were told that the purpose of the experiment was to test the quality of the audio equipment and were instructed to move their heads repeatedly to check for any distortions of sound. Half the participants were told to nod their head up and down while others were told to shake it side to side. The messages they heard were radio editorials. Those who nodded (a yes gesture) tended to accept the message they heard, but those who shook their head tended to reject it. Again, there was no awareness, just a habitual connection between an attitude of rejection or acceptance and its common physical expression. You can see why the common admonition to "act calm and kind regardless of how you feel" is very good advice: you are likely to be rewarded by actually feeling calm and kind.

Primes That Guide Us

Studies of priming effects have yielded discoveries that threaten our self-image as conscious and autonomous authors of our judgments and our choices. For instance, most of us think of voting as a deliberate act that reflects our values and our assessments of policies and is not influenced by irrelevancies. Our vote should not be affected by the location of the polling station, for example, but it is. A study of voting patterns in precincts of Arizona in 2000 showed that the support for propositions to increase the funding of schools was significantly greater when the polling station was in a school than when it was in a nearby location. A separate experiment showed that exposing people to images of classrooms and school lockers also increased the tendency of participants to support a school initiative. The effect of the images was larger than the difference between parents and other voters! The study of priming has come some way from the initial demonstrations that reminding people of old age makes them walk more slowly. We now know that the effects of priming can reach into every corner of our lives.

Reminders of money produce some troubling effects. Participants in one experiment were shown a list of five words from which they were required to construct a four-word phrase that had a money theme ("high a salary desk paying" became "a high-paying salary"). Other primes were much more subtle, including the presence of an irrelevant money-related object in the background, such as a stack of Monopoly money on a table, or a computer with a screen saver of dollar bills floating in water.

Money-primed people become more independent than they would be without the associative trigger. They persevered almost twice as long in trying to solve a very difficult problem before they asked the experimenter for help, a crisp demonstration of increased self-reliance. Money-primed people are also more selfish: they were much less willing to spend time helping another student who pretended to be confused about an experimental task. When an experimenter clumsily dropped a bunch of pencils on the floor, the participants with money (unconsciously) on their mind picked up fewer pencils. In another experiment in the series, participants were told that they would shortly have a get-acquainted conversation with another person and were asked to set up two chairs while the experimenter left to retrieve that person. Participants primed by money chose in the exto stay much farther apart than their nonprimed peers (118 vs. 80 centimeters). Money-primed undergraduates also showed a greater preference for being alone.

The general theme of these findings is that the idea of money primes individualism: a reluctance to be involved with others, to depend on others, or to accept demands from others. The psychologist who has done this remarkable research, Kathleen Vohs, has been laudably restrained in discussing the implications of her findings, leaving the task to her readers. Her experiments are profound—her findings suggest that living in a culture that surrounds us with reminders of money may shape our behavior and our attitudes in ways that we do not know about and of which we may not be proud. Some cultures provide frequent reminders of respect, others constantly remind their members of God, and some societies prime obedience by large images of the Dear Leader. Can there be any doubt that the ubiquitous portraits of the national leader in dictatorial societies not only convey the feeling that "Big Brother Is Watching" but also lead to an actual reduction in spontaneous thought and independent action?

The evidence of priming studies suggests that reminding people of their mortality increases the appeal of authoritarian ideas, which may become reassuring in the context of the terror of death. Other experiments have confirmed Freudian insights about the role of symbols and metaphors in unconscious associations. For example, consider the ambiguous word fragments W_ _ H and S_ _ P. People who were recently asked to think of an action of which they are ashamed are more likely to complete those fragments as WASH and SOAP and less likely to see WISH and SOUP. Furthermore, merely thinking about stabbing a coworker in the back leaves people more inclined to buy soap, disinfectant, or detergent than batteries, juice, or candy bars. Feeling that one's soul is stained appears to trigger a desire to cleanse one's body, an impulse that has been dubbed the "Lady Macbeth effect."

The cleansing is highly specific to the body parts involved in a sin. Participants in an experiment were induced to "lie" to an imaginary person, either on the phone or in e-mail. In a subsequent test of the desirability of various products, people who had lied on the phone preferred mouthwash over soap, and those who had lied in e-mail preferred soap to mouthwash.

When I describe priming studies to audiences, the reaction is often disbelief. This is not a surprise: System 2 believes that it is in charge and that it knows the reasons for its choices. Questions are probably cropping up in your mind as well: How is it possible for such trivial manipulations of the context to have such large effects? Do these experiments demonstrate that we are completely at the mercy of whatever primes the environment provides at any moment? Of course not. The effects of the primes are robust but not necessarily large. Among a hundred voters, only a few whose initial preferences were uncertain will vote differently about a school issue if their precinct is located in a school rather than in a church—but a few percent could tip an election.

The idea you should focus on, however, is that disbelief is not an option. The results are not made up, nor are they statistical flukes. You have no choice but to accept that the major conclusions of these studies are true. More important, you must accept that they are true about *you*. If you had been exposed to a screen saver of floating dollar bills, you too would likely have picked up fewer pencils to help a clumsy stranger. You do not believe that these results apply to you because they correspond to nothing in your subjective experience. But your subjective expefteelief. Trience consists largely of the story that your System 2 tells itself about what is going on. Priming phenomena arise in System 1, and you have no conscious access to them.

I conclude with a perfect demonstration of a priming effect, which was conducted in an office kitchen at a British university. For many years members of that office had paid for the tea or coffee to which they helped themselves during the day by dropping money into an "honesty box." A list of suggested prices was posted. One day a banner poster was displayed just above the price list, with no warning or explanation. For a period of ten weeks a new image was presented each week, either flowers or eyes that appeared to be looking directly at the observer. No one commented on the new decorations, but the contributions to the honesty box changed significantly. The posters and the amounts that people put into the cash box (relative to the amount they consumed) are shown in figure 4. They deserve a close look.

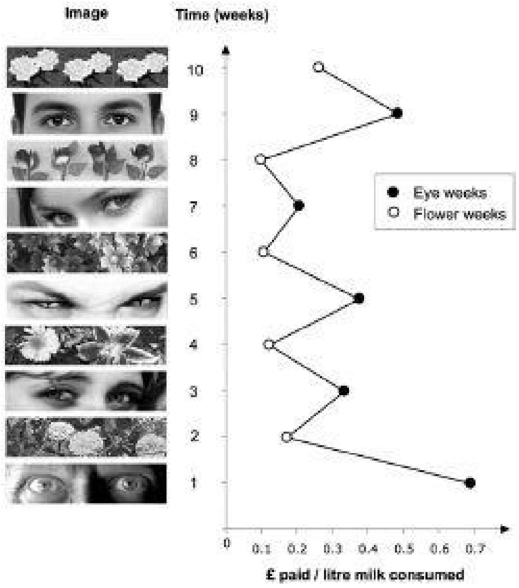


Figure 4

On the first week of the experiment (which you can see at the bottom of the figure), two wide-open eyes stare at the coffee or tea drinkers, whose average contribution was 70 pence per liter of milk. On week 2, the poster shows flowers and average contributions drop to about 15 pence. The trend continues. On average, the users of the kitchen contributed almost three times as much in "eye weeks" as they did in "flower weeks." Evidently, a purely symbolic reminder of being watched prodded people into improved behavior. As we expect at this point, the effect occurs without any awareness. Do you now believe that you would also fall into the same pattern?

Some years ago, the psychologist Timothy Wilson wrote a book with the evocative title *Strangers to Ourselves*. You have now been introduced to that stranger in you, which may be in control of much of what you do, although you rarely have a glimpse of it. System 1 provides the impressions that often turn into your beliefs, and is the source of the impulses that often become your choices and your actions. It offers a tacit interpretation of what happens to you and around you, linking the present with the recent past and with expectations about the near future. It contains the model of the world that instantly evaluates events as normal or surprising. It is the source of your rapid and often precise intuitive judgments. And it does most of this without your conscious awareness of its activities. System 1 is also, as we will see in the following chapters, the origin of many of the systematic errors in your intuitions.

Speaking of Priming

"The sight of all these people in uniforms does not prime creativity."

“The world makes much less sense than you think. The coherence comes mostly from the way your mind works.”

“They were primed to find flaws, and this is exactly what they found.”

“His System 1 constructed a story, and his System 2 believed it. It happens to allel

“I made myself smile and I’m actually feeling better!”

P

Cognitive Ease

Whenever you are conscious, and perhaps even when you are not, multiple computations are going on in your brain, which maintain and update current answers to some key questions: Is anything new going on? Is there a threat? Are things going well? Should my attention be redirected? Is more effort needed for this task? You can think of a cockpit, with a set of dials that indicate the current values of each of these essential variables. The assessments are carried out automatically by System 1, and one of their functions is to determine whether extra effort is required from System 2.

One of the dials measures *cognitive ease*, and its range is between “Easy” and “Strained.” Easy is a sign that things are going well—no threats, no major news, no need to redirect attention or mobilize effort. Strained indicates that a problem exists, which will require increased mobilization of System 2. Conversely, you experience *cognitive strain*. Cognitive strain is affected by both the current level of effort and the presence of unmet demands. The surprise is that a single dial of cognitive ease is connected to a large network of diverse inputs and outputs. Figure 5 tells the story.

The figure suggests that a sentence that is printed in a clear font, or has been repeated, or has been primed, will be fluently processed with cognitive ease. Hearing a speaker when you are in a good mood, or even when you have a pencil stuck crosswise in your mouth to make you “smile,” also induces cognitive ease. Conversely, you experience cognitive strain when you read instructions in a poor font, or in faint colors, or worded in complicated language, or when you are in a bad mood, and even when you frown.

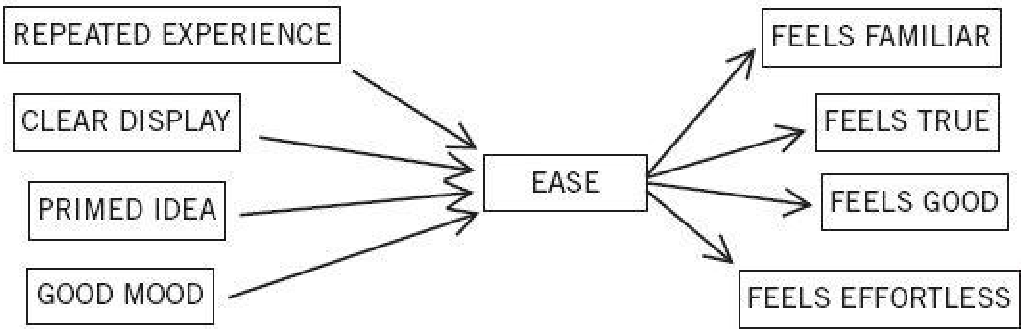


Figure 5. Causes and Consequences of Cognitive Ease

The various causes of ease or strain have interchangeable effects. When you are in a state of cognitive ease, you are probably in a good mood, like what you see, believe what you hear, trust your intuitions, and feel that the current situation is comfortably familiar. You are also likely to be relatively casual and superficial in your thinking. When you feel strained, you are more likely to be vigilant and suspicious, invest more effort in what you are doing, feel less comfortable, and make fewer errors, but you also are less intuitive and less creative than usual.

Illusions of Remembering

The word *illusion* brings visual illusions to mind, because we are all familiar with pictures that mislead. But vision is not the only domain of illusions; memory is also susceptible to them, as is thinking more generally.

David Stenbill, Monica Bigoutski, Sh”imight=s is pictana Tirana. I just made up these names. If you encounter any of them within the next few minutes you are likely to remember where you saw them. You know, and will know for a while, that these are not the names of minor celebrities. But suppose that a few days from now you are shown a long list of names, including some minor celebrities and “new” names of people that you have never heard of; your task will be to check every name of a celebrity in the list. There is a substantial probability that you will identify David Stenbill as a well-known person, although you will not (of course) know whether you encountered his name in the context of movies, sports, or politics. Larry Jacoby, the psychologist who first demonstrated this memory illusion in the laboratory, titled his article “Becoming Famous Overnight.” How does this happen? Start by asking yourself how you know whether or not someone is famous. In some cases of truly famous people (or of celebrities in an area you follow), you have a mental file with rich information about a person—think Albert Einstein, Bono, Hillary Clinton. But you will have no file of information about David Stenbill if you encounter his name in a few days. All you will have is a sense of familiarity—you have seen this name somewhere.

Jacoby nicely stated the problem: “The experience of familiarity has a simple but powerful quality of ‘pastness’ that seems to indicate that it is a direct reflection of prior experience.” This quality of pastness is an illusion. The truth is, as Jacoby and many followers have shown, that the name David Stenbill will look familiar when you see it *because you will see it more clearly*. Words that you have seen before become easier to see again—you can identify them better than other words when they are shown very briefly or masked by noise, and you will be quicker (by a few hundredths of a second) to read them than to read other words. In short, you experience greater cognitive ease in perceiving a word you have seen earlier, and it is this sense of ease that gives you the impression of familiarity.

Figure 5 suggests a way to test this. Choose a completely new word, make it easier to see, and it will be more likely to have the quality of pastness. Indeed, a new word is more likely to be recognized as familiar if it is unconsciously primed by showing it for a few milliseconds just before the test, or if it is shown in sharper contrast than some other words in the list. The link also operates in the other direction. Imagine you are shown a list of words that are more or less out of focus. Some of the words are severely blurred, others less so, and your task is to identify the words that are shown more clearly. A word that you have seen recently will appear to be clearer than unfamiliar words. As figure 5 indicates, the various ways of inducing cognitive ease or strain are interchangeable; you may not know precisely what it is that makes things cognitively easy or strained. This is how the illusion of familiarity comes about.

Illusions of Truth

"New York is a large city in the United States." "The moon revolves around Earth." "A chicken has four legs." In all these cases, you quickly retrieved a great deal of related information, almost all pointing one way or another. You knew soon after reading them that the first two statements are true and the last one is false. Note, however, that the statement "A chicken has three legs" is more obviously false than "A chicken has four legs." Your associative machinery slows the judgment of the latter sentence by delivering the fact that many animals have four legs, and perhaps also that supermarkets often sell chickenordblurred, legs in packages of four. System 2 was involved in sifting that information, perhaps raising the issue of whether the question about New York was too easy, or checking the meaning of *revolves*.

Think of the last time you took a driving test. Is it true that you need a special license to drive a vehicle that weighs more than three tons? Perhaps you studied seriously and can remember the side of the page on which the answer appeared, as well as the logic behind it. This is certainly not how I passed driving tests when I moved to a new state. My practice was to read the booklet of rules quickly once and hope for the best. I knew some of the answers from the experience of driving for a long time. But there were questions where no good answer came to mind, where all I had to go by was cognitive ease. If the answer felt familiar, I assumed that it was probably true. If it looked new (or improbably extreme), I rejected it. The impression of familiarity is produced by System 1, and System 2 relies on that impression for a true/false judgment.

The lesson of figure 5 is that predictable illusions inevitably occur if a judgment is based on an impression of cognitive ease or strain. Anything that makes it easier for the associative machine to run smoothly will also bias beliefs. A reliable way to make people believe in falsehoods is frequent repetition, because familiarity is not easily distinguished from truth. Authoritarian institutions and marketers have always known this fact. But it was psychologists who discovered that you do not have to repeat the entire statement of a fact or idea to make it appear true. People who were repeatedly exposed to the phrase "the body temperature of a chicken" were more likely to accept as true the statement that "the body temperature of a chicken is 144°" (or any other arbitrary number). The familiarity of one phrase in the statement sufficed to make the whole statement feel familiar, and therefore true. If you cannot remember the source of a statement, and have no way to relate it to other things you know, you have no option but to go with the sense of cognitive ease.

How to Write a Persuasive Message

Suppose you must write a message that you want the recipients to believe. Of course, your message will be true, but that is not necessarily enough for people to believe that it is true. It is entirely legitimate for you to enlist cognitive ease to work in your favor, and studies of *truth illusions* provide specific suggestions that may help you achieve this goal.

The general principle is that anything you can do to reduce cognitive strain will help, so you should first maximize legibility. Compare these two statements:

Adolf Hitler was born in 1892.

Adolf Hitler was born in 1887.

Both are false (Hitler was born in 1889), but experiments have shown that the first is more likely to be believed. More advice: if your message is to be printed, use high-quality paper to maximize the contrast between characters and their background. If you use color, you are more likely to be believed if your text is printed in bright blue or red than in middling shades of green, yellow, or pale blue.

If you care about being thought credible and intelligent, do not use complex language where simpler language will do. My Princeton ton colleague Danny Oppenheimer refuted a myth prevalent a wo ton colmong undergraduates about the vocabulary that professors find most impressive. In an article titled "Consequences of Erudite Vernacular Utilized Irrespective of Necessity: Problems with Using Long Words Needlessly," he showed that couching familiar ideas in pretentious language is taken as a sign of poor intelligence and low credibility.

In addition to making your message simple, try to make it memorable. Put your ideas in verse if you can; they will be more likely to be taken as truth. Participants in a much cited experiment read dozens of unfamiliar aphorisms, such as:

Woes unite foes.

Little strokes will tumble great oaks.

A fault confessed is half redressed.

Other students read some of the same proverbs transformed into nonrhyming versions:

Woes unite enemies.

Little strokes will tumble great trees.

A fault admitted is half redressed.

The aphorisms were judged more insightful when they rhymed than when they did not.

Finally, if you quote a source, choose one with a name that is easy to pronounce.

Participants in an experiment were asked to evaluate the prospects of fictitious Turkish companies on the basis of reports from two brokerage firms. For each stock, one of the reports came from an easily pronounced name (e.g., Artan) and the other report came from a firm with an unfortunate name (e.g., Taahhut). The reports sometimes disagreed. The best procedure for the observers would have been to average the two reports, but this is not what they did. They gave much more weight to the report from Artan than to the report from Taahhut. Remember that System 2 is lazy and that mental effort is aversive. If possible, the recipients of your message want to stay away from anything that reminds them of effort, including a source with a complicated name.

All this is very good advice, but we should not get carried away. High-quality paper, bright colors, and rhyming or simple language will not be much help if your message is obviously nonsensical, or if it contradicts facts that your audience knows to be true. The psychologists who do these experiments do not believe that people are stupid or infinitely gullible. What psychologists do believe is that all of us live much of our life guided by the impressions of System 1—and we often do not know the source of these impressions. How do you know that a statement is true? If it is strongly linked by logic or association to other beliefs or preferences you hold, or comes from a source you trust and like, you will feel a sense of cognitive ease. The trouble is that there may be other causes for your feeling of ease—including the quality of the font and the appealing rhythm of the prose—and you have no simple way of tracing your feelings to their source. This is the message of figure 5: the sense of ease or strain has multiple causes, and it is difficult to tease them apart. Difficult, but not impossible. People can overcome some of the superficial factors that produce illusions of truth when strongly motivated to do so. On most occasions, however, the lazy System 2 will adopt the suggestions of System 1 and march on.

Strain and Effort

The symmetry of many associative connections was a dominant theme in the discussion of associative coherence. As we saw earlier, people who are made to “smile” or “frown” by sticking a pencil in their mouth or holding a ball between their furrowed brows are prone to experience the emotions that frowning and smiling normally express. The same self-reinforcing reciprocity is found in studies of cognitive ease. On the one hand, cognitive strain is experienced when the effortful operations of System 2 are engaged. On the other hand, the experience of cognitive strain, whatever its source, tends to mobilize System 2, shifting people’s approach to problems from a casual intuitive mode to a more engaged and analytic mode.

The bat-and-ball problem was mentioned earlier as a test of people’s tendency to answer questions with the first idea that comes to their mind, without checking it. Shane Frederick’s Cognitive Reflection Test consists of the bat-and-ball problem and two others, all chosen because they evoke an immediate intuitive answer that is incorrect. The other two items in the CRT are:

> If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?
>
> 100 minutes OR 5 minutes

In a lake, there is a patch of lily pads. Every day, the patch doubles in size.

If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?

24 days OR 47 days

The correct answers to both problems are in a footnote at the bottom of the page.[*] The experimenters recruited 40 Princeton students to take the CRT. Half of them saw the puzzles in a small font in washed-out gray print. The puzzles were legible, but the font induced cognitive strain. The results tell a clear story: 90% of the students who saw the CRT in normal font made at least one mistake in the test, but the proportion dropped to 35% when the font was barely legible. You read this correctly: performance was better with the bad font. Cognitive strain, whatever its source, mobilizes System 2, which is more likely to reject the intuitive answer suggested by System 1.

The Pleasure of Cognitive Ease

An article titled "Mind at Ease Puts a Smile on the Face" describes an experiment in which participants were briefly shown pictures of objects. Some of these pictures were made easier to recognize by showing the outline of the object just before the complete image was shown, so briefly that the contours were never noticed. Emotional reactions were measured by recording electrical impulses from facial muscles, registering changes of expression that are too slight and too brief to be detectable by observers. As expected, people showed a faint smile and relaxed brows when the pictures were easier to see. It appears to be a feature of System 1 that cognitive ease is associated with good feelings.

As expected, easily pronounced words evoke a favorable attitude. Companies with pronounceable names dmisorrectlo better than others for the first week after the stock is issued, though the effect disappears over time. Stocks with pronounceable trading symbols (like KAR or LUNMOO) outperform those with tongue-twisting tickers like PXG or RDO —and they appear to retain a small advantage over some time. A study conducted in Switzerland found that investors believe that stocks with fluent names like Emmi, Swissfirst, and Comet will earn higher returns than those with clunky labels like Geberit and Ypsomed.

As we saw in figure 5, repetition induces cognitive ease and a comforting feeling of familiarity. The famed psychologist Robert Zajonc dedicated much of his career to the study of the link between the repetition of an arbitrary stimulus and the mild affection that people eventually have for it. Zajonc called it the *mere exposure effect*. A demonstration conducted in the student newspapers of the University of Michigan and of Michigan State University is one of my favorite experiments. For a period of some weeks, an ad-like box appeared on the front page of the paper, which contained one of the following Turkish (or

Turkish-sounding) words: *kadirga*, *saricik*, *biwonjni*, *nansoma*, and *iktitaf*. The frequency with which the words were repeated varied: one of the words was shown only once, the others appeared on two, five, ten, or twenty-five separate occasions. (The words that were presented most often in one of the university papers were the least frequent in the other.) No explanation was offered, and readers' queries were answered by the statement that "the purchaser of the display wished for anonymity."

When the mysterious series of ads ended, the investigators sent questionnaires to the university communities, asking for impressions of whether each of the words "means something 'good' or something 'bad.'" The results were spectacular: the words that were presented more frequently were rated much more favorably than the words that had been shown only once or twice. The finding has been confirmed in many experiments, using Chinese ideographs, faces, and randomly shaped polygons.

The mere exposure effect does not depend on the conscious experience of familiarity. In fact, the effect does not depend on consciousness at all: it occurs even when the repeated words or pictures are shown so quickly that the observers never become aware of having seen them. They still end up liking the words or pictures that were presented more frequently. As should be clear by now, System 1 can respond to impressions of events of which System 2 is unaware. Indeed, the mere exposure effect is actually stronger for stimuli that the individual never consciously sees.

Zajonc argued that the effect of repetition on liking is a profoundly important biological fact, and that it extends to all animals. To survive in a frequently dangerous world, an organism should react cautiously to a novel stimulus, with withdrawal and fear. Survival prospects are poor for an animal that is not suspicious of novelty. However, it is also adaptive for the initial caution to fade if the stimulus is actually safe. The mere exposure effect occurs, Zajonc claimed, because the repeated exposure of a stimulus is followed by nothing bad. Such a stimulus will eventually become a safety signal, and safety is good. Obviously, this argument is not restricted to humans. To make that point, one of Zajonc's associates exposed two sets of fertile chicken eggs to different tones. After they hatched, the chicks consistently emitted fewer distress calls when exposed to the tone they had heard while inhabiting the shell.

Zajonc offered an eloquent summary of hing icts program of research:

> The consequences of repeated exposures benefit the organism in its relations to the immediate animate and inanimate environment. They allow the organism to distinguish objects and habitats that are safe from those that are not, and they are the most primitive basis of social attachments. Therefore, they form the basis for social organization and cohesion—the basic sources of psychological and social stability.

The link between positive emotion and cognitive ease in System 1 has a long evolutionary history.

Ease, Mood, and Intuition

Around 1960, a young psychologist named Sarnoff Mednick thought he had identified the essence of creativity. His idea was as simple as it was powerful: creativity is associative memory that works exceptionally well. He made up a test, called the Remote Association Test (RAT), which is still often used in studies of creativity.

For an easy example, consider the following three words:

cottage Swiss cake

Can you think of a word that is associated with all three? You probably worked out that the answer is *cheese*. Now try this:

dive light rocket

This problem is much harder, but it has a unique correct answer, which every speaker of English recognizes, although less than 20% of a sample of students found it within 15 seconds. The answer is *sky*. Of course, not every triad of words has a solution. For example, the words *dream*, *ball*, *book* do not have a shared association that everyone will recognize as valid.

Several teams of German psychologists that have studied the RAT in recent years have come up with remarkable discoveries about cognitive ease. One of the teams raised two questions: Can people feel that a triad of words has a solution before they know what the solution is? How does mood influence performance in this task? To find out, they first made some of their subjects happy and others sad, by asking them to think for several minutes about happy or sad episodes in their lives. Then they presented these subjects with a series of triads, half of them linked (such as *dive*, *light*, *rocket*) and half unlinked (such as *dream*, *ball*, *book*), and instructed them to press one of two keys very quickly to indicate their guess about whether the triad was linked. The time allowed for this guess, 2 seconds, was much too short for the actual solution to come to anyone's mind.

The first surprise is that people's guesses are much more accurate than they would be by chance. I find this astonishing. A sense of cognitive ease is apparently generated by a very faint signal from the associative machine, which "knows" that the three words are coherent (share an association) long before the association is retrieved. The role of cognitive ease in the judgment was confirmed experimentally by another German team: manipulations that increase cognitive ease (priming, a clear font, pre-exposing words) all increase the tendency to see the words as linked.

Another remarkable discovery is the powerful effect of mood on this intuitive performance. The experimentershape tende computed an "intuition index" to measure accuracy. They found that putting the participants in a good mood before the test by having them think happy thoughts more than doubled accuracy. An even more striking result is that unhappy subjects were completely incapable of performing the intuitive task accurately; their guesses were no better than random. Mood evidently affects the operation of System 1: when we are uncomfortable and unhappy, we lose touch with our intuition.

These findings add to the growing evidence that good mood, intuition, creativity, gullibility, and increased reliance on System 1 form a cluster. At the other pole, sadness, vigilance, suspicion, an analytic approach, and increased effort also go together. A happy mood loosens the control of System 2 over performance: when in a good mood, people become more intuitive and more creative but also less vigilant and more prone to logical errors. Here again, as in the mere exposure effect, the connection makes biological sense. A good mood is a signal that things are generally going well, the environment is safe, and it is all right to let one's guard down. A bad mood indicates that things are not going very well, there may be a threat, and vigilance is required. Cognitive ease is both a cause and a consequence of a pleasant feeling.

The Remote Association Test has more to tell us about the link between cognitive ease and positive affect. Briefly consider two triads of words:

sleep mail switch

salt deep foam

You could not know it, of course, but measurements of electrical activity in the muscles of your face would probably have shown a slight smile when you read the second triad, which is coherent (*sea* is the solution). This smiling reaction to coherence appears in subjects who are told nothing about common associates; they are merely shown a vertically arranged triad of words and instructed to press the space bar after they have read it. The impression of cognitive ease that comes with the presentation of a coherent triad appears to be mildly pleasurable in itself.

The evidence that we have about good feelings, cognitive ease, and the intuition of coherence is, as scientists say, correlational but not necessarily causal. Cognitive ease and smiling occur together, but do the good feelings actually lead to intuitions of coherence? Yes, they do. The proof comes from a clever experimental approach that has become increasingly popular. Some participants were given a cover story that provided an alternative interpretation for their good feeling: they were told about music played in their earphones that "previous research showed that this music influences the emotional reactions of individuals." This story completely eliminates the intuition of coherence. The finding shows that the brief emotional response that follows the presentation of a triad of words (pleasant if the triad is coherent, unpleasant otherwise) is actually the basis of judgments of coherence. There is nothing here that System 1 cannot do. Emotional changes are now expected, and because they are unsurprising they are not linked causally to the words.

This is as good as psychological research ever gets, in its combination of experimental techniques and in its results, which are both robust and extremely surprising. We have learned a great deal about the automatic workings of System 1 in the last decades. Much of what we now know would have sounded like science fiction thirty or forty years ago. It was beyond imagining that bad font influences judgments of truth and improves cognitive performance, or that an emotional response to the cognitive ease of a tri pr that aad of words mediates impressions of coherence. Psychology has come a long way.

Speaking of Cognitive Ease

"Let's not dismiss their business plan just because the font makes it hard to read."

“We must be inclined to believe it because it has been repeated so often, but let’s think it through again.”

“Familiarity breeds liking. This is a mere exposure effect.”

“I’m in a very good mood today, and my System 2 is weaker than usual. I should be extra careful.”

P

Norms, Surprises, and Causes

The central characteristics and functions of System 1 and System 2 have now been introduced, with a more detailed treatment of System 1. Freely mixing metaphors, we have in our head a remarkably powerful computer, not fast by conventional hardware standards, but able to represent the structure of our world by various types of associative links in a vast network of various types of ideas. The spreading of activation in the associative machine is automatic, but we (System 2) have some ability to control the search of memory, and also to program it so that the detection of an event in the environment can attract attention. We next go into more detail of the wonders and limitation of what System 1 can do.

Assessing Normality

The main function of System 1 is to maintain and update a model of your personal world, which represents what is normal in it. The model is constructed by associations that link ideas of circumstances, events, actions, and outcomes that co-occur with some regularity, either at the same time or within a relatively short interval. As these links are formed and strengthened, the pattern of associated ideas comes to represent the structure of events in your life, and it determines your interpretation of the present as well as your expectations of the future.

A capacity for surprise is an essential aspect of our mental life, and surprise itself is the most sensitive indication of how we understand our world and what we expect from it. There are two main varieties of surprise. Some expectations are active and conscious—you know you are waiting for a particular event to happen. When the hour is near, you may be expecting the sound of the door as your child returns from school; when the door opens you expect the sound of a familiar voice. You will be surprised if an actively expected event does not occur. But there is a much larger category of events that you expect passively; you don't wait for them, but you are not surprised when they happen. These are events that are normal in a situation, though not sufficiently probable to be actively expected.

A single incident may make a recurrence less surprising. Some years ago, my wife and I were of dealWhen normvacationing in a small island resort on the Great Barrier Reef. There are only forty guest rooms on the island. When we came to dinner, we were surprised to meet an acquaintance, a psychologist named Jon. We greeted each other warmly and commented on the coincidence. Jon left the resort the next day. About two weeks later, we were in a theater in London. A latecomer sat next to me after the lights went down. When the lights came up for the intermission, I saw that my neighbor was Jon. My wife and I commented later that we were simultaneously conscious of two facts: first, this was a more remarkable coincidence than the first meeting; second, we were distinctly *less* surprised to meet Jon on the second occasion than we had been on the first. Evidently, the first meeting had somehow changed the idea of Jon in our minds. He was now “the psychologist who shows up when we travel abroad.” We (System 2) knew this was a ludicrous idea, but our System 1 had made it seem almost normal to meet Jon in strange places. We would have experienced much more surprise if we had met any acquaintance other than Jon in the next seat of a London theater. By any measure of probability, meeting Jon in the theater was much less likely than meeting any one of our hundreds of acquaintances—yet meeting Jon seemed more normal.

Under some conditions, passive expectations quickly turn active, as we found in another coincidence. On a Sunday evening some years ago, we were driving from New York City to Princeton, as we had been doing every week for a long time. We saw an unusual sight: a car on fire by the side of the road. When we reached the same stretch of road the following Sunday, another car was burning there. Here again, we found that we were distinctly less surprised on the second occasion than we had been on the first. This was now “the place where cars catch fire.” Because the circumstances of the recurrence were the same, the second incident was sufficient to create an active expectation: for months, perhaps for years, after the event we were reminded of burning cars whenever we reached that spot of the road and were quite prepared to see another one (but of course we never did).

The psychologist Dale Miller and I wrote an essay in which we attempted to explain how events come to be perceived as normal or abnormal. I will use an example from our description of “norm theory,” although my interpretation of it has changed slightly:

> An observer, casually watching the patrons at a neighboring table in a fashionable restaurant, notices that the first guest to taste the soup winces, as if in pain. The normality of a multitude of events will be altered by this incident. It is now unsurprising for the guest who first tasted the soup to startle violently when touched by a waiter; it is also unsurprising for another guest to stifle a cry when tasting soup from the same tureen. These events and many others appear more normal than they would have otherwise, but not necessarily because they confirm advance expectations. Rather, they appear normal because they recruit the original episode, retrieve it from memory, and are interpreted in conjunction with it.

Imagine yourself the observer at the restaurant. You were surprised by the first guest’s unusual reaction to the soup, and surprised again by the startled response to the waiter’s touch. However, the second abnormal event will retrieve the first from memory, and both make sense together. The two events fit into a pattern, in which the guest is an exceptionally tense person. On the other hand, if the next thing that happens after the first guest’s grimace is that another customer rejects the soup, these two surprises will be linked and thehinsur soup will surely be blamed.

“How many animals of each kind did Moses take into the ark?” The number of people who detect what is wrong with this question is so small that it has been dubbed the “Moses illusion." Moses took no animals into the ark; Noah did. Like the incident of the wincing soup eater, the Moses illusion is readily explained by norm theory. The idea of animals going into the ark sets up a biblical context, and Moses is not abnormal in that context. You did not positively expect him, but the mention of his name is not surprising. It also helps that Moses and Noah have the same vowel sound and number of syllables. As with the triads that produce cognitive ease, you unconsciously detect associative coherence between "Moses" and "ark" and so quickly accept the question. Replace Moses with George W. Bush in this sentence and you will have a poor political joke but no illusion.

When something cement does not fit into the current context of activated ideas, the system detects an abnormality, as you just experienced. You had no particular idea of what was coming after *something*, but you knew when the word *cement* came that it was abnormal in that sentence. Studies of brain responses have shown that violations of normality are detected with astonishing speed and subtlety. In a recent experiment, people heard the sentence "Earth revolves around the trouble every year." A distinctive pattern was detected in brain activity, starting within two-tenths of a second of the onset of the odd word. Even more remarkable, the same brain response occurs at the same speed when a male voice says, "I believe I am pregnant because I feel sick every morning," or when an upper-class voice says, "I have a large tattoo on my back." A vast amount of world knowledge must instantly be brought to bear for the incongruity to be recognized: the voice must be identified as upper-class English and confronted with the generalization that large tattoos are uncommon in the upper class.

We are able to communicate with each other because our knowledge of the world and our use of words are largely shared. When I mention a table, without specifying further, you understand that I mean a normal table. You know with certainty that its surface is approximately level and that it has far fewer than 25 legs. We have *norms* for a vast number of categories, and these norms provide the background for the immediate detection of anomalies such as pregnant men and tattooed aristocrats.

To appreciate the role of norms in communication, consider the sentence "The large mouse climbed over the trunk of the very small elephant." I can count on your having norms for the size of mice and elephants that are not too far from mine. The norms specify a typical or average size for these animals, and they also contain information about the range or variability within the category. It is very unlikely that either of us got the image in our mind's eye of a mouse larger than an elephant striding over an elephant smaller than a mouse. Instead, we each separately but jointly visualized a mouse smaller than a shoe clambering over an elephant larger than a sofa. System 1, which understands language, has access to norms of categories, which specify the range of plausible values as well as the most typical cases.

Seeing Causes and Intentions

"Fred's parents arrived late. The caterers were expected soon. Fred was angry." You know why Fred was angry, and it is not because the caterers were expected soon. In your network of associationsmals in co, anger and lack of punctuality are linked as an effect and its possible cause, but there is no such link between anger and the idea of expecting caterers. A coherent story was instantly constructed as you read; you immediately knew the cause of Fred’s anger. Finding such causal connections is part of understanding a story and is an automatic operation of System 1. System 2, your conscious self, was offered the causal interpretation and accepted it.

A story in Nassim Taleb’s *The Black Swan* illustrates this automatic search for causality. He reports that bond prices initially rose on the day of Saddam Hussein’s capture in his hiding place in Iraq. Investors were apparently seeking safer assets that morning, and the Bloomberg News service flashed this headline: U.S. TREASURIES RISE; HUSSEIN CAPTURE MAY NOT CURB TERRORISM. Half an hour later, bond prices fell back and the revised headline read: U.S. TREASURIES FALL; HUSSEIN CAPTURE BOOSTS ALLURE OF RISKY ASSETS. Obviously, Hussein’s capture was the major event of the day, and because of the way the automatic search for causes shapes our thinking, that event was destined to be the explanation of whatever happened in the market on that day. The two headlines look superficially like explanations of what happened in the market, but a statement that can explain two contradictory outcomes explains nothing at all. In fact, all the headlines do is satisfy our need for coherence: a large event is supposed to have consequences, and consequences need causes to explain them. We have limited information about what happened on a day, and System 1 is adept at finding a coherent causal story that links the fragments of knowledge at its disposal.

Read this sentence:

> After spending a day exploring beautiful sights in the crowded streets of New York, Jane discovered that her wallet was missing.

When people who had read this brief story (along with many others) were given a surprise recall test, the word *pickpocket* was more strongly associated with the story than the word *sights*, even though the latter was actually in the sentence while the former was not. The rules of associative coherence tell us what happened. The event of a lost wallet could evoke many different causes: the wallet slipped out of a pocket, was left in the restaurant, etc. However, when the ideas of lost wallet, New York, and crowds are juxtaposed, they jointly evoke the explanation that a pickpocket caused the loss. In the story of the startling soup, the outcome—whether another customer wincing at the taste of the soup or the first person’s extreme reaction to the waiter’s touch—brings about an associatively coherent interpretation of the initial surprise, completing a plausible story.

The aristocratic Belgian psychologist Albert Michotte published a book in 1945 (translated into English in 1963) that overturned centuries of thinking about causality, going back at least to Hume’s examination of the association of ideas. The commonly accepted wisdom was that we infer physical causality from repeated observations of correlations among events. We have had myriad experiences in which we saw one object in motion touching another object, which immediately starts to move, often (but not always) in the same direction. This is what happens when a billiard ball hits another, and it is also what happens when you knock over a vase by brushing against it. Michotte had a different idea: he argued that we *see* causality, just as directly as we see color. To make his point, he created episodes in n ttiowhich a black square drawn on paper is seen in motion; it comes into contact with another square, which immediately begins to move. The observers know that there is no real physical contact, but they nevertheless have a powerful “illusion of causality.” If the second object starts moving instantly, they describe it as having been “launched” by the first. Experiments have shown that six-month-old infants see the sequence of events as a cause-effect scenario, and they indicate surprise when the sequence is altered. We are evidently ready from birth to have *impressions* of causality, which do not depend on reasoning about patterns of causation. They are products of System 1.

In 1944, at about the same time as Michotte published his demonstrations of physical causality, the psychologists Fritz Heider and Mary-Ann Simmel used a method similar to Michotte’s to demonstrate the perception of *intentional* causality. They made a film, which lasts all of one minute and forty seconds, in which you see a large triangle, a small triangle, and a circle moving around a shape that looks like a schematic view of a house with an open door. Viewers see an aggressive large triangle bullying a smaller triangle, a terrified circle, the circle and the small triangle joining forces to defeat the bully; they also observe much interaction around a door and then an explosive finale. The perception of intention and emotion is irresistible; only people afflicted by autism do not experience it. All this is entirely in your mind, of course. Your mind is ready and even eager to identify agents, assign them personality traits and specific intentions, and view their actions as expressing individual propensities. Here again, the evidence is that we are born prepared to make intentional attributions: infants under one year old identify bullies and victims, and expect a pursuer to follow the most direct path in attempting to catch whatever it is chasing.

The experience of freely willed action is quite separate from physical causality. Although it is your hand that picks up the salt, you do not think of the event in terms of a chain of physical causation. You experience it as caused by a decision that a disembodied *you* made, because you wanted to add salt to your food. Many people find it natural to describe their soul as the source and the cause of their actions. The psychologist Paul Bloom, writing in *The Atlantic* in 2005, presented the provocative claim that our inborn readiness to separate physical and intentional causality explains the near universality of religious beliefs. He observes that “we perceive the world of objects as essentially separate from the world of minds, making it possible for us to envision soulless bodies and bodiless souls.” The two modes of causation that we are set to perceive make it natural for us to accept the two central beliefs of many religions: an immaterial divinity is the ultimate cause of the physical world, and immortal souls temporarily control our bodies while we live and leave them behind as we die. In Bloom’s view, the two concepts of causality were shaped separately by evolutionary forces, building the origins of religion into the structure of System 1.

The prominence of causal intuitions is a recurrent theme in this book because people are prone to apply causal thinking inappropriately, to situations that require statistical reasoning. Statistical thinking derives conclusions about individual cases from properties of categories and ensembles. Unfortunately, System 1 does not have the capability for this mode of reasoning; System 2 can learn to think statistically, but few people receive the necessary training.

The psychology of causality was the basis of my decision to describe psycl c to thinhological processes by metaphors of agency, with little concern for consistency. I sometimes refer to System 1 as an agent with certain traits and preferences, and sometimes as an associative machine that represents reality by a complex pattern of links. The system and the machine are fictions; my reason for using them is that they fit the way we think about causes. Heider's triangles and circles are not really agents—it is just very easy and natural to think of them that way. It is a matter of mental economy. I assume that you (like me) find it easier to think about the mind if we describe what happens in terms of traits and intentions (the two systems) and sometimes in terms of mechanical regularities (the associative machine). I do not intend to convince you that the systems are real, any more than Heider intended you to believe that the large triangle is really a bully.

Speaking of Norms and Causes

"When the second applicant also turned out to be an old friend of mine, I wasn't quite as surprised. Very little repetition is needed for a new experience to feel normal!"

"When we survey the reaction to these products, let's make sure we don't focus exclusively on the average. We should consider the entire range of normal reactions."

"She can't accept that she was just unlucky; she needs a causal story. She will end up thinking that someone intentionally sabotaged her work."

P

A Machine for Jumping to Conclusions

The great comedian Danny Kaye had a line that has stayed with me since my adolescence. Speaking of a woman he dislikes, he says, “Her favorite position is beside herself, and her favorite sport is jumping to conclusions.” The line came up, I remember, in the initial conversation with Amos Tversky about the rationality of statistical intuitions, and now I believe it offers an apt description of how System 1 functions. Jumping to conclusions is efficient if the conclusions are likely to be correct and the costs of an occasional mistake acceptable, and if the jump saves much time and effort. Jumping to conclusions is risky when the situation is unfamiliar, the stakes are high, and there is no time to collect more information. These are the circumstances in which intuitive errors are probable, which may be prevented by a deliberate intervention of System 2.

Neglect of Ambiguity and Suppression of Doubt

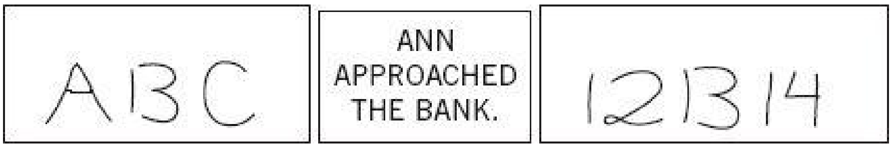


Figure 6

What do the three exhibits in figure 6 have in common? The answer is that all are ambiguous. You almost certainly read the display on the left as A B C and the one on the right as 12 13 14, but the middle items in both displays are identical. You could just as well have read e iom prthe cve them as A 13 C or 12 B 14, but you did not. Why not? The same shape is read as a letter in a context of letters and as a number in a context of numbers. The entire context helps determine the interpretation of each element. The shape is ambiguous, but you jump to a conclusion about its identity and do not become aware of the ambiguity that was resolved.

As for Ann, you probably imagined a woman with money on her mind, walking toward a building with tellers and secure vaults. But this plausible interpretation is not the only possible one; the sentence is ambiguous. If an earlier sentence had been “They were floating gently down the river,” you would have imagined an altogether different scene. When you have just been thinking of a river, the word *bank* is not associated with money. In the absence of an explicit context, System 1 generated a likely context on its own. We know that it is System 1 because you were not aware of the choice or of the possibility of another interpretation. Unless you have been canoeing recently, you probably spend more time going to banks than floating on rivers, and you resolved the ambiguity accordingly. When uncertain, System 1 bets on an answer, and the bets are guided by experience. The rules of the betting are intelligent: recent events and the current context have the most weight in determining an interpretation. When no recent event comes to mind, more distant memories govern. Among your earliest and most memorable experiences was singing your ABCs; you did not sing your A13Cs.

The most important aspect of both examples is that a definite choice was made, but you did not know it. Only one interpretation came to mind, and you were never aware of the ambiguity. System 1 does not keep track of alternatives that it rejects, or even of the fact that there were alternatives. Conscious doubt is not in the repertoire of System 1; it requires maintaining incompatible interpretations in mind at the same time, which demands mental effort. Uncertainty and doubt are the domain of System 2.

A Bias to Believe and Confirm

The psychologist Daniel Gilbert, widely known as the author of *Stumbling to Happiness*, once wrote an essay, titled "How Mental Systems Believe," in which he developed a theory of believing and unbelieving that he traced to the seventeenth-century philosopher Baruch Spinoza. Gilbert proposed that understanding a statement must begin with an attempt to believe it: you must first know what the idea would mean if it were true. Only then can you decide whether or not to *unbelieve* it. The initial attempt to believe is an automatic operation of System 1, which involves the construction of the best possible interpretation of the situation. Even a nonsensical statement, Gilbert argues, will evoke initial belief. Try his example: "whitefish eat candy." You probably were aware of vague impressions of fish and candy as an automatic process of associative memory searched for links between the two ideas that would make sense of the nonsense.

Gilbert sees unbelieving as an operation of System 2, and he reported an elegant experiment to make his point. The participants saw nonsensical assertions, such as "a dinca is a flame," followed after a few seconds by a single word, "true" or "false." They were later tested for their memory of which sentences had been labeled "true." In one condition of the experiment subjects were required to hold digits in memory during the task. The disruption of System 2 had a selective effect: it made it difficult for people to "unbelieve" false sentences. In a later test of memory, the depleted par muumbling toticipants ended up thinking that many of the false sentences were true. The moral is significant: when System 2 is otherwise engaged, we will believe almost anything. System 1 is gullible and biased to believe, System 2 is in charge of doubting and unbelieving, but System 2 is sometimes busy, and often lazy. Indeed, there is evidence that people are more likely to be influenced by empty persuasive messages, such as commercials, when they are tired and depleted.

The operations of associative memory contribute to a general *confirmation bias*. When asked, "Is Sam friendly?" different instances of Sam's behavior will come to mind than would if you had been asked "Is Sam unfriendly?" A deliberate search for confirming evidence, known as *positive test strategy*, is also how System 2 tests a hypothesis. Contrary to the rules of philosophers of science, who advise testing hypotheses by trying to refute them, people (and scientists, quite often) seek data that are likely to be compatible with the beliefs they currently hold. The confirmatory bias of System 1 favors uncritical acceptance of suggestions and exaggeration of the likelihood of extreme and improbable events. If you are asked about the probability of a tsunami hitting California within the next thirty years, the images that come to your mind are likely to be images of tsunamis, in the manner Gilbert proposed for nonsense statements such as "whitefish eat candy." You will be prone to overestimate the probability of a disaster.

Exaggerated Emotional Coherence (Halo Effect)

If you like the president's politics, you probably like his voice and his appearance as well. The tendency to like (or dislike) everything about a person—including things you have not observed—is known as the halo effect. The term has been in use in psychology for a century, but it has not come into wide use in everyday language. This is a pity, because the halo effect is a good name for a common bias that plays a large role in shaping our view of people and situations. It is one of the ways the representation of the world that System 1 generates is simpler and more coherent than the real thing.

You meet a woman named Joan at a party and find her personable and easy to talk to. Now her name comes up as someone who could be asked to contribute to a charity. What do you know about Joan's generosity? The correct answer is that you know virtually nothing, because there is little reason to believe that people who are agreeable in social situations are also generous contributors to charities. But you like Joan and you will retrieve the feeling of liking her when you think of her. You also like generosity and generous people. By association, you are now predisposed to believe that Joan is generous. And now that you believe she is generous, you probably like Joan even better than you did earlier, because you have added generosity to her pleasant attributes.

Real evidence of generosity is missing in the story of Joan, and the gap is filled by a guess that fits one's emotional response to her. In other situations, evidence accumulates gradually and the interpretation is shaped by the emotion attached to the first impression. In an enduring classic of psychology, Solomon Asch presented descriptions of two people and asked for comments on their personality. What do you think of Alan and Ben?

Alan: intelligent—industrious—impulsive—critical—stubborn—envious

Ben: envious—The#82stubborn—critical—impulsive—industrious—intelligent

If you are like most of us, you viewed Alan much more favorably than Ben. The initial traits in the list change the very meaning of the traits that appear later. The stubbornness of an intelligent person is seen as likely to be justified and may actually evoke respect, but intelligence in an envious and stubborn person makes him more dangerous. The halo effect is also an example of suppressed ambiguity: like the word *bank*, the adjective *stubborn* is ambiguous and will be interpreted in a way that makes it coherent with the context.

There have been many variations on this research theme. Participants in one study first considered the first three adjectives that describe Alan; then they considered the last three, which belonged, they were told, to another person. When they had imagined the two individuals, the participants were asked if it was plausible for all six adjectives to describe the same person, and most of them thought it was impossible!

The sequence in which we observe characteristics of a person is often determined by chance. Sequence matters, however, because the halo effect increases the weight of first impressions, sometimes to the point that subsequent information is mostly wasted. Early in my career as a professor, I graded students' essay exams in the conventional way. I would pick up one test booklet at a time and read all that student's essays in immediate succession, grading them as I went. I would then compute the total and go on to the next student. I eventually noticed that my evaluations of the essays in each booklet were strikingly homogeneous. I began to suspect that my grading exhibited a halo effect, and that the first question I scored had a disproportionate effect on the overall grade. The mechanism was simple: if I had given a high score to the first essay, I gave the student the benefit of the doubt whenever I encountered a vague or ambiguous statement later on. This seemed reasonable. Surely a student who had done so well on the first essay would not make a foolish mistake in the second one! But there was a serious problem with my way of doing things. If a student had written two essays, one strong and one weak, I would end up with different final grades depending on which essay I read first. I had told the students that the two essays had equal weight, but that was not true: the first one had a much greater impact on the final grade than the second. This was unacceptable.

I adopted a new procedure. Instead of reading the booklets in sequence, I read and scored all the students' answers to the first question, then went on to the next one. I made sure to write all the scores on the inside back page of the booklet so that I would not be biased (even unconsciously) when I read the second essay. Soon after switching to the new method, I made a disconcerting observation: my confidence in my grading was now much lower than it had been. The reason was that I frequently experienced a discomfort that was new to me. When I was disappointed with a student's second essay and went to the back page of the booklet to enter a poor grade, I occasionally discovered that I had given a top grade to the same student's first essay. I also noticed that I was tempted to reduce the discrepancy by changing the grade that I had not yet written down, and found it hard to follow the simple rule of never yielding to that temptation. My grades for the essays of a single student often varied over a considerable range. The lack of coherence left me uncertain and frustrated.

I was now less happy with and less confident in my grades than I had been earlier, but I recognized that thass confthis was a good sign, an indication that the new procedure was superior. The consistency I had enjoyed earlier was spurious; it produced a feeling of cognitive ease, and my System 2 was happy to lazily accept the final grade. By allowing myself to be strongly influenced by the first question in evaluating subsequent ones, I spared myself the dissonance of finding the same student doing very well on some questions and badly on others. The uncomfortable inconsistency that was revealed when I switched to the new procedure was real: it reflected both the inadequacy of any single question as a measure of what the student knew and the unreliability of my own grading.

The procedure I adopted to tame the halo effect conforms to a general principle: decorrelate error! To understand how this principle works, imagine that a large number of observers are shown glass jars containing pennies and are challenged to estimate the number of pennies in each jar. As James Surowiecki explained in his best-selling *The Wisdom of Crowds*, this is the kind of task in which individuals do very poorly, but pools of individual judgments do remarkably well. Some individuals greatly overestimate the true number, others underestimate it, but when many judgments are averaged, the average tends to be quite accurate. The mechanism is straightforward: all individuals look at the same jar, and all their judgments have a common basis. On the other hand, the errors that individuals make are independent of the errors made by others, and (in the absence of a systematic bias) they tend to average to zero. However, the magic of error reduction works well only when the observations are independent and their errors uncorrelated. If the observers share a bias, the aggregation of judgments will not reduce it. Allowing the observers to influence each other effectively reduces the size of the sample, and with it the precision of the group estimate.

To derive the most useful information from multiple sources of evidence, you should always try to make these sources independent of each other. This rule is part of good police procedure. When there are multiple witnesses to an event, they are not allowed to discuss it before giving their testimony. The goal is not only to prevent collusion by hostile witnesses, it is also to prevent unbiased witnesses from influencing each other. Witnesses who exchange their experiences will tend to make similar errors in their testimony, reducing the total value of the information they provide. Eliminating redundancy from your sources of information is always a good idea.

The principle of independent judgments (and decorrelated errors) has immediate applications for the conduct of meetings, an activity in which executives in organizations spend a great deal of their working days. A simple rule can help: before an issue is discussed, all members of the committee should be asked to write a very brief summary of their position. This procedure makes good use of the value of the diversity of knowledge and opinion in the group. The standard practice of open discussion gives too much weight to the opinions of those who speak early and assertively, causing others to line up behind them.

What You See is All There is (Wysiati)

One of my favorite memories of the early years of working with Amos is a comedy routine he enjoyed performing. In a perfect impersonation of one of the professors with whom he had studied philosophy as an undergraduate, Amos would growl in Hebrew marked by a thick German accent: “You must never forget the *Primat of the Is*.” What exactly his teacher had meant by that phrase never became clear to me (or to Amos, I believe), but Amos’s jokes always maht=cipde a point. He was reminded of the old phrase (and eventually I was too) whenever we encountered the remarkable asymmetry between the ways our mind treats information that is currently available and information we do not have.

An essential design feature of the associative machine is that it represents only activated ideas. Information that is not retrieved (even unconsciously) from memory might as well not exist. System 1 excels at constructing the best possible story that incorporates ideas currently activated, but it does not (cannot) allow for information it does not have.

The measure of success for System 1 is the coherence of the story it manages to create. The amount and quality of the data on which the story is based are largely irrelevant. When information is scarce, which is a common occurrence, System 1 operates as a machine for jumping to conclusions. Consider the following: "Will Mindik be a good leader? She is intelligent and strong…" An answer quickly came to your mind, and it was yes. You picked the best answer based on the very limited information available, but you jumped the gun. What if the next two adjectives were *corrupt* and *cruel*?

Take note of what you did *not* do as you briefly thought of Mindik as a leader. You did not start by asking, "What would I need to know before I formed an opinion about the quality of someone's leadership?" System 1 got to work on its own from the first adjective: intelligent is good, intelligent and strong is very good. This is the best story that can be constructed from two adjectives, and System 1 delivered it with great cognitive ease. The story will be revised if new information comes in (such as Mindik is corrupt), but there is no waiting and no subjective discomfort. And there also remains a bias favoring the first impression.

The combination of a coherence-seeking System 1 with a lazy System 2 implies that System 2 will endorse many intuitive beliefs, which closely reflect the impressions generated by System 1. Of course, System 2 also is capable of a more systematic and careful approach to evidence, and of following a list of boxes that must be checked before making a decision—think of buying a home, when you deliberately seek information that you don't have. However, System 1 is expected to influence even the more careful decisions. Its input never ceases.

Jumping to conclusions on the basis of limited evidence is so important to an understanding of intuitive thinking, and comes up so often in this book, that I will use a cumbersome abbreviation for it: WYSIATI, which stands for what you see is all there is. System 1 is radically insensitive to both the quality and the quantity of the information that gives rise to impressions and intuitions.

Amos, with two of his graduate students at Stanford, reported a study that bears directly on WYSIATI, by observing the reaction of people who are given one-sided evidence and know it. The participants were exposed to legal scenarios such as the following:

> On September 3, plaintiff David Thornton, a forty-three-year-old union field representative, was present in Thrifty Drug Store #168, performing a routine union visit. Within ten minutes of his arrival, a store manager confronted him and told him he could no longer speak with the union employees on the floor of the store. Instead, he would have to see them in a back room while they were on break. Such a request is allowed by the union contract with Thrifty Drug but had never before been enforced. When Mr. Thornton objected, he was told that he had the choice of conto room whilforming to these requirements, leaving the store, or being arrested. At this

point, Mr. Thornton indicated to the manager that he had always been allowed to speak to employees on the floor for as much as ten minutes, as long as no business was disrupted, and that he would rather be arrested than change the procedure of his routine visit. The manager then called the police and had Mr. Thornton handcuffed in the store for trespassing. After he was booked and put into a holding cell for a brief time, all charges were dropped. Mr. Thornton is suing Thrifty Drug for false arrest.

In addition to this background material, which all participants read, different groups were exposed to presentations by the lawyers for the two parties. Naturally, the lawyer for the union organizer described the arrest as an intimidation attempt, while the lawyer for the store argued that having the talk in the store was disruptive and that the manager was acting properly. Some participants, like a jury, heard both sides. The lawyers added no useful information that you could not infer from the background story.

The participants were fully aware of the setup, and those who heard only one side could easily have generated the argument for the other side. Nevertheless, the presentation of one-sided evidence had a very pronounced effect on judgments. Furthermore, participants who saw one-sided evidence were more confident of their judgments than those who saw both sides. This is just what you would expect if the confidence that people experience is determined by the coherence of the story they manage to construct from available information. It is the consistency of the information that matters for a good story, not its completeness. Indeed, you will often find that knowing little makes it easier to fit everything you know into a coherent pattern.

WY SIATI facilitates the achievement of coherence and of the cognitive ease that causes us to accept a statement as true. It explains why we can think fast, and how we are able to make sense of partial information in a complex world. Much of the time, the coherent story we put together is close enough to reality to support reasonable action. However, I will also invoke WY SIATI to help explain a long and diverse list of biases of judgment and choice, including the following among many others:

- Overconfidence: As the WY SIATI rule implies, neither the quantity nor the quality of the evidence counts for much in subjective confidence. The confidence that individuals have in their beliefs depends mostly on the quality of the story they can tell about what they see, even if they see little. We often fail to allow for the possibility that evidence that should be critical to our judgment is missing—what we see is all there is. Furthermore, our associative system tends to settle on a coherent pattern of activation and suppresses doubt and ambiguity.
- Framing effects: Different ways of presenting the same information often evoke different emotions. The statement that “the odds of survival one month after surgery are 90%” is more reassuring than the equivalent statement that “mortality within one month of surgery is 10%.” Similarly, cold cuts described as “90% fat-free” are more

attractive than when they are described as “10% fat.” The equivalence of the alternative formulations is transparent, but an individual normally sees only one formulation, and what she sees is all there is.

- Base-rate neglect: Recall Steve, the meek and tidy soul who is often believed to be a librarian. The personality description is salient and vivid, and although you surely know that there are more male farm mu
- Base-rers than male librarians, that statistical fact almost certainly did not come to your mind when you first considered the question. What you saw was all there was.

Speaking of Jumping to Conclusions

“She knows nothing about this person’s management skills. All she is going by is the halo effect from a good presentation.”

“Let’s decorrelate errors by obtaining separate judgments on the issue before any discussion. We will get more information from independent assessments.”

“They made that big decision on the basis of a good report from one consultant. WYSIATI—what you see is all there is. They did not seem to realize how little information they had.”

“They didn’t want more information that might spoil their story. WYSIATI.”

P

How Judgments Happen

There is no limit to the number of questions you can answer, whether they are questions someone else asks or questions you ask yourself. Nor is there a limit to the number of attributes you can evaluate. You are capable of counting the number of capital letters on this page, comparing the height of the windows of your house to the one across the street, and assessing the political prospects of your senator on a scale from excellent to disastrous. The questions are addressed to System 2, which will direct attention and search memory to find the answers. System 2 receives questions or generates them: in either case it directs attention and searches memory to find the answers. System 1 operates differently. It continuously monitors what is going on outside and inside the mind, and continuously generates assessments of various aspects of the situation without specific intention and with little or no effort. These *basic assessments* play an important role in intuitive judgment, because they are easily substituted for more difficult questions—this is the essential idea of the heuristics and biases approach. Two other features of System 1 also support the substitution of one judgment for another. One is the ability to translate values across dimensions, which you do in answering a question that most people find easy: "If Sam were as tall as he is intelligent, how tall would he be?" Finally, there is the mental shotgun. An intention of System 2 to answer a specific question or evaluate a particular attribute of the situation automatically triggers other computations, including basic assessments.

Basic Assessments

System 1 has been shaped by evolution to provide a continuous assessment of the main problems that an organism must solve to survive: How are things going? Is there a threat or a major opportunity? Is everything normal? Should I approach or avoid? The questions are perhaps less urgent for a human in a city environment than for a gazelle on the savannah, aalenc and e: How , but we have inherited the neural mechanisms that evolved to provide ongoing assessments of threat level, and they have not been turned off. Situations are constantly evaluated as good or bad, requiring escape or permitting approach. Good mood and cognitive ease are the human equivalents of assessments of safety and familiarity.

For a specific example of a basic assessment, consider the ability to discriminate friend from foe at a glance. This contributes to one's chances of survival in a dangerous world, and such a specialized capability has indeed evolved. Alex Todorov, my colleague at Princeton, has explored the biological roots of the rapid judgments of how safe it is to interact with a stranger. He showed that we are endowed with an ability to evaluate, in a single glance at a stranger's face, two potentially crucial facts about that person: how dominant (and therefore potentially threatening) he is, and how trustworthy he is, whether his intentions are more likely to be friendly or hostile. The shape of the face provides the cues for assessing dominance: a “strong” square chin is one such cue. Facial expression (smile or frown) provides the cues for assessing the stranger’s intentions. The combination of a square chin with a turned-down mouth may spell trouble. The accuracy of face reading is far from perfect: round chins are not a reliable indicator of meekness, and smiles can (to some extent) be faked. Still, even an imperfect ability to assess strangers confers a survival advantage.

This ancient mechanism is put to a novel use in the modern world: it has some influence on how people vote. Todorov showed his students pictures of men’s faces, sometimes for as little as one-tenth of a second, and asked them to rate the faces on various attributes, including likability and competence. Observers agreed quite well on those ratings. The faces that Todorov showed were not a random set: they were the campaign portraits of politicians competing for elective office. Todorov then compared the results of the electoral races to the ratings of competence that Princeton students had made, based on brief exposure to photographs and without any political context. In about 70% of the races for senator, congressman, and governor, the election winner was the candidate whose face had earned a higher rating of competence. This striking result was quickly confirmed in national elections in Finland, in zoning board elections in England, and in various electoral contests in Australia, Germany, and Mexico. Surprisingly (at least to me), ratings of competence were far more predictive of voting outcomes in Todorov’s study than ratings of likability.

Todorov has found that people judge competence by combining the two dimensions of strength and trustworthiness. The faces that exude competence combine a strong chin with a slight confident-appearing smile. There is no evidence that these facial features actually predict how well politicians will perform in office. But studies of the brain’s response to winning and losing candidates show that we are biologically predisposed to reject candidates who lack the attributes we value—in this research, losers evoked stronger indications of (negative) emotional response. This is an example of what I will call a *judgment heuristic* in the following chapters. Voters are attempting to form an impression of how good a candidate will be in office, and they fall back on a simpler assessment that is made quickly and automatically and is available when System 2 must make its decision.

Political scientists followed up on Todorov’s initial research by identifying a category of voters for whom the automatic preferences of System 1 are particularly likely to play a large role. They found what they were looking for among politicalr m=“5%”>Todoly uninformed voters who watch a great deal of television. As expected, the effect of facial competence on voting is about three times larger for information-poor and TV-prone voters than for others who are better informed and watch less television. Evidently, the relative importance of System 1 in determining voting choices is not the same for all people. We will encounter other examples of such individual differences.

System 1 understands language, of course, and understanding depends on the basic assessments that are routinely carried out as part of the perception of events and the comprehension of messages. These assessments include computations of similarity and representativeness, attributions of causality, and evaluations of the availability of associations and exemplars. They are performed even in the absence of a specific task set, although the results are used to meet task demands as they arise.

The list of basic assessments is long, but not every possible attribute is assessed. For an example, look briefly at figure 7.

A glance provides an immediate impression of many features of the display. You know that the two towers are equally tall and that they are more similar to each other than the tower on the left is to the array of blocks in the middle. However, you do not immediately know that the number of blocks in the left-hand tower is the same as the number of blocks arrayed on the floor, and you have no impression of the height of the tower that you could build from them. To confirm that the numbers are the same, you would need to count the two sets of blocks and compare the results, an activity that only System 2 can carry out.

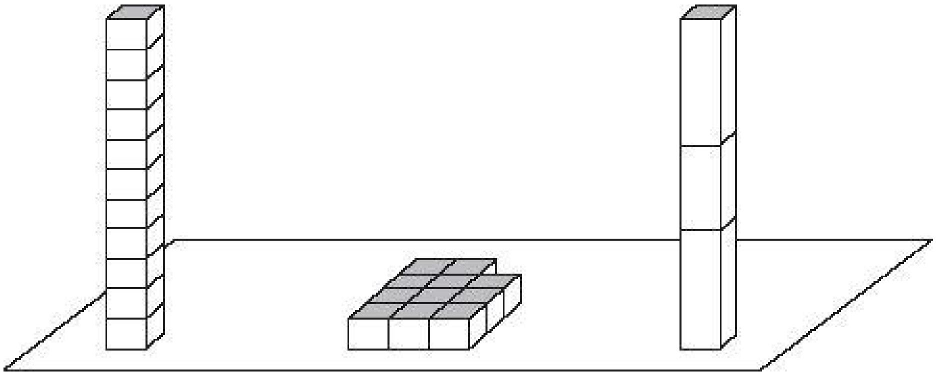


Figure 7

Sets and Prototypes

For another example, consider the question: What is the average length of the lines in figure 8?

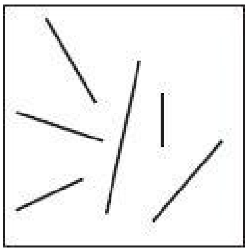


Figure 8

This question is easy and System 1 answers it without prompting. Experiments have shown that a fraction of a second is sufficient for people to register the average length of an array of lines with considerable precision. Furthermore, the accuracy of these judgments is not impaired when the observer is cognitively busy with a memory task. They do not necessarily know how to describe the average in inches or centimeters, but they will be very accurate in adjusting the length of another line to match the average. System 2 is not needed to form an impression of the norm of length for an array. System 1 does it, automatically and effortlessly, just as it registers the color of the lines and the fact that they are not parallel. We also can form an immediate impression of the number of objects in an array—precisely if there are four or fewer objects, crudely if there are more.

Now to another question: What is the total length of the lines in figure 8? This is a different experience, because System 1 has no suggestions to offer. The only way you can answer this question is by activating System 2, which will laboriously estimate the average, estimate or count the lines, and multiply average length by the number of lines.

estimaight=“0%”>

The failure of System 1 to compute the total length of a set of lines at a glance may look obvious to you; you never thought you could do it. It is in fact an instance of an important limitation of that system. Because System 1 represents categories by a prototype or a set of typical exemplars, it deals well with averages but poorly with sums. The size of the category, the number of instances it contains, tends to be ignored in judgments of what I will call *sum-like variables*.

Participants in one of the numerous experiments that were prompted by the litigation following the disastrous *Exxon Valdez* oil spill were asked their willingness to pay for nets to cover oil ponds in which migratory birds often drown. Different groups of participants stated their willingness to pay to save 2,000, 20,000, or 200,000 birds. If saving birds is an economic good it should be a sum-like variable: saving 200,000 birds should be worth much more than saving 2,000 birds. In fact, the average contributions of the three groups were $80, $78, and $88 respectively. The number of birds made very little difference. What the participants reacted to, in all three groups, was a prototype—the awful image of a helpless bird drowning, its feathers soaked in thick oil. The almost complete neglect of quantity in such emotional contexts has been confirmed many times.

Intensity Matching

Questions about your happiness, the president’s popularity, the proper punishment of financial evildoers, and the future prospects of a politician share an important characteristic: they all refer to an underlying dimension of intensity or amount, which permits the use of the word *more*: more happy, more popular, more severe, or more powerful (for a politician). For example, a candidate’s political future can range from the low of “She will be defeated in the primary” to a high of “She will someday be president of the United States.”

Here we encounter a new aptitude of System 1. An underlying scale of intensity allows *matching* across diverse dimensions. If crimes were colors, murder would be a deeper shade of red than theft. If crimes were expressed as music, mass murder would be played fortissimo while accumulating unpaid parking tickets would be a faint pianissimo. And of course you have similar feelings about the intensity of punishments. In classic experiments, people adjusted the loudness of a sound to the severity of crimes; other people adjusted loudness to the severity of legal punishments. If you heard two notes, one for the crime and one for the punishment, you would feel a sense of injustice if one tone was much louder than the other.

Consider an example that we will encounter again later:

Julie read fluently when she was four years old.

Now match Julie's reading prowess as a child to the following intensity scales:

How tall is a man who is as tall as Julie was precocious?

What do you think of 6 feet? Obviously too little. What about 7 feet? Probably too much. You are looking for a height that is as remarkable as the achievement of reading at age four. Fairly remarkable, but not extraordinary. Reading at fifteen months would be extraordinary, perhaps like a man who is 7′8″.

What level of income in your profession matches Julie's reading achievement?

Which crime is as severe as Julie was precocious?

Which graduating GPA in an Ivy League college matches Julie's reading?

Not very hard, was it? Furthermore, you can be assured that your matches will be quite close to those of other people in your cultural milieu. We will see that when people are asked to predict Julie's GPA from the information about the age at which she learned to read, they answer by translating from one scale to another and pick the matching GPA. And we will also see why this mode of prediction by matching is statistically wrong—although it is perfectly natural to System 1, and for most people except statisticians it is also acceptable to System 2.

The Mental Shotgun

System 1 carries out many computations at any one time. Some of these are routine assessments that go on continuously. Whenever your eyes are open, your brain computes a three-dimensional representation of what is in your field of vision, complete with the shape of objects, their position in space, and their identity. No intention is needed to trigger this operation or the continuous monitoring for violated expectations. In contrast to these routine assessments, other computations are undertaken only when needed: you do not maintain a continuous evaluation of how happy or wealthy you are, and even if you are a political addict you do not continuously assess the president's prospects. The occasional judgments are voluntary. They occur only when you intend them to do so.

You do not automatically count the number of syllables of every word you read, but you can do it if you so choose. However, the control over intended computations is far from precise: we often compute much more than we want or need. I call this excess computation the *mental shotgun*. It is impossible to aim at a single point with a shotgun because it shoots pellets that scatter, and it seems almost equally difficult for System 1 not to do more than System 2 charges it to do. Two experiments that I read long ago suggested this image.

Participants in one experiment listened to pairs of words, with the instruction to press a key as quickly as possible whenever they detected that the words rhymed. The words rhyme in both these pairs:

VOTE—NOTE

VOTE—GOAT

The difference is obvious to you because you see the two pairs. VOTE and GOAT rhyme, but they are spelled differently. The participants only heard the words, but they were also influenced by the spelling. They were distinctly slower to recognize the words as rhyming if their spelling was discrepant. Although the instructions required only a comparison of sounds, the participants also compared their spelling, and the mismatch on the irrelevant dimension slowed them down. An intention to answer one question evoked another, which was not only superfluous but actually detrimental to the main task.

In another study, people listened to a series of sentences, with the instruction to press one key as quickly as post="lly desible to indicate if the sentence was literally true, and another key if the sentence was not literally true. What are the correct responses for the following sentences?

Some roads are snakes.

Some jobs are snakes.

Some jobs are jails.

All three sentences are literally false. However, you probably noticed that the second sentence is more obviously false than the other two—the reaction times collected in the experiment confirmed a substantial difference. The reason for the difference is that the two difficult sentences can be metaphorically true. Here again, the intention to perform one computation evoked another. And here again, the correct answer prevailed in the conflict, but the conflict with the irrelevant answer disrupted performance. In the next chapter we will see that the combination of a mental shotgun with intensity matching explains why we have intuitive judgments about many things that we know little about.

Speaking of Judgment

> "Evaluating people as attractive or not is a basic assessment. You do that automatically whether or not you want to, and it influences you."

> "There are circuits in the brain that evaluate dominance from the shape of the face. He looks the part for a leadership role."

> "The punishment won't feel just unless its intensity matches the crime. Just like you can match the loudness of a sound to the brightness of a light."

> "This was a clear instance of a mental shotgun. He was asked whether he thought the company was financially sound, but he couldn't forget that he likes their product."

P

Answering an Easier Question

A remarkable aspect of your mental life is that you are rarely stumped. True, you occasionally face a question such as 17 × 24 = ? to which no answer comes immediately to mind, but these dumbfounded moments are rare. The normal state of your mind is that you have intuitive feelings and opinions about almost everything that comes your way. You like or dislike people long before you know much about them; you trust or distrust strangers without knowing why; you feel that an enterprise is bound to succeed without analyzing it. Whether you state them or not, you often have answers to questions that you do not completely understand, relying on evidence that you can neither explain nor defend.

Substituting Questions

I propose a simple account of how we generate intuitive opinions on complex matters. If a satisfactory answer to a hard question isebr ques D not found quickly, System 1 will find a related question that is easier and will answer it. I call the operation of answering one question in place of another *substitution*. I also adopt the following terms:

The target question is the assessment you intend to produce.

The heuristic question is the simpler question that you answer instead.

The technical definition of *heuristic* is a simple procedure that helps find adequate, though often imperfect, answers to difficult questions. The word comes from the same root as *eureka*.

The idea of substitution came up early in my work with Amos, and it was the core of what became the heuristics and biases approach. We asked ourselves how people manage to make judgments of probability without knowing precisely what probability is. We concluded that people must somehow simplify that impossible task, and we set out to find how they do it. Our answer was that when called upon to judge probability, people actually judge something else and believe they have judged probability. System 1 often makes this move when faced with difficult target questions, if the answer to a related and easier heuristic question comes readily to mind.

Substituting one question for another can be a good strategy for solving difficult problems, and George Pólya included substitution in his classic *How to Solve It*: "If you can't solve a problem, then there is an easier problem you can solve: find it." Pólya's heuristics are strategic procedures that are deliberately implemented by System 2. But the heuristics that I discuss in this chapter are not chosen; they are a consequence of the mental shotgun, the imprecise control we have over targeting our responses to questions.

Consider the questions listed in the left-hand column of table 1. These are difficult questions, and before you can produce a reasoned answer to any of them you must deal with other difficult issues. What is the meaning of happiness? What are the likely political developments in the next six months? What are the standard sentences for other financial crimes? How strong is the competition that the candidate faces? What other environmental or other causes should be considered? Dealing with these questions seriously is completely impractical. But you are not limited to perfectly reasoned answers to questions. There is a heuristic alternative to careful reasoning, which sometimes works fairly well and sometimes leads to serious errors.

Target Question	*Heuristic Question*
How much would you contribute to save an endangered species?	How much emotion do I feel when I think of dying dolphins?
How happy are you with your life these days?	What is my mood right now?
How popular is the president right now?	How popular will the president be six months from now?
How should financial advisers who prey on the elderly be punished?	How much anger do I feel when I think of financial predators?
This woman is running for the primary. How far will she go in politics?	Does this woman look like a political winner?

Table 1

The mental shotgun makes it easy to generate quick answers to difficult questions without imposing much hard work on your lazy System 2. The right-hand counterpart of each of the left-hand questions is very likely to be evoked and very easily answered. Your feelings about dolphins and financial crooks, your current mood, your impressions of the political skill of the primary candidate, or the current standing of the president will readily come to mind. The heuristic questions provide an off-the-shelf answer to each of the difficult target questions.

Something is still missing from this story: the answers need to be fitted to the original questions. For example, my feelings about dying dolphins must be expressed in dollars. Another capability of System 1, intensity matching, is available to solve that problem. Recall that both feelings and contribution dollars are intensity scales. I can feel more or less strongly about dolphins and there is a contribution that matches the intensity of my feelings. The dollar amount that will come to my mind is the matching amount. Similar intensity matches are possible for all the questions. For example, the political skills of a candidate can range from pathetic to extraordinarily impressive, and the scale of political success can range from the low of “She will be defeated in the primary” to a high of “She will someday be president of the United States.”

The automatic processes of the mental shotgun and intensity matching often make available one or more answers to easy questions that could be mapped onto the target question. On some occasions, substitution will occur and a heuristic answer will be endorsed by System 2. Of course, System 2 has the opportunity to reject this intuitive answer, or to modify it by incorporating other information. However, a lazy System 2 often follows the path of least effort and endorses a heuristic answer without much scrutiny of whether it is truly appropriate. You will not be stumped, you will not have to work very her p wheard, and you may not even notice that you did not answer the question you were asked. Furthermore, you may not realize that the target question was difficult, because an intuitive answer to it came readily to mind.

The 3-D Heuristic

Have a look at the picture of the three men and answer the question that follows.

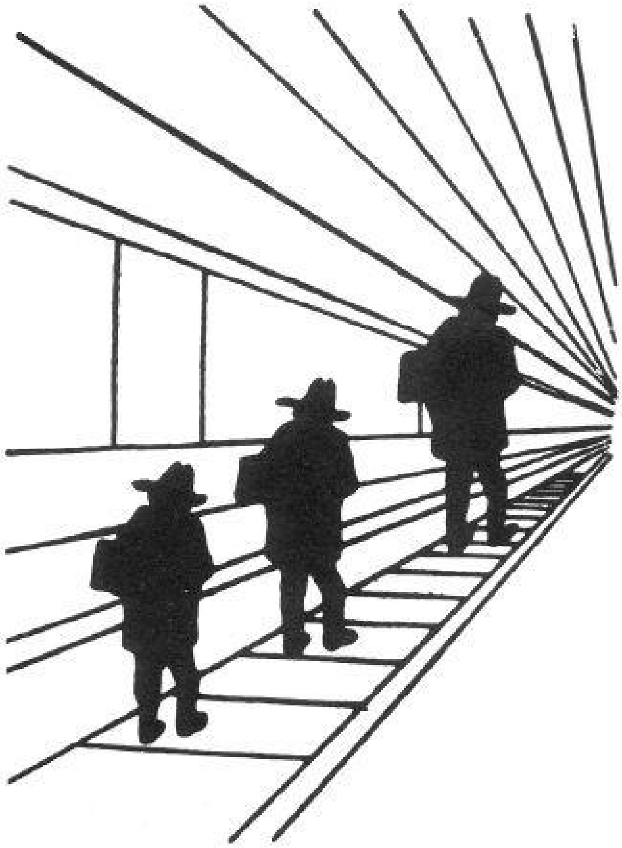


Figure 9

As printed on the page, is the figure on the right larger than the figure on the left?

The obvious answer comes quickly to mind: the figure on the right is larger. If you take a ruler to the two figures, however, you will discover that in fact the figures are exactly the same size. Your impression of their relative size is dominated by a powerful illusion, which neatly illustrates the process of substitution.

The corridor in which the figures are seen is drawn in perspective and appears to go into the depth plane. Your perceptual system automatically interprets the picture as a three-dimensional scene, not as an image printed on a flat paper surface. In the 3-D interpretation, the person on the right is both much farther away and much larger than the person on the left. For most of us, this impression of 3-D size is overwhelming. Only visual artists and experienced photographers have developed the skill of seeing the drawing as an object on the page. For the rest of us, substitution occurs: the dominant impression of 3-D size dictates the judgment of 2-D size. The illusion is due to a 3-D heuristic.

What happens here is a true illusion, not a misunderstanding of the question. You knew that the question was about the size of the figures in the picture, as printed on the page. If you had been asked to estimate the size of the figures, we know from experiments that your answer would have been in inches, not feet. You were not confused about the question, but you were influenced by the answer to a question that you were not asked: “How tall are the three people?”

The essential step in the heuristic—the substitution of three-dimensional for two-dimensional size—occurred automatically. The picture contains cues that suggest a 3-D interpretation. These cues are irrelevant to the task at hand—the judgment of size of the figure on the page—and you should have ignored them, but you could not. The bias associated with the heuristic is that objects that appear to be more distant also appear to be larger on the page. As this example illustrates, a judgment that is based on substitution will inevitably be biased in predictable ways. In this case, it happens so deep in the perceptual system that you simply cannot help it.

The Mood Heuristic for Happiness

A survey of German students is one of the best examples of substitution. The survey that the young participants completed included the following two questions:

How happy are you these days?

How many dates did you have last month?

< stpr to a p height=“0%” width=“0%”>The experimenters were interested in the correlation between the two answers. Would the students who reported many dates say that they were happier than those with fewer dates? Surprisingly, no: the correlation between the answers was about zero. Evidently, dating was not what came first to the students’ minds when they were asked to assess their happiness. Another group of students saw the same two questions, but in reverse order:

How many dates did you have last month?

How happy are you these days?

The results this time were completely different. In this sequence, the correlation between the number of dates and reported happiness was about as high as correlations between psychological measures can get. What happened?

The explanation is straightforward, and it is a good example of substitution. Dating was apparently not the center of these students’ life (in the first survey, happiness and dating were uncorrelated), but when they were asked to think about their romantic life, they certainly had an emotional reaction. The students who had many dates were reminded of a happy aspect of their life, while those who had none were reminded of loneliness and rejection. The emotion aroused by the dating question was still on everyone’s mind when the query about general happiness came up.

The psychology of what happened is precisely analogous to the psychology of the size illusion in figure 9. “Happiness these days” is not a natural or an easy assessment. A good answer requires a fair amount of thinking. However, the students who had just been asked about their dating did not need to think hard because they already had in their mind an answer to a related question: how happy they were with their love life. They substituted the question to which they had a readymade answer for the question they were asked.

Here again, as we did for the illusion, we can ask: Are the students confused? Do they really think that the two questions—the one they were asked and the one they answer—are synonymous? Of course not. The students do not temporarily lose their ability to distinguish romantic life from life as a whole. If asked about the two concepts, they would say they are different. But they were not asked whether the concepts are different. They were asked how happy they were, and System 1 has a ready answer.

Dating is not unique. The same pattern is found if a question about the students’ relations with their parents or about their finances immediately precedes the question about general happiness. In both cases, satisfaction in the particular domain dominates happiness reports. Any emotionally significant question that alters a person’s mood will have the same effect. WYSIATI. The present state of mind looms very large when people evaluate their happiness.

The Affect Heuristic

The dominance of conclusions over arguments is most pronounced where emotions are involved. The psychologist Paul Slovic has proposed an *affect heuristic* in which people let their likes and dislikes determine their beliefs about the world. Your political preference determines the arguments that you find compelling. If you like the current health policy, you believe its benefits are substantial and its costs more manageable than the costs of alternatives. If you are a hawk in your attitude toward other nations, you probablthep“0%y think they are relatively weak and likely to submit to your country’s will. If you are a dove, you probably think they are strong and will not be easily coerced. Your emotional attitude to such things as irradiated food, red meat, nuclear power, tattoos, or motorcycles drives your beliefs about their benefits and their risks. If you dislike any of these things, you probably believe that its risks are high and its benefits negligible.

The primacy of conclusions does not mean that your mind is completely closed and that your opinions are wholly immune to information and sensible reasoning. Your beliefs, and even your emotional attitude, may change (at least a little) when you learn that the risk of an activity you disliked is smaller than you thought. However, the information about lower risks will also change your view of the benefits (for the better) even if nothing was said about benefits in the information you received.

We see here a new side of the “personality” of System 2. Until now I have mostly described it as a more or less acquiescent monitor, which allows considerable leeway to System 1. I have also presented System 2 as active in deliberate memory search, complex computations, comparisons, planning, and choice. In the bat-and-ball problem and in many other examples of the interplay between the two systems, it appeared that System 2 is ultimately in charge, with the ability to resist the suggestions of System 1, slow things down, and impose logical analysis. Self-criticism is one of the functions of System 2. In the context of attitudes, however, System 2 is more of an apologist for the emotions of System 1 than a critic of those emotions—an endorser rather than an enforcer. Its search for information and arguments is mostly constrained to information that is consistent with existing beliefs, not with an intention to examine them. An active, coherence-seeking System 1 suggests solutions to an undemanding System 2.

Speaking of Substitution and Heuristics

> "Do we still remember the question we are trying to answer? Or have we substituted an easier one?"

> "The question we face is whether this candidate can succeed. The question we seem to answer is whether she interviews well. Let's not substitute."

> "He likes the project, so he thinks its costs are low and its benefits are high. Nice example of the affect heuristic."

> "We are using last year's performance as a heuristic to predict the value of the firm several years from now. Is this heuristic good enough? What other information do we need?"

The table below contains a list of features and activities that have been attributed to System 1. Each of the active sentences replaces a statement, technically more accurate but harder to understand, to the effect that a mental event occurs automatically and fast. My hope is that the list of traits will help you develop an intuitive sense of the "personality" of the fictitious System 1. As happens with other characters you know, you will have hunches about what System 1 would do under different circumstances, and most of your hunches will be correct.

Characteristics of System 1

- generates impressions, feelings, and inclinations; when endorsed by System 2 these become beliefs, attitudes, and intentions
- operates automatically and quickly, with little or no effort, and no sense of voluntary control
- can be programmed by System 2 to mobilize attention when a particular pattern is detected (search)
- executes skilled responses and generates skilled intuitions, after adequate training
- creates a coherent pattern of activated ideas in associative memory
- links a sense of cognitive ease to illusions of truth, pleasant feelings, and reduced vigilance
- distinguishes the surprising from the normal
- infers and invents causes and intentions
- neglects ambiguity and suppresses doubt
- is biased to believe and confirm
- exaggerates emotional consistency (halo effect)
- focuses on existing evidence and ignores absent evidence (WYSIATI)

- generates a limited set of basic assessments
- represents sets by norms and prototypes, does not integrate

- matches intensities across scales (e.g., size to loudness)
- computes more than intended (mental shotgun)
- sometimes substitutes an easier question for a difficult one (heuristics)
- is more sensitive to changes than to states (prospect theory)[*]
- overweights low probabilities[*]
- shows diminishing sensitivity to quantity (psychophysics)[*]
- responds more strongly to losses than to gains (loss aversion)[*]
- frames decision problems narrowly, in isolation from one another[*]

P

Part 2

P

Heuristics and Biases

P

The Law of Small Numbers

A study of the incidence of kidney cancer in the 3,141 counties of the United a>< НЉStates reveals a remarkable pattern. The counties in which the incidence of kidney cancer is lowest are mostly rural, sparsely populated, and located in traditionally Republican states in the Midwest, the South, and the West. What do you make of this?

Your mind has been very active in the last few seconds, and it was mainly a System 2 operation. You deliberately searched memory and formulated hypotheses. Some effort was involved; your pupils dilated, and your heart rate increased measurably. But System 1 was not idle: the operation of System 2 depended on the facts and suggestions retrieved from associative memory. You probably rejected the idea that Republican politics provide protection against kidney cancer. Very likely, you ended up focusing on the fact that the counties with low incidence of cancer are mostly rural. The witty statisticians Howard Wainer and Harris Zwerling, from whom I learned this example, commented, “It is both easy and tempting to infer that their low cancer rates are directly due to the clean living of the rural lifestyle—no air pollution, no water pollution, access to fresh food without additives.” This makes perfect sense.

Now consider the counties in which the incidence of kidney cancer is highest. These ailing counties tend to be mostly rural, sparsely populated, and located in traditionally Republican states in the Midwest, the South, and the West. Tongue-in-cheek, Wainer and Zwerling comment: “It is easy to infer that their high cancer rates might be directly due to the poverty of the rural lifestyle—no access to good medical care, a high-fat diet, and too much alcohol, too much tobacco.” Something is wrong, of course. The rural lifestyle cannot explain both very high and very low incidence of kidney cancer.

The key factor is not that the counties were rural or predominantly Republican. It is that rural counties have small populations. And the main lesson to be learned is not about epidemiology, it is about the difficult relationship between our mind and statistics. System 1 is highly adept in one form of thinking—it automatically and effortlessly identifies causal connections between events, sometimes even when the connection is spurious. When told about the high-incidence counties, you immediately assumed that these counties are different from other counties for a reason, that there must be a cause that explains this difference. As we shall see, however, System 1 is inept when faced with “merely statistical” facts, which change the probability of outcomes but do not cause them to happen.

A random event, by definition, does not lend itself to explanation, but collections of random events do behave in a highly regular fashion. Imagine a large urn filled with marbles. Half the marbles are red, half are white. Next, imagine a very patient person (or a robot) who blindly draws 4 marbles from the urn, records the number of red balls in the sample, throws the balls back into the urn, and then does it all again, many times. If you summarize the results, you will find that the outcome “2 red, 2 white” occurs (almost exactly) 6 times as often as the outcome “4 red” or “4 white.” This relationship is a mathematical fact. You can predict the outcome of repeated sampling from an urn just as confidently as you can predict what will happen if you hit an egg with a hammer. You cannot predict every detail of how the shell will shatter, but you can be sure of the general idea. There is a difference: the satisfying sense of causation that you experience when thinking of a hammer hitting an egg is altogether absent when you think about sampling.

A related statistical fact is relevant to the cancer example. From the same urn, two very patient marble counters thatpy dake turns. Jack draws 4 marbles on each trial, Jill draws 7. They both record each time they observe a homogeneous sample—all white or all red. If they go on long enough, Jack will observe such extreme outcomes more often than Jill—by a factor of 8 (the expected percentages are 12.5% and 1.56%). Again, no hammer, no causation, but a mathematical fact: samples of 4 marbles yield extreme results more often than samples of 7 marbles do.

Now imagine the population of the United States as marbles in a giant urn. Some marbles are marked KC, for kidney cancer. You draw samples of marbles and populate each county in turn. Rural samples are smaller than other samples. Just as in the game of Jack and Jill, extreme outcomes (very high and/or very low cancer rates) are most likely to be found in sparsely populated counties. This is all there is to the story.

We started from a fact that calls for a cause: the incidence of kidney cancer varies widely across counties and the differences are systematic. The explanation I offered is statistical: extreme outcomes (both high and low) are more likely to be found in small than in large samples. This explanation is not causal. The small population of a county neither causes nor prevents cancer; it merely allows the incidence of cancer to be much higher (or much lower) than it is in the larger population. The deeper truth is that there is nothing to explain. The incidence of cancer is not truly lower or higher than normal in a county with a small population, it just appears to be so in a particular year because of an accident of sampling. If we repeat the analysis next year, we will observe the same general pattern of extreme results in the small samples, but the counties where cancer was common last year will not necessarily have a high incidence this year. If this is the case, the differences between dense and rural counties do not really count as facts: they are what scientists call artifacts, observations that are produced entirely by some aspect of the method of research—in this case, by differences in sample size.

The story I have told may have surprised you, but it was not a revelation. You have long known that the results of large samples deserve more trust than smaller samples, and even people who are innocent of statistical knowledge have heard about this law of large numbers. But “knowing” is not a yes-no affair and you may find that the following statements apply to you:

- The feature “sparsely populated” did not immediately stand out as relevant when you read the epidemiological story.
- You were at least mildly surprised by the size of the difference between samples of 4

and samples of 7.

- Even now, you must exert some mental effort to see that the following two statements mean exactly the same thing:
- Large samples are more precise than small samples.
- Small samples yield extreme results more often than large samples do.

The first statement has a clear ring of truth, but until the second version makes intuitive sense, you have not truly understood the first.

The bottom line: yes, you did know that the results of large samples are more precise, but you may now realize that you did not know it very well. You are not alone. The first study that Amos and I did together showed that even sophisticated researchers have poor intuitions and a wobbly understanding of sampling effects.

The Law of Small Numbers

My collaboration with Amos in the early 1970s began with a discussion of the claim that people who have had no training in statistics are good "intuitive statisticians." He told my seminar and me of researchers at the University of Michigan who were generally optimistic about intuitive statistics. I had strong feelings about that claim, which I took personally: I had recently discovered that I was not a good intuitive statistician, and I did not believe that I was worse than others.

For a research psychologist, sampling variation is not a curiosity; it is a nuisance and a costly obstacle, which turns the undertaking of every research project into a gamble. Suppose that you wish to confirm the hypothesis that the vocabulary of the average six-year-old girl is larger than the vocabulary of an average boy of the same age. The hypothesis is true in the population; the average vocabulary of girls is indeed larger. Girls and boys vary a great deal, however, and by the luck of the draw you could select a sample in which the difference is inconclusive, or even one in which boys actually score higher. If you are the researcher, this outcome is costly to you because you have wasted time and effort, and failed to confirm a hypothesis that was in fact true. Using a sufficiently large sample is the only way to reduce the risk. Researchers who pick too small a sample leave themselves at the mercy of sampling luck.

The risk of error can be estimated for any given sample size by a fairly simple procedure. Traditionally, however, psychologists do not use calculations to decide on a sample size. They use their judgment, which is commonly flawed. An article I had read shortly before the debate with Amos demonstrated the mistake that researchers made (they still do) by a dramatic observation. The author pointed out that psychologists commonly chose samples so small that they exposed themselves to a 50% risk of failing to confirm their true hypotheses! No researcher in his right mind would accept such a risk. A plausible explanation was that psychologists' decisions about sample size reflected prevalent intuitive misconceptions of the extent of sampling variation.

The article shocked me, because it explained some troubles I had had in my own research. Like most research psychologists, I had routinely chosen samples that were too small and had often obtained results that made no sense. Now I knew why: the odd results were actually artifacts of my research method. My mistake was particularly embarrassing because I taught statistics and knew how to compute the sample size that would reduce the risk of failure to an acceptable level. But I had never chosen a sample size by computation. Like my colleagues, I had trusted tradition and my intuition in planning my experiments and had never thought seriously about the issue. When Amos visited the seminar, I had already reached the conclusion that my intuitions were deficient, and in the course of the seminar we quickly agreed that the Michigan optimists were wrong.

Amos and I set out to examine whether I was the only fool or a member of a majority of fools, by testing whether researchers selected for mathematical expertise would make similar mistakes. We developed a questionnaire that described realistic research situations, including replications of successful experiments. It asked the researchers to choose sample sizes, to assess the risks of failure to which their decisions exposed them, and to provide advice to hypothetical graduate students planning their research. Amos collected the responses of a group of sophisticated participants (including authors of two statistical textbooks) at a meetatipp>

Amos and I called our first joint article “Belief in the Law of Small Numbers.” We explained, tongue-in-cheek, that “intuitions about random sampling appear to satisfy the law of small numbers, which asserts that the law of large numbers applies to small numbers as well.” We also included a strongly worded recommendation that researchers regard their “statistical intuitions with proper suspicion and replace impression formation by computation whenever possible.”

A Bias of Confidence Over Doubt

In a telephone poll of 300 seniors, 60% support the president.

If you had to summarize the message of this sentence in exactly three words, what would they be? Almost certainly you would choose “elderly support president.” These words provide the gist of the story. The omitted details of the poll, that it was done on the phone with a sample of 300, are of no interest in themselves; they provide background information that attracts little attention. Your summary would be the same if the sample size had been different. Of course, a completely absurd number would draw your attention (“a telephone poll of 6 [or 60 million] elderly voters…”). Unless you are a professional, however, you may not react very differently to a sample of 150 and to a sample of 3,000. That is the meaning of the statement that “people are not adequately sensitive to sample size.”

The message about the poll contains information of two kinds: the story and the source of the story. Naturally, you focus on the story rather than on the reliability of the results. When the reliability is obviously low, however, the message will be discredited. If you are told that "a partisan group has conducted a flawed and biased poll to show that the elderly support the president…" you will of course reject the findings of the poll, and they will not become part of what you believe. Instead, the partisan poll and its false results will become a new story about political lies. You can choose to disbelieve a message in such clear-cut cases. But do you discriminate sufficiently between "I read in *The New York Times*…" and "I heard at the watercooler…"? Can your System 1 distinguish degrees of belief? The principle of WY SIATI suggests that it cannot.

As I described earlier, System 1 is not prone to doubt. It suppresses ambiguity and spontaneously constructs stories that are as coherent as possible. Unless the message is immediately negated, the associations that it evokes will spread as if the message were true. System 2 is capable of doubt, because it can maintain incompatible possibilities at the same time. However, sustaining doubt is harder work than sliding into certainty. The law of small numbers is a manifestation of a general bias that favors certainty over doubt, which will turn up in many guises in following chapters.

The strong bias toward believing that small samples closely resemble the population from which they are drawn is also part of a larger story: we are prone to exaggerate the consistency and coherence of what we see. The exaggerated faith of researchers in what can be learned from a few observations is closely related to the halo effect thphe , the sense we often get that we know and understand a person about whom we actually know very little. System 1 runs ahead of the facts in constructing a rich image on the basis of scraps of evidence. A machine for jumping to conclusions will act as if it believed in the law of small numbers. More generally, it will produce a representation of reality that makes too much sense.

Cause and Chance

The associative machinery seeks causes. The difficulty we have with statistical regularities is that they call for a different approach. Instead of focusing on how the event at hand came to be, the statistical view relates it to what could have happened instead. Nothing in particular caused it to be what it is—chance selected it from among its alternatives.

Our predilection for causal thinking exposes us to serious mistakes in evaluating the randomness of truly random events. For an example, take the sex of six babies born in sequence at a hospital. The sequence of boys and girls is obviously random; the events are independent of each other, and the number of boys and girls who were born in the hospital in the last few hours has no effect whatsoever on the sex of the next baby. Now consider three possible sequences:

BBBGGG

GGGGGG

BGBBGB

Are the sequences equally likely? The intuitive answer—"of course not!"—is false.

Because the events are independent and because the outcomes B and G are (approximately) equally likely, then any possible sequence of six births is as likely as any other. Even now that you know this conclusion is true, it remains counterintuitive, because only the third sequence appears random. As expected, BGBBGB is judged much more likely than the other two sequences. We are pattern seekers, believers in a coherent world, in which regularities (such as a sequence of six girls) appear not by accident but as a result of mechanical causality or of someone's intention. We do not expect to see regularity produced by a random process, and when we detect what appears to be a rule, we quickly reject the idea that the process is truly random. Random processes produce many sequences that convince people that the process is not random after all. You can see why assuming causality could have had evolutionary advantages. It is part of the general vigilance that we have inherited from ancestors. We are automatically on the lookout for the possibility that the environment has changed. Lions may appear on the plain at random times, but it would be safer to notice and respond to an apparent increase in the rate of appearance of prides of lions, even if it is actually due to the fluctuations of a random process.

The widespread misunderstanding of randomness sometimes has significant consequences. In our article on representativeness, Amos and I cited the statistician William Feller, who illustrated the ease with which people see patterns where none exists. During the intensive rocket bombing of London in World War II, it was generally believed that the bombing could not be random because a map of the hits revealed conspicuous gaps. Some suspected that German spies were located in the unharmed areas. A careful statistical analysis revealed that the distribution of hits was typical of a random process—and typical as well in evoking a strong impression that it was not random. "To the untrained eye," Feller remarks, "randomness appears as regularity or tendency to cluster."

I soon had an occasion to apply what I had learned frpeaprainom Feller. The Yom Kippur War broke out in 1973, and my only significant contribution to the war effort was to advise high officers in the Israeli Air Force to stop an investigation. The air war initially went quite badly for Israel, because of the unexpectedly good performance of Egyptian ground-to-air missiles. Losses were high, and they appeared to be unevenly distributed. I was told of two squadrons flying from the same base, one of which had lost four planes while the other had lost none. An inquiry was initiated in the hope of learning what it was that the unfortunate squadron was doing wrong. There was no prior reason to believe that one of the squadrons was more effective than the other, and no operational differences were found, but of course the lives of the pilots differed in many random ways, including, as I recall, how often they went home between missions and something about the conduct of debriefings. My advice was that the command should accept that the different outcomes were due to blind luck, and that the interviewing of the pilots should stop. I reasoned that luck was the most likely answer, that a random search for a nonobvious cause was hopeless, and that in the meantime the pilots in the squadron that had sustained losses did not need the extra burden of being made to feel that they and their dead friends were at fault.

Some years later, Amos and his students Tom Gilovich and Robert Vallone caused a stir with their study of misperceptions of randomness in basketball. The "fact" that players occasionally acquire a hot hand is generally accepted by players, coaches, and fans. The inference is irresistible: a player sinks three or four baskets in a row and you cannot help forming the causal judgment that this player is now hot, with a temporarily increased propensity to score. Players on both teams adapt to this judgment—teammates are more likely to pass to the hot scorer and the defense is more likely to doubleteam. Analysis of thousands of sequences of shots led to a disappointing conclusion: there is no such thing as a hot hand in professional basketball, either in shooting from the field or scoring from the foul line. Of course, some players are more accurate than others, but the sequence of successes and missed shots satisfies all tests of randomness. The hot hand is entirely in the eye of the beholders, who are consistently too quick to perceive order and causality in randomness. The hot hand is a massive and widespread cognitive illusion.

The public reaction to this research is part of the story. The finding was picked up by the press because of its surprising conclusion, and the general response was disbelief. When the celebrated coach of the Boston Celtics, Red Auerbach, heard of Gilovich and his study, he responded, “Who is this guy? So he makes a study. I couldn’t care less.” The tendency to see patterns in randomness is overwhelming—certainly more impressive than a guy making a study.

The illusion of pattern affects our lives in many ways off the basketball court. How many good years should you wait before concluding that an investment adviser is unusually skilled? How many successful acquisitions should be needed for a board of directors to believe that the CEO has extraordinary flair for such deals? The simple answer to these questions is that if you follow your intuition, you will more often than not err by misclassifying a random event as systematic. We are far too willing to reject the belief that much of what we see in life is random.

I began this chapter with the example of cancer incidence across the United States. The example appears in a book intended for statistics teachers, but I learned about it from an amusing article by the two statisticians I quoted earlier, Howard Wainer and Harris Zwerling. Their essay focused on a large iivepothersnvestment, some $1.7 billion, which the Gates Foundation made to follow up intriguing findings on the characteristics of the most successful schools. Many researchers have sought the secret of successful education by identifying the most successful schools in the hope of discovering what distinguishes them from others. One of the conclusions of this research is that the most successful schools, on average, are small. In a survey of 1,662 schools in Pennsylvania, for instance, 6 of the top 50 were small, which is an overrepresentation by a factor of 4. These data encouraged the Gates Foundation to make a substantial investment in the creation of small schools, sometimes by splitting large schools into smaller units. At least half a dozen other prominent institutions, such as the Annenberg Foundation and the Pew Charitable Trust, joined the effort, as did the U.S. Department of Education’s Smaller Learning Communities Program.

This probably makes intuitive sense to you. It is easy to construct a causal story that explains how small schools are able to provide superior education and thus produce high-achieving scholars by giving them more personal attention and encouragement than they could get in larger schools. Unfortunately, the causal analysis is pointless because the facts are wrong. If the statisticians who reported to the Gates Foundation had asked about the characteristics of the worst schools, they would have found that bad schools also tend to be smaller than average. The truth is that small schools are not better on average; they are simply more variable. If anything, say Wainer and Zwerling, large schools tend to produce better results, especially in higher grades where a variety of curricular options is valuable.

Thanks to recent advances in cognitive psychology, we can now see clearly what Amos and I could only glimpse: the law of small numbers is part of two larger stories about the workings of the mind.

- The exaggerated faith in small samples is only one example of a more general illusion—we pay more attention to the content of messages than to information about their reliability, and as a result end up with a view of the world around us that is simpler and more coherent than the data justify. Jumping to conclusions is a safer sport in the world of our imagination than it is in reality.
- Statistics produce many observations that appear to beg for causal explanations but do not lend themselves to such explanations. Many facts of the world are due to chance, including accidents of sampling. Causal explanations of chance events are inevitably wrong.

Speaking of the Law of Small Numbers

"Yes, the studio has had three successful films since the new CEO took over. But it is too early to declare he has a hot hand."

"I won't believe that the new trader is a genius before consulting a statistician who could estimate the likelihood of his streak being a chance event."

"The sample of observations is too small to make any inferences. Let's not follow the law of small numbers."

"I plan to keep the results of the experiment secret until we have a sufficiently large sample. Otherwisortpxpere we will face pressure to reach a conclusion prematurely."

P

Anchors

Amos and I once rigged a wheel of fortune. It was marked from 0 to 100, but we had it built so that it would stop only at 10 or 65. We recruited students of the University of Oregon as participants in our experiment. One of us would stand in front of a small group, spin the wheel, and ask them to write down the number on which the wheel stopped, which of course was either 10 or 65. We then asked them two questions:

> Is the percentage of African nations among UN members larger or smaller than the number you just wrote?

> What is your best guess of the percentage of African nations in the UN?

The spin of a wheel of fortune—even one that is not rigged—cannot possibly yield useful information about anything, and the participants in our experiment should simply have ignored it. But they did not ignore it. The average estimates of those who saw 10 and 65 were 25% and 45%, respectively.

The phenomenon we were studying is so common and so important in the everyday world that you should know its name: it is an *anchoring effect*. It occurs when people consider a particular value for an unknown quantity before estimating that quantity. What happens is one of the most reliable and robust results of experimental psychology: the estimates stay close to the number that people considered—hence the image of an anchor. If you are asked whether Gandhi was more than 114 years old when he died you will end up with a much higher estimate of his age at death than you would if the anchoring question referred to death at 35. If you consider how much you should pay for a house, you will be influenced by the asking price. The same house will appear more valuable if its listing price is high than if it is low, even if you are determined to resist the influence of this number; and so on—the list of anchoring effects is endless. Any number that you are asked to consider as a possible solution to an estimation problem will induce an anchoring effect.

We were not the first to observe the effects of anchors, but our experiment was the first demonstration of its absurdity: people's judgments were influenced by an obviously uninformative number. There was no way to describe the anchoring effect of a wheel of fortune as reasonable. Amos and I published the experiment in our *Science* paper, and it is one of the best known of the findings we reported there.

There was only one trouble: Amos and I did not fully agree on the psychology of the anchoring effect. He supported one interpretation, I liked another, and we never found a way to settle the argument. The problem was finally solved decades later by the efforts of numerous investigators. It is now clear that Amos and I were both right. Two different mechanisms produce anchoring effects—one for each system. There is a form of anchoring that occurs in a deliberate process of adjustment, an operation of System 2. And there is anchoring that occurs by a priming effect, an automatic manifestation of System 1.

Anchoring as Adjustment

Amos liked the idea of an adjust-and-anchor heuristic as a strategy for estimating uncertain quantities: start from an anchoring number, assess whether it is too high or too low, and gradually adjust your estimate by mentally "moving" from the anchor. The adjustment typically ends prematurely, because people stop when they are no longer certain that they should move farther. Decades after our disagreement, and years after Amos's death, convincing evidence of such a process was offered independently by two psychologists who had worked closely with Amos early in their careers: Eldar Shafir and Tom Gilovich together with their own students—Amos's intellectual grandchildren!

To get the idea, take a sheet of paper and draw a 2½-inch line going up, starting at the bottom of the page—without a ruler. Now take another sheet, and start at the top and draw a line going down until it is 2½ inches from the bottom. Compare the lines. There is a good chance that your first estimate of 2½ inches was shorter than the second. The reason is that you do not know exactly what such a line looks like; there is a range of uncertainty. You stop near the bottom of the region of uncertainty when you start from the bottom of the page and near the top of the region when you start from the top. Robyn Le Boeuf and Shafir found many examples of that mechanism in daily experience. Insufficient adjustment neatly explains why you are likely to drive too fast when you come off the highway onto city streets—especially if you are talking with someone as you drive. Insufficient adjustment is also a source of tension between exasperated parents and teenagers who enjoy loud music in their room. Le Boeuf and Shafir note that a "well-intentioned child who turns down exceptionally loud music to meet a parent's demand that it be played at a 'reasonable' volume may fail to adjust sufficiently from a high anchor, and may feel that genuine attempts at compromise are being overlooked." The driver and the child both deliberately adjust down, and both fail to adjust enough.

Now consider these questions:

When did George Washington become president?

What is the boiling temperature of water at the top of Mount Everest?

The first thing that happens when you consider each of these questions is that an anchor comes to your mind, and you know both that it is wrong and the direction of the correct answer. You know immediately that George Washington became president after 1776, and you also know that the boiling temperature of water at the top of Mount Everest is lower than 100°C. You have to adjust in the appropriate direction by finding arguments to move away from the anchor. As in the case of the lines, you are likely to stop when you are no longer sure you should go farther—at the near edge of the region of uncertainty.

Nick Epley and Tom Gilovich found evidence that adjustment is a deliberate attempt to find reasons to move away from the anchor: people who are instructed to shake their head when they hear the anchor, as if they rejected it, move farther from the anchor, and people who nod their head show enhanced anchoring. Epley and Gilovich also confirmed that adjustment is an effortful operation. People adjust less (stay closer to the anchor) when their mental resources are depleted, either because their memory is loaded with dighdth=igits or because they are slightly drunk. Insufficient adjustment is a failure of a weak or lazy System 2.

So we now know that Amos was right for at least some cases of anchoring, which involve a deliberate System 2 adjustment in a specified direction from an anchor.

Anchoring as Priming Effect

When Amos and I debated anchoring, I agreed that adjustment sometimes occurs, but I was uneasy. Adjustment is a deliberate and conscious activity, but in most cases of anchoring there is no corresponding subjective experience. Consider these two questions:

Was Gandhi more or less than 144 years old when he died?

How old was Gandhi when he died?

Did you produce your estimate by adjusting down from 144? Probably not, but the absurdly high number still affected your estimate. My hunch was that anchoring is a case of suggestion. This is the word we use when someone causes us to see, hear, or feel something by merely bringing it to mind. For example, the question “Do you now feel a slight numbness in your left leg?” always prompts quite a few people to report that their left leg does indeed feel a little strange.

Amos was more conservative than I was about hunches, and he correctly pointed out that appealing to suggestion did not help us understand anchoring, because we did not know how to explain suggestion. I had to agree that he was right, but I never became enthusiastic about the idea of insufficient adjustment as the sole cause of anchoring effects. We conducted many inconclusive experiments in an effort to understand anchoring, but we failed and eventually gave up the idea of writing more about it.

The puzzle that defeated us is now solved, because the concept of suggestion is no longer obscure: suggestion is a priming effect, which selectively evokes compatible evidence. You did not believe for a moment that Gandhi lived for 144 years, but your associative machinery surely generated an impression of a very ancient person. System 1 understands sentences by trying to make them true, and the selective activation of compatible thoughts produces a family of systematic errors that make us gullible and prone to believe too strongly whatever we believe. We can now see why Amos and I did not realize that there were two types of anchoring: the research techniques and theoretical ideas we needed did not yet exist. They were developed, much later, by other people. A process that resembles suggestion is indeed at work in many situations: System 1 tries its best to construct a world in which the anchor is the true number. This is one of the manifestations of associative coherence that I described in the first part of the book.

The German psychologists Thomas Mussweiler and Fritz Strack offered the most compelling demonstrations of the role of associative coherence in anchoring. In one experiment, they asked an anchoring question about temperature: "Is the annual mean temperature in Germany higher or lower than 20°C (68°F)?" or "Is the annual mean temperature in Germany higher or lower than 5°C (40°F)?"

All participants were then briefly shown words that they were asked to identify. The researchers found that 68°F made it easier to recognize summer words (like *sun* and *beach*), and 40°F facilitated winter words (like *frost* and *ski*). The selective activation of compatible memories explains anchoring: the high and the low numbers activate different sets of ideas in memory. The estimates of annual temperature draw on these biased samples of ideas and are therefore biased as well. In another elegant study in the same vein, participants were asked about the average price of German cars. A high anchor selectively primed the names of luxury brands (Mercedes, Audi), whereas the low anchor primed brands associated with mass-market cars (Volkswagen). We saw earlier that any prime will tend to evoke information that is compatible with it. Suggestion and anchoring are both explained by the same automatic operation of System 1. Although I did not know how to prove it at the time, my hunch about the link between anchoring and suggestion turned out to be correct.

The Anchoring Index

Many psychological phenomena can be demonstrated experimentally, but few can actually be measured. The effect of anchors is an exception. Anchoring can be measured, and it is an impressively large effect. Some visitors at the San Francisco Exploratorium were asked the following two questions:

Is the height of the tallest redwood more or less than 1,200 feet?

What is your best guess about the height of the tallest redwood?

The "high anchor" in this experiment was 1,200 feet. For other participants, the first question referred to a "low anchor" of 180 feet. The difference between the two anchors was 1,020 feet.

As expected, the two groups produced very different mean estimates: 844 and 282 feet. The difference between them was 562 feet. The anchoring index is simply the ratio of the two differences (562/1,020) expressed as a percentage: 55%. The anchoring measure would be 100% for people who slavishly adopt the anchor as an estimate, and zero for people who are able to ignore the anchor altogether. The value of 55% that was observed in this example is typical. Similar values have been observed in numerous other problems.

The anchoring effect is not a laboratory curiosity; it can be just as strong in the real world. In an experiment conducted some years ago, real-estate agents were given an opportunity to assess the value of a house that was actually on the market. They visited the house and studied a comprehensive booklet of information that included an asking price. Half the agents saw an asking price that was substantially higher than the listed price of the house; the other half saw an asking price that was substantially lower. Each agent gave her opinion about a reasonable buying price for the house and the lowest price at which she would agree to sell the house if she owned it. The agents were then asked about the factors that had affected their judgment. Remarkably, the asking price was not one of these factors; the agents took pride in their ability to ignore it. They insisted that the listing price had no effect on their responses, but they were wrong: the anchoring effect was 41%. Indeed, the professionals were almost as susceptible to anchoring effects as business school students with no real-estate experience, whose anchoring index was 48%. The only difference between the two groups was that the students conceded that they were influenced by the anchor, while the professionals denied that influence.

Powerful anchoring effects are found in decisions that people make about money, such as when they choose how much to contribute al.ls denied to a cause. To demonstrate this effect, we told participants in the Exploratorium study about the environmental damage caused by oil tankers in the Pacific Ocean and asked about their willingness to make an annual contribution “to save 50,000 offshore Pacific Coast seabirds from small offshore oil spills, until ways are found to prevent spills or require tanker owners to pay for the operation.” This question requires intensity matching: the respondents are asked, in effect, to find the dollar amount of a contribution that matches the intensity of their feelings about the plight of the seabirds. Some of the visitors were first asked an anchoring question, such as, “Would you be willing to pay $5…,” before the point-blank question of how much they would contribute.

When no anchor was mentioned, the visitors at the Exploratorium—generally an environmentally sensitive crowd—said they were willing to pay $64, on average. When the anchoring amount was only $5, contributions averaged $20. When the anchor was a rather extravagant $400, the willingness to pay rose to an average of $143.

The difference between the high-anchor and low-anchor groups was $123. The anchoring effect was above 30%, indicating that increasing the initial request by $100 brought a return of $30 in average willingness to pay.

Similar or even larger anchoring effects have been obtained in numerous studies of estimates and of willingness to pay. For example, French residents of the heavily polluted Marseilles region were asked what increase in living costs they would accept if they could live in a less polluted region. The anchoring effect was over 50% in that study. Anchoring effects are easily observed in online trading, where the same item is often offered at different “buy now” prices. The “estimate” in fine-art auctions is also an anchor that influences the first bid.

There are situations in which anchoring appears reasonable. After all, it is not surprising that people who are asked difficult questions clutch at straws, and the anchor is a plausible straw. If you know next to nothing about the trees of California and are asked whether a redwood can be taller than 1,200 feet, you might infer that this number is not too far from the truth. Somebody who knows the true height thought up that question, so the anchor may be a valuable hint. However, a key finding of anchoring research is that anchors that are obviously random can be just as effective as potentially informative anchors. When we used a wheel of fortune to anchor estimates of the proportion of African nations in the UN, the anchoring index was 44%, well within the range of effects observed with anchors that could plausibly be taken as hints. Anchoring effects of similar size have been observed in experiments in which the last few digits of the respondent's Social Security number was used as the anchor (e.g., for estimating the number of physicians in their city). The conclusion is clear: anchors do not have their effects because people believe they are informative.

The power of random anchors has been demonstrated in some unsettling ways. German judges with an average of more than fifteen years of experience on the bench first read a description of a woman who had been caught shoplifting, then rolled a pair of dice that were loaded so every roll resulted in either a 3 or a 9. As soon as the dice came to a stop, the judges were asked whether they would sentence the woman to a term in prison greater or lesser, in months, than the number showing on the dice. Finally, the judges were instructed to specify the exact prison sentence they would give to the shoplifter. On average, those who had rolled a 9 said they would sentence her to 8 months; those who rolled a 3 saidthif Africa they would sentence her to 5 months; the anchoring effect was 50%.

Uses and Abuses of Anchors

By now you should be convinced that anchoring effects—sometimes due to priming, sometimes to insufficient adjustment—are everywhere. The psychological mechanisms that produce anchoring make us far more suggestible than most of us would want to be. And of course there are quite a few people who are willing and able to exploit our gullibility.

Anchoring effects explain why, for example, arbitrary rationing is an effective marketing ploy. A few years ago, supermarket shoppers in Sioux City, Iowa, encountered a sales promotion for Campbell's soup at about 10% off the regular price. On some days, a sign on the shelf said limit of 12 per person. On other days, the sign said no limit per person. Shoppers purchased an average of 7 cans when the limit was in force, twice as many as they bought when the limit was removed. Anchoring is not the sole explanation. Rationing also implies that the goods are flying off the shelves, and shoppers should feel some urgency about stocking up. But we also know that the mention of 12 cans as a possible purchase would produce anchoring even if the number were produced by a roulette wheel.

We see the same strategy at work in the negotiation over the price of a home, when the seller makes the first move by setting the list price. As in many other games, moving first is an advantage in single-issue negotiations—for example, when price is the only issue to be settled between a buyer and a seller. As you may have experienced when negotiating for the first time in a bazaar, the initial anchor has a powerful effect. My advice to students when I taught negotiations was that if you think the other side has made an outrageous proposal, you should not come back with an equally outrageous counteroffer, creating a gap that will be difficult to bridge in further negotiations. Instead you should make a scene, storm out or threaten to do so, and make it clear—to yourself as well as to the other side—that you will not continue the negotiation with that number on the table.

The psychologists Adam Galinsky and Thomas Mussweiler proposed more subtle ways to resist the anchoring effect in negotiations. They instructed negotiators to focus their attention and search their memory for arguments against the anchor. The instruction to activate System 2 was successful. For example, the anchoring effect is reduced or eliminated when the second mover focuses his attention on the minimal offer that the opponent would accept, or on the costs to the opponent of failing to reach an agreement. In general, a strategy of deliberately "thinking the opposite" may be a good defense against anchoring effects, because it negates the biased recruitment of thoughts that produces these effects.

Finally, try your hand at working out the effect of anchoring on a problem of public policy: the size of damages in personal injury cases. These awards are sometimes very large. Businesses that are frequent targets of such lawsuits, such as hospitals and chemical companies, have lobbied to set a cap on the awards. Before you read this chapter you might have thought that capping awards is certainly good for potential defendants, but now you should not be so sure. Consider the effect of capping awards at $1 million. This rule would eliminate all larger awards, but the anchor would also pull up the size of many awards that would otherwise be much smaller. It would almost certainly benefit serious offenders and large firms much more than small ones.

Anchoring and the Two Systems

The effects of random anchors have much to tell us about the relationship between System 1 and System 2. Anchoring effects have always been studied in tasks of judgment and choice that are ultimately completed by System 2. However, System 2 works on data that is retrieved from memory, in an automatic and involuntary operation of System 1. System 2 is therefore susceptible to the biasing influence of anchors that make some information easier to retrieve. Furthermore, System 2 has no control over the effect and no knowledge of it. The participants who have been exposed to random or absurd anchors (such as Gandhi's death at age 144) confidently deny that this obviously useless information could have influenced their estimate, and they are wrong.

We saw in the discussion of the law of small numbers that a message, unless it is immediately rejected as a lie, will have the same effect on the associative system regardless of its reliability. The gist of the message is the story, which is based on whatever information is available, even if the quantity of the information is slight and its quality is poor: WYSIATI. When you read a story about the heroic rescue of a wounded mountain climber, its effect on your associative memory is much the same if it is a news report or the synopsis of a film. Anchoring results from this associative activation. Whether the story is true, or believable, matters little, if at all. The powerful effect of random anchors is an extreme case of this phenomenon, because a random anchor obviously provides no information at all.

Earlier I discussed the bewildering variety of priming effects, in which your thoughts and behavior may be influenced by stimuli to which you pay no attention at all, and even by stimuli of which you are completely unaware. The main moral of priming research is that our thoughts and our behavior are influenced, much more than we know or want, by the environment of the moment. Many people find the priming results unbelievable, because they do not correspond to subjective experience. Many others find the results upsetting, because they threaten the subjective sense of agency and autonomy. If the content of a screen saver on an irrelevant computer can affect your willingness to help strangers without your being aware of it, how free are you? Anchoring effects are threatening in a similar way. You are always aware of the anchor and even pay attention to it, but you do not know how it guides and constrains your thinking, because you cannot imagine how you would have thought if the anchor had been different (or absent). However, you should assume that any number that is on the table has had an anchoring effect on you, and if the stakes are high you should mobilize yourself (your System 2) to combat the effect.

Speaking of Anchors

"The firm we want to acquire sent us their business plan, with the revenue they expect. We shouldn't let that number influence our thinking. Set it aside."

"Plans are best-case scenarios. Let's avoid anchoring on plans when we forecast actual outcomes. Thinking about ways the plan could go wrong is one way to do it."

"Our aim in the negotiation is to get them anchored on this number."

& st

"The defendant's lawyers put in a frivolous reference in which they mentioned a ridiculously low amount of damages, and they got the judge anchored on it!"

P

The Science of Availability

Amos and I had our most productive year in 1971–72, which we spent in Eugene, Oregon. We were the guests of the Oregon Research Institute, which housed several future stars of all the fields in which we worked—judgment, decision making, and intuitive prediction. Our main host was Paul Slovic, who had been Amos's classmate at Ann Arbor and remained a lifelong friend. Paul was on his way to becoming the leading psychologist among scholars of risk, a position he has held for decades, collecting many honors along the way. Paul and his wife, Roz, introduced us to life in Eugene, and soon we were doing what people in Eugene do—jogging, barbecuing, and taking children to basketball games. We also worked very hard, running dozens of experiments and writing our articles on judgment heuristics. At night I wrote *Attention and Effort*. It was a busy year.

One of our projects was the study of what we called the *availability heuristic*. We thought of that heuristic when we asked ourselves what people actually do when they wish to estimate the frequency of a category, such as "people who divorce after the age of 60" or "dangerous plants." The answer was straightforward: instances of the class will be retrieved from memory, and if retrieval is easy and fluent, the category will be judged to be large. We defined the availability heuristic as the process of judging frequency by "the ease with which instances come to mind." The statement seemed clear when we formulated it, but the concept of availability has been refined since then. The two-system approach had not yet been developed when we studied availability, and we did not attempt to determine whether this heuristic is a deliberate problem-solving strategy or an automatic operation. We now know that both systems are involved.

A question we considered early was how many instances must be retrieved to get an impression of the ease with which they come to mind. We now know the answer: none. For an example, think of the number of words that can be constructed from the two sets of letters below.

XUZONLCJM

TAPCERHOB

You knew almost immediately, without generating any instances, that one set offers far more possibilities than the other, probably by a factor of 10 or more. Similarly, you do not need to retrieve specific news stories to have a good idea of the relative frequency with which different countries have appeared in the news during the past year (Belgium, China, France, Congo, Nicaragua, Romania…).

The availability heuristic, like other heuristics of judgment, substitutes one question for another: you wish to estimate the size se ost c d of a category or the frequency of an event, but you report an impression of the ease with which instances come to mind. Substitution of questions inevitably produces systematic errors. You can discover how the heuristic leads to biases by following a simple procedure: list factors other than frequency that make it easy to come up with instances. Each factor in your list will be a potential source of bias. Here are some examples:

- A salient event that attracts your attention will be easily retrieved from memory. Divorces among Hollywood celebrities and sex scandals among politicians attract much attention, and instances will come easily to mind. You are therefore likely to exaggerate the frequency of both Hollywood divorces and political sex scandals.
- A dramatic event temporarily increases the availability of its category. A plane crash that attracts media coverage will temporarily alter your feelings about the safety of flying. Accidents are on your mind, for a while, after you see a car burning at the side of the road, and the world is for a while a more dangerous place.
- Personal experiences, pictures, and vivid examples are more available than incidents that happened to others, or mere words, or statistics. A judicial error that affects you will undermine your faith in the justice system more than a similar incident you read about in a newspaper.

Resisting this large collection of potential availability biases is possible, but tiresome. You must make the effort to reconsider your impressions and intuitions by asking such questions as, "Is our belief that theft s by teenagers are a major problem due to a few recent instances in our neighborhood?" or "Could it be that I feel no need to get a flu shot because none of my acquaintances got the flu last year?" Maintaining one's vigilance against biases is a chore—but the chance to avoid a costly mistake is sometimes worth the effort.

One of the best-known studies of availability suggests that awareness of your own biases can contribute to peace in marriages, and probably in other joint projects. In a famous study, spouses were asked, "How large was your personal contribution to keeping the place tidy, in percentages?" They also answered similar questions about "taking out the garbage," "initiating social engagements," etc. Would the self-estimated contributions add up to 100%, or more, or less? As expected, the self-assessed contributions added up to more than 100%. The explanation is a simple *availability bias*: both spouses remember their own individual efforts and contributions much more clearly than those of the other, and the difference in availability leads to a difference in judged frequency. The bias is not necessarily self-serving: spouses also overestimated their contribution to causing quarrels, although to a smaller extent than their contributions to more desirable outcomes. The same bias contributes to the common observation that many members of a collaborative team feel they have done more than their share and also feel that the others are not adequately grateful for their individual contributions.

I am generally not optimistic about the potential for personal control of biases, but this is an exception. The opportunity for successful debiasing exists because the circumstances in which issues of credit allocation come up are easy to identify, the more so because tensions often arise when several people at once feel that their efforts are not adequately recognized. The mere observation that there is usually more than 100% credit to go around is sometimes sufficient to defuse the situation. In any eve#82ght=nt, it is a good thing for every individual to remember. You will occasionally do more than your share, but it is useful to know that you are likely to have that feeling even when each member of the team feels the same way.

The Psychology of Availability

A major advance in the understanding of the availability heuristic occurred in the early 1990s, when a group of German psychologists led by Norbert Schwarz raised an intriguing question: How will people's impressions of the frequency of a category be affected by a requirement to list a specified number of instances? Imagine yourself a subject in that experiment:

First, list six instances in which you behaved assertively.

Next, evaluate how assertive you are.

Imagine that you had been asked for twelve instances of assertive behavior (a number most people find difficult). Would your view of your own assertiveness be different?

Schwarz and his colleagues observed that the task of listing instances may enhance the judgments of the trait by two different routes:

- the number of instances retrieved
- the ease with which they come to mind

The request to list twelve instances pits the two determinants against each other. On the one hand, you have just retrieved an impressive number of cases in which you were assertive. On the other hand, while the first three or four instances of your own assertiveness probably came easily to you, you almost certainly struggled to come up with the last few to complete a set of twelve; fluency was low. Which will count more—the amount retrieved or the ease and fluency of the retrieval?

The contest yielded a clear-cut winner: people who had just listed twelve instances rated themselves as less assertive than people who had listed only six. Furthermore, participants who had been asked to list twelve cases in which they had *not* behaved assertively ended up thinking of themselves as quite assertive! If you cannot easily come up with instances of meek behavior, you are likely to conclude that you are not meek at all. Self-ratings were dominated by the ease with which examples had come to mind. The experience of fluent retrieval of instances trumped the number retrieved.

An even more direct demonstration of the role of fluency was offered by other psychologists in the same group. All the participants in their experiment listed six instances of assertive (or nonassertive) behavior, while maintaining a specified facial expression. “Smilers” were instructed to contract the zygomaticus muscle, which produces a light smile; “frowners” were required to furrow their brow. As you already know, frowning normally accompanies cognitive strain and the effect is symmetric: when people are instructed to frown while doing a task, they actually try harder and experience greater cognitive strain. The researchers anticipated that the frowners would have more difficulty retrieving examples of assertive behavior and would therefore rate themselves as relatively lacking in assertiveness. And so it was.

Psychologists enjoy experiments that yield paradoxical results, and they have appliserv heighted Schwarz’s discovery with gusto. For example, people:

- believe that they use their bicycles less often after recalling many rather than few instances
- are less confident in a choice when they are asked to produce more arguments to support it
- are less confident that an event was avoidable after listing more ways it could have been avoided
- are less impressed by a car after listing many of its advantages

A professor at UCLA found an ingenious way to exploit the availability bias. He asked different groups of students to list ways to improve the course, and he varied the required number of improvements. As expected, the students who listed more ways to improve the class rated it higher!

Perhaps the most interesting finding of this paradoxical research is that the paradox is not always found: people sometimes go by content rather than by ease of retrieval. The proof that you truly understand a pattern of behavior is that you know how to reverse it. Schwarz and his colleagues took on this challenge of discovering the conditions under which this reversal would take place.

The ease with which instances of assertiveness come to the subject's mind changes during the task. The first few instances are easy, but retrieval soon becomes much harder. Of course, the subject also expects fluency to drop gradually, but the drop of fluency between six and twelve instances appears to be steeper than the participant expected. The results suggest that the participants make an inference: if I am having so much more trouble than expected coming up with instances of my assertiveness, then I can't be very assertive. Note that this inference rests on a surprise—fluency being worse than expected. The availability heuristic that the subjects apply is better described as an "unexplained unavailability" heuristic.

Schwarz and his colleagues reasoned that they could disrupt the heuristic by providing the subjects with an explanation for the fluency of retrieval that they experienced. They told the participants they would hear background music while recalling instances and that the music would affect performance in the memory task. Some subjects were told that the music would help, others were told to expect diminished fluency. As predicted, participants whose experience of fluency was "explained" did not use it as a heuristic; the subjects who were told that music would make retrieval more difficult rated themselves as equally assertive when they retrieved twelve instances as when they retrieved six. Other cover stories have been used with the same result: judgments are no longer influenced by ease of retrieval when the experience of fluency is given a spurious explanation by the presence of curved or straight text boxes, by the background color of the screen, or by other irrelevant factors that the experimenters dreamed up.

As I have described it, the process that leads to judgment by availability appears to involve a complex chain of reasoning. The subjects have an experience of diminishing fluency as they produce instances. They evidently have expectations about the rate at which fluency decreases, and those expectations are wrong: the difficulty of coming up with new instances increases more rapidly than they expect. It is the unexpectedly low fluency that causes people who were asked for twelve instances to describe themselves as unassertive. When the surprise is eliminated, low fluency no longer influences the judgment. The process appears to consist of a sophisticatedriethe subj set of inferences. Is the automatic System 1 capable of it?

The answer is that in fact no complex reasoning is needed. Among the basic features of System 1 is its ability to set expectations and to be surprised when these expectations are violated. The system also retrieves possible causes of a surprise, usually by finding a possible cause among recent surprises. Furthermore, System 2 can reset the expectations of System 1 on the fly, so that an event that would normally be surprising is now almost normal. Suppose you are told that the three-year-old boy who lives next door frequently wears a top hat in his stroller. You will be far less surprised when you actually see him with his top hat than you would have been without the warning. In Schwarz's experiment, the background music has been mentioned as a possible cause of retrieval problems. The difficulty of retrieving twelve instances is no longer a surprise and therefore is less likely to be evoked by the task of judging assertiveness.

Schwarz and his colleagues discovered that people who are personally involved in the judgment are more likely to consider the number of instances they retrieve from memory and less likely to go by fluency. They recruited two groups of students for a study of risks to cardiac health. Half the students had a family history of cardiac disease and were expected to take the task more seriously than the others, who had no such history. All were asked to recall either three or eight behaviors in their routine that could affect their cardiac health (some were asked for risky behaviors, others for protective behaviors). Students with no family history of heart disease were casual about the task and followed the availability heuristic. Students who found it difficult to find eight instances of risky behavior felt themselves relatively safe, and those who struggled to retrieve examples of safe behaviors felt themselves at risk. The students with a family history of heart disease showed the opposite pattern—they felt safer when they retrieved many instances of safe behavior and felt greater danger when they retrieved many instances of risky behavior. They were also more likely to feel that their future behavior would be affected by the experience of evaluating their risk.

The conclusion is that the ease with which instances come to mind is a System 1 heuristic, which is replaced by a focus on content when System 2 is more engaged. Multiple lines of evidence converge on the conclusion that people who let themselves be guided by System 1 are more strongly susceptible to availability biases than others who are in a state of higher vigilance. The following are some conditions in which people "go with the flow" and are affected more strongly by ease of retrieval than by the content they retrieved:

- when they are engaged in another effortful task at the same time
- when they are in a good mood because they just thought of a happy episode in their life
- if they score low on a depression scale
- if they are knowledgeable novices on the topic of the task, in contrast to true experts
- when they score high on a scale of faith in intuition
- if they are (or are made to feel) powerful

I find the last finding particularly intriguing. The authors introduce their article with a famous quote: "I don't spend a lot of time taking polls around the world to tell me what I think is the right way to act. I've just got to know how I feel" (Georgee e the w W. Bush, November 2002). They go on to show that reliance on intuition is only in part a personality trait. Merely reminding people of a time when they had power increases their apparent trust in their own intuition.

Speaking of Availability

> "Because of the coincidence of two planes crashing last month, she now prefers to take the train. That's silly. The risk hasn't really changed; it is an availability bias."