

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 the *insur* 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 associations, 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 which 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 psychological 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.”