

The Feeling of Knowing

I AM STUCK IN AN OBLIGATORY NEIGHBORHOOD COCKTAIL party during the first week of the U.S. invasion of Iraq. A middle-aged, pin-striped lawyer announces that he'd love to be in the front lines when the troops reach Baghdad. "Door-to-door fighting," he says, puffing up his chest. He says he's certain he could shoot an Iraqi soldier, although he's never been in a conflict bigger than a schoolyard brawl.

"I don't know," I say. "I'd have trouble shooting some young kid who was being forced to fight."

"Not me. We're down to dog-eat-dog."

He nods at his frowning wife, who's anti-invasion. "All's fair in love and war." Then back to me. "You're not one of those peacenik softies, are you?"

"It wouldn't bother you to kill someone?"

"Not a bit."

"You're sure?"

"Absolutely."

He's a neighbor and I can't escape. So I tell him one of my father's favorite self-mocking stories.

During the 1930s and '40s, my father had a pharmacy in one of the tougher areas of San Francisco. He kept a small revolver hidden beneath the back cash register. One night, a man approached, pulled out a knife, and demanded all the money in the register. My father reached under the counter, grabbed his gun, and aimed it at the robber.

"Drop it," the robber said, his knife at my father's throat. "You're not going to shoot me, but I *will* kill you."

For a moment it was a Hollywood standoff, *mano a mano*. Then my father put down his gun, emptied out the register, and handed over the money.

"What's your point?" the lawyer asks. "Your father should have shot him."

"Just the obvious," I say. "You don't always know what you're going to do until you're in the moment."

"Sure you do. I know with absolute certainty that I'd shoot anyone who was threatening me."

"No chance of any hesitation?"

"None at all. I know myself. I know what I would do. End of discussion."

MY MIND REELS with seemingly impossible questions. What kind of knowledge is "I know myself and what I would do"? Is it a conscious decision based upon deep self-contemplation or is it a "gut feeling"? But what is a gut feeling—an unconscious decision, a mood or emotion, an ill-defined but clearly recognizable mental state, or a combination of all these ingredients? If we are to

understand how we know what we know, we first need some ground rules, including a general classification of mental states that create our sense of knowledge about our knowledge.

For simplicity, I have chosen to lump together the closely allied feelings of certainty, rightness, conviction, and correctness under the all-inclusive term, the *feeling of knowing*. Whether or not these are separate sensations or merely shades or degrees of a common feeling isn't important. What they do share is a common quality: Each is a form of metaknowledge—knowledge about our knowledge—that qualifies or colors our thoughts, imbuing them with a sense of rightness or wrongness. When focusing on the phenomenology (how these sensations *feel*), I've chosen to use the term the *feeling of knowing* (in italics). However, when talking about the underlying science, I'll use *knowing* (in italics). Later I will expand this category to include feelings of familiarity and realness—qualities that enhance our sense of correctness.

EVERYONE IS FAMILIAR with the most commonly recognized *feeling of knowing*. When asked a question, you feel strongly that you know an answer that you cannot immediately recall. Psychologists refer to this hard-to-describe but easily recognizable feeling as a tip-of-the-tongue sensation. The frequent accompanying comment as you scan your mental Rolodex for the forgotten name or phone number: "I know it, but I just can't think of it." In this example, you are aware of knowing something, without knowing what this sense of knowing refers to.

Anyone who's been frustrated with a difficult math problem has appreciated the delicious moment of relief when an incomprehensible equation suddenly *makes sense*. We "see the light." This *aha* is a notification from a subterranean portion of our

mind, an involuntary all-clear signal that we have grasped the heart of a problem. It isn't just that we can solve the problem; we also "know" that we understand it.

Most *feelings of knowing* are far less dramatic. We don't ordinarily sense them as spontaneous emotions or moods like love or happiness; rather they feel like thoughts—elements of a correct line of reasoning. We learn to add $2 + 2$. Our teacher tells us that 4 is the correct answer. Yes, we hear a portion of our mind say. Something within us tells us that we "know" that our answer is correct. At this simplest level of understanding, there are two components to our understanding—the knowledge that $2 + 2 = 4$, and the judgment or assessment of this understanding. We know that our understanding that $2 + 2 = 4$ is itself correct.

The *feeling of knowing* is also commonly recognized by its absence. Most of us are all too familiar with the frustration of being able to operate a computer without having any "sense" of how the computer really works. Or learning physics despite having no "feeling" for the rightness of what you've learned. I can fix a frayed electrical cord, yet am puzzled by the very essence of electricity. I can pick up iron filings with a magnet without having the slightest sense of what magnetism "is."

At a deeper level, most of us have agonized over those sickening "crises of faith" when firmly held personal beliefs are suddenly stripped of a visceral sense of correctness, rightness, or meaning. Our most considered beliefs suddenly don't "feel right." Similarly, most of us have been shocked to hear that a close friend or relative has died unexpectedly, and yet we "feel" that he is still alive. Such upsetting news often takes time to "sink in." This disbelief associated with hearing about a death is an example of the sometimes complete disassociation between intellectual and felt knowledge.

To begin our discussion of the *feeling of knowing*, read the following excerpt at normal speed. Don't skim, give up halfway through, or skip to the explanation. Because this experience can't be duplicated once you know the explanation, take a moment to ask yourself how you feel about the paragraph. After reading the clarifying word, reread the paragraph. As you do so, please pay close attention to the shifts in your mental state and your feeling about the paragraph.

A newspaper is better than a magazine. A seashore is a better place than the street. At first it is better to run than to walk. You may have to try several times. It takes some skill, but it is easy to learn. Even young children can enjoy it. Once successful, complications are minimal. Birds seldom get too close. Rain, however, soaks in very fast. Too many people doing the same thing can also cause problems. One needs lots of room. If there are no complications, it can be very peaceful. A rock will serve as an anchor. If things break loose from it, however, you will not get a second chance.

Is this paragraph comprehensible or meaningless? Feel your mind sort through potential explanations. Now watch what happens with the presentation of a single word: kite. As you reread the paragraph, feel the prior discomfort of something amiss shifting to a pleasing sense of rightness. Everything fits; every sentence works and has meaning. Reread the paragraph again; it is impossible to regain the sense of not understanding. In an instant, without due conscious deliberation, the paragraph has been irreversibly infused with a *feeling of knowing*.

Try to imagine other interpretations for the paragraph. Suppose

[6] On Being Certain

I tell you that this is a collaborative poem written by a third-grade class, or a collage of strung-together fortune cookie quotes. Your mind balks. The presence of this *feeling of knowing* makes contemplating alternatives physically difficult.

Each of us probably read the paragraph somewhat differently, but certain features seem universal. After seeing the word *kite*, we quickly go back and reread the paragraph, testing the sentences against this new piece of information. At some point, we are convinced. But when and how?

The kite paragraph raises several questions central to our understanding of how we “know” something. Though each will be discussed at greater length in subsequent chapters, here’s a sneak preview.

- Did you consciously “decide” that *kite* was the correct explanation for the paragraph, or did this decision occur involuntarily, outside of conscious awareness?
- What brain mechanism(s) created the shift from not knowing to *knowing*?
- When did this shift take place? (Did you know that the explanation was correct before, during, or after you reread the paragraph?)
- After rereading the paragraph, are you able to consciously separate out the *feeling of knowing* that *kite* is the correct answer from a reasoned understanding that the answer is correct?
- Are you sure that *kite* is the correct answer? If so, how do you know?

2

How Do We Know What We Know?

PARENTS' AND TEACHERS' CUSTOMARY ADVICE FOR "NOT getting" math and physics is to study harder and think more deeply about the problem. Their assumption is that more effort will bridge the gap between dry knowledge and felt understanding. Without this assumption, we would give up every time we failed to understand something at first glance. But for those "what's the point of it all" existential moments—when formerly satisfactory feelings of purpose and meaning no longer "feel right"—history and experience have taught us differently. Logic and reason rarely are "convincing." (In this context, "convincing" is synonymous with reviving this missing "*feeling of knowing* what life is about.") Instead, we conjure up images of ascetics, mystics, and spiritual seekers—those who have donned hair shirts, trekked through the desert à la St. Jerome, huddled in caves or under trees, or sought isolation and silence in monasteries. Eastern religions emphasize a "stillness of the mind" rather than actively thinking about the missing sense of meaning.

So, which is it? Should the remedy for the absence of the *feeling of knowing* be more conscious effort and hard thought, or less? Or are both of these common teachings at odds with more basic neurobiology? Consider the curious phenomenon of *blindsight*, perhaps the best-studied example of the lack of the *feeling of knowing* in the presence of a state of knowledge.

Out of Sight Is Not Out of Mind

A patient has a stroke that selectively destroys his occipital cortex—the portion of the brain that receives primary visual inputs. His retina still records incoming information, but his malfunctioning visual cortex does not register the images sent from the retina. The result is that the patient consciously sees nothing. Now flash a light in various quadrants of his visual field. The patient reports that he sees nothing, yet he can fairly accurately localize the flashing light to the appropriate quadrant. He feels that he is guessing and is unaware that he is performing any better than by chance.

How is this possible?

First, let's trace the pathway of the "unseen" light. Some fibers from the retina proceed directly to the primary visual cortex in the occipital lobe. But other fibers bypass the region responsible for conscious "seeing" and instead project to subcortical and upper brain stem regions that do not produce a visual image. These lower brain areas are primarily concerned with automatic, reflexive functions such as fight-or-flight. Quickly approaching or looming objects cause the body to swing the head into position so that the eyes can examine the threat. An immediate reflexive action has clear evolutionary benefits over more time-consuming conscious

perception and deliberation. In the broadest sense, you could say that these subcortical regions “see” the threat without sending a visual image into awareness.

Blindsight is a primitive unconscious visual localization and navigation system uncovered by the patient’s cortical blindness. The patient’s subliminal knowledge of the location of the flashing light doesn’t trigger the *feeling of knowing* because news of this knowledge can’t reach the higher cortical regions that generate the feeling. As a result, the patient swears that he hasn’t seen a flashing light, yet he clearly possesses a subliminal knowledge of the light’s location. When he chooses the proper visual field for the flashing light, he has no feeling that this is a correct answer. *He does not know what he knows.*¹

With blindsight, we see the disconnect between knowledge and awareness of this knowledge as being related to a fundamental flaw in our circuitry. This broken connection cannot be restored either through conscious effort or stiling of the mind—the problem is not within our control.

Though clinically apparent blindsight is a rare event usually caused by a stroke that interferes with the blood supply to the occipital cortex, faulty expressions of the *feeling of knowing* are everyday occurrences. Let’s begin with our own memories.

The Challenger Study

Try to remember where you were when Kennedy was assassinated, the *Challenger* blew up, or the World Trade Center was attacked. Now ask yourself how certain you are of those memories. If you believe that you are quite sure of where you were when you heard the news, keep that feeling in mind as you read about

the *Challenger* study in the next pages. If you don't remember where you were, ask yourself how you know that you don't remember. (Keep in mind the blindsight example when asking this question.) Either way, try to understand the feeling and your degree of certainty of this memory.

At my most recent med school reunion dinner, several former classmates were recalling where they were when Kennedy was assassinated. We had been in the second year of medical school, which meant that we all went to the same classes. Wherever one was, we probably all were. But the recollections were strikingly different; after dinner the discussion was becoming increasingly heated, as though each classmate's mind was on trial. A urologist thought we were at lunch, an internist said we were in the lab. A pathologist remembered being at a pub down the street from the med center. "That can't be true," the urologist said. "The assassination was at noon, Dallas time. You didn't go to the bars 'til after class."

I laughed and briefly described the *Challenger* study.²

Within one day of the space shuttle *Challenger* explosion, Ulric Neisser, a psychologist studying "flashbulb" memories (the recall of highly dramatic events), asked his class of 106 students to write down exactly how they'd heard about the explosion, where they were, what they'd been doing, and how they felt. Two and a half years later they were again interviewed. Twenty-five percent of the students' subsequent accounts were strikingly different than their original journal entries. More than half the people had lesser degrees of error, and less than ten percent had all the details correct. (Prior to seeing their original journals, most students presumed that their memories were correct.)

Most of us reluctantly admit that memory changes³ over time.

As kids, we saw how a story changed with retellings around a campfire. We have been at enough family reunions to hear once-familiar shared events morphed into unrecognizable and often contradictory descriptions. So, seeing that your journal entries were different than your recollection a couple of years later shouldn't be surprising. What startled me about the *Challenger* study were the students' responses when confronted with their conflicting accounts. Many expressed a high level of confidence that their false recollections were correct, despite being confronted with their own handwritten journals. The most unnerving was one student's comment, "That's my handwriting, but that's not what happened."

Why wouldn't the students consider their journal entries written shortly after the event to be more accurate than a recollection pulled up several years later? Pride, stubbornness, or fear of admitting an error? Not remembering the details of the *Challenger* explosion doesn't imply some massive personal failing that would make resistance to contrary evidence so overwhelming. Conversely, wouldn't pride in being logical and rational steer the students toward choosing their own handwriting over memories that they know might have been altered with time?

The inflamed urologist interrupted me, insisting the pathologist concede that he was wrong. The pathologist refused, turned to me, and said, "You tell them, Burton. You were there in the bar with me."

"Beats me. I just don't remember."

"That's not possible," the two warring doctors said simultaneously. "Everyone remembers the Kennedy assassination."

I shrugged and silently marveled at the vehemence of my classmates' convictions. Even telling them of the *Challenger* study

persuaded no one, as though they were intent upon reproducing the very study that questioned their recollections. All felt that they were right, that they absolutely *knew* where they were and what they were doing when Kennedy was assassinated.

Cognitive Dissonance

In 1957, Stanford professor of social psychology Leon Festinger introduced the term *cognitive dissonance* to describe the distressing mental state in which people “find themselves doing things that don’t fit with what they know, or having opinions that do not fit with other opinions they hold.”³ In a series of clever experiments, Festinger demonstrated that such tensions were more often minimized or resolved through changes in personal attitudes than by relinquishing the dissonant belief or opinion.

As an example, Festinger and his associates described a cult that believed that the earth was going to be destroyed by a flood. When the flood did not happen, those less involved with the cult were more inclined to recognize that they had been wrong. The more invested members who had given up their homes and jobs to work for the cult were more likely to reinterpret the evidence to show that they were right all along, but that the earth was not destroyed because of their faithfulness.⁴

Festinger’s seminal observation: The more committed we are to a belief, the harder it is to relinquish, even in the face of overwhelming contradictory evidence. Instead of acknowledging an error in judgment and abandoning the opinion, we tend to develop a new attitude or belief that will justify retaining it. By giving us a model to consider how we deal with conflicting values, the theory of cognitive dissonance has become one of the most

influential theories in social psychology. Yet it fails to convincingly answer why it is so difficult to relinquish unreasonable opinions, especially in light of seemingly convincing contrary evidence. It is easy to dismiss such behavior in cult members and others “on the fringe,” but what about those of us who presume ourselves to be less flaky, those of us who pride ourselves on being levelheaded and reasonable?

WE MIGHT THINK of the *Challenger* study as an oddity, but here are additional examples of consciously choosing a false belief because it *feels* correct even when we know better. I have chosen the first example as a prelude to a later discussion in chapter 13 of the deeply rooted biological component of the science-versus-religion struggle. The second example, highlighting the cognitive dissonance of the placebo effect, introduces the idea that an unjustified *feeling of knowing* can have a clear adaptive benefit.

A Scientist Contemplates Creationism

Kurt Wise, with a B.A. in geophysics from the University of Chicago, a Ph.D. in geology from Harvard, where he studied under Steven Jay Gould, and a professorship at Bryan College in Dayton, Tennessee, writes of his personal conflict between science and religion.⁵

I had to make a decision between evolution and Scripture. Either the Scripture was true and evolution was wrong or evolution was true and I must toss out the Bible. . . . It was there that night that I accepted the Word of God and rejected all that would ever

counter it, including evolution. With that, in great sorrow, I tossed into the fire all my dreams and hopes in science. . . . *If all the evidence in the universe turns against creationism, I would be the first to admit it, but I would still be a creationist because that is what the Word of God seems to indicate.* (Italics mine.)

A Patient Confronts the Placebo Effect

In a study of 180 people with osteoarthritis of the knee, a team of Houston surgeons headed by Bruce Moseley, M.D., found that patients who had “sham” arthroscopic surgery reported as much pain relief and improved mobility as patients who actually underwent the procedure.⁶

Mr. A, a seventy-six-year-old retired World War II veteran with a five-year history of disabling knee pain from X-ray-documented degenerative osteoarthritis was assigned to the placebo group (sham surgery in which general anesthesia was given, superficial incisions were made in the skin over the knee, but no actual surgical repair was performed). After the procedure, Mr. A was informed that he had received sham surgery; the procedure was described in detail. Nevertheless, he dramatically improved; for the first time in years he was able to walk without a cane. When questioned, he both fully understood what sham surgery meant and fully believed that his knee had been fixed.

“The surgery was two years ago and the knee has never bothered me since. It’s just like my other knee now. I give a whole lot of credit to Dr. Moseley. Whenever I see him on the TV, I call the wife in and say, ‘Hey, there’s the doctor that fixed my knee!’ ”⁷

Our creationist geologist cringes at his own irrationality and yet

declares that he does not have a choice. A patient “knows” that he hasn’t had any reparative surgery performed, yet insists that the doctor fixed his knee. What if we could find patients who developed similar difficulties with reason as the result of specific brain insults (lesions)? If brain malfunctions can produce a similar flawed logic, what might that tell us about the biological underpinnings of cognitive dissonances?

Cotard’s Syndrome

Ms. B, a twenty-nine-year-old grad student hospitalized for an acute viral encephalitis (a viral inflammation of the brain) complained: “Nothing feels real. I am dead.” The patient refused any medical care. “There is no point in treating a dead person,” she insisted. Her internist tried to reason with her. He asked her to put her hand on her chest and feel her heart beating. She did, and agreed that her heart was beating. He suggested that the presence of a pulse must mean that she was not dead. The patient countered that, since she was dead, her beating heart could not be evidence for being alive. She said she recognized that there was a logical inconsistency between being dead and being able to feel her beating heart, but that being dead felt more “real” than any contrary evidence that she was alive.

Weeks later, Ms. B began to recover; eventually she no longer believed that she was dead. She was able to make a distinction between her recovered “reality” and her prior delusions, yet she continued to believe that it must be possible to feel one’s heart beat after death. After all, it had happened to her.

Cotard’s syndrome—*le délire de négation*—is attributed to a French psychiatrist, Jules Cotard, who in 1882 described several

patients with delusions of self-negation. These ranged from the belief that parts of the body were missing, or had putrefied, to the complete denial of bodily existence. The syndrome has been described with a variety of brain injuries, strokes, and dementia, as well as severe psychiatric disorders. The most extraordinary element of the syndrome is the patient's unshakable belief in being dead that overpowers any logical counterconclusion. Feeling one's beating heart isn't sufficient evidence to overcome the more powerful sense of the reality of being dead.

Other delusional syndromes associated with acute brain lesions include believing that a friend or a relative has been replaced by an impostor, or a double, or has taken on different appearances or identities, or that an inanimate object has been replaced by an inferior copy. The clinical feature common to all of these syndromes is the inability of the patient to shake a belief that he logically knows is wrong.

Mr. C, an elegant retired art dealer, was hospitalized overnight with a small stroke. The next morning, he felt well and was discharged. Within moments of returning home, he phoned my office in a panic. He was certain that his favorite antique desk had been replaced by a cheap Levitz reproduction. "Hurry over and see for yourself." He lived near my office; I dropped by at lunchtime. The desk in question was a massive eighteenth-century Italian refectory table that took up most of his den. It could easily seat a dozen; just lifting it would require a minimum of several men. And it was far too wide to fit through the doorway without removing the French doors. I quickly pointed out the impossibility of someone sneaking in, moving out the desk, and substituting a fake. Mr. C shook his head. "Yes, I admit that it is physically impossible that the desk has been replaced. But it has. You have to

take my word for it. I know real when I see real, and this desk isn't real." He ran his hand along the grain, repeatedly fingering a couple of prominent wormholes. "It's funny," he said with a puzzled expression. "These are exact replicas of the holes in my desk. But they don't feel the least bit familiar. No," he announced emphatically, "someone must have replaced it." He then delivered the cognitive checkmate: "After all, I know what I know."

Although not restricted to a single area of the brain or a single definitive physiology, the most striking shared characteristic of these *delusional misidentification syndromes* is that the conflict between logic and a contrary *feeling of knowing* tends to be resolved in favor of feeling. Rather than rejecting ideas and beliefs that defy common sense and overwhelming contrary evidence, such patients end up using tortured logic to justify the more powerful sense of *knowing what they know*.⁸

Mr. C's statements also point out that *knowing* may also involve additional hard-to-define mental states such as a sense of familiarity and feelings of "realness." Like the tip-of-the-tongue sensation or the feeling of *déjà vu*, a sense of being familiar suggests some prior experience or knowledge. When stumped on a multiple choice test question, we tend to choose the answer that feels most familiar. Though we have no justification, we presume that such answers are more likely to be correct than those that we don't recognize or seem unfamiliar. Mr. C's "I know real when I see real" points out how a sense of "realness" might also bias us toward believing that an idea is correct. Patients with delusional misidentification syndromes often use "correct" interchangeably with "real."

It is likely that Mr. C's stroke affected his ability to appropriately

experience feelings of familiarity and "realness." When neither the sight nor the feel of the desk triggered these feelings, he was forced to conclude that this desk could not be the original. Such delusions might be seen as an attempt to resolve a cognitive dissonance between hard evidence (the table is too big to move) and the absence of any feeling of familiarity and realness when Mr. C examined his desk.

In chapter 3 we shall see that the mental states of familiarity, "realness," conviction, truth, *déjà vu*, and tip of the tongue share a similar physiology with the *feeling of knowing*, including the ability to be directly triggered with electrical stimulation of the brain's limbic system.

It May Be Right, But It's Not *Right*

The other day, at a downtown garage, I left my car with valet parking. I returned, started to drive away, but felt something was wrong. I questioned the attendant's gaze, wondering if I'd paid too much. I checked the gas and oil gauge, and whether one of the doors was ajar. Then I realized that the seat had been readjusted by the attendant. It was a nominal difference, the seat was at most a half inch higher than usual. My *derriere* knew immediately; it took me considerably longer.

I was reminded of a story attributed to Ludwig Wittgenstein.

A man walks into a tailor's shop. The sign over the front door reads: CUSTOMER SATISFACTION GUARANTEED. The man orders a custom-made suit that should fit exactly like the one he is wearing. The tailor painstakingly measures every detail and jots them down in a notebook. A week later the customer returns to try on the new suit.

"It's not right," the customer says with annoyance. . . .

"Of course it is," the tailor says. "Here, I'll show you." The tailor takes out his measuring tape, compares the suit's readings with those in his notebook. "See, they're identical."

The customer shifts in his new suit but is still uncomfortable and displeased. "It may be right, but it's not *right*." He refuses to pay for the suit and storms out.

In the case of my car seat, I was forced to think through all the possible reasons that I sensed something was wrong. Fortunately, there was something measurable (the new angle of the car seat) that *explained* what I was *feeling*. With the tailor example, the customer's sense of something amiss is a matter of taste, of inexpressible or subconscious aesthetics. No matter what the measurements, the suit does not *feel* right.

The tailor demands his money; the customer admits that the suit was to his specifications, but not to his liking, and therefore he is under no obligation to buy the suit. Each *feels* that he is right. Hence that irritating popular refrain—end of discussion. We often talk about gut feelings. There is now extensive literature on the neuroenteric brain, as though some form of thought might actually originate in the pit of your stomach. Maybe so. And maybe my body just *knew* that my car seat was out of whack. But whatever the origin of the sensation, the key feature is that there seems to be an underlying *sense* or *feeling* that something is either correct or incorrect.

Consider the similarity in tone between the *Challenger* study student who said, "That's my writing, but that's not what happened," and the suit customer's "It might be right, but it's not *right*." When such a sense of conviction overrides obvious logical inconsistencies or scientific evidence, what is happening? Is it possible that

[20] **On Being Certain**

there is an underlying neurophysiological basis for the specific sensation of *feeling right* or of *being right* that is so powerful that ordinary rational thought *feels* either wrong or irrelevant? Conviction versus knowledge—is the jury rigged, the game fixed by a basic physiology hidden beneath awareness?