

# "The Power of Reflection"

by Stephen M. Fleming

Judith Keppel was a single question away from taking home £1 million. She was in the final round of a British quiz show, and she had to face one last challenge to become the show's first victor: "Which king was married to Eleanor of Aquitaine?"

After a brief discussion with the show's host, Chris Tarrant, she settled on Henry II. Then Tarrant asked her the killer question, the moment when contestants often agonize the most: "Final answer?" Without missing a beat, Keppel confirmed. The audience broke into cheers as it was revealed she had won.

Keppel did not waver on that November day in 2000 because of her metacognition. The term, coined by psychologist John Flavell in the 1970s, refers to our ability to evaluate our own thinking. Is the answer to a password prompt correct? Is a particular memory from childhood accurate? Will we find it easy or difficult to learn a new language? Metacognition is an internal tribunal that rules on the soundness of our mental representations, such as a memory or judgment. Keppel's metacognition gave her answer a resounding endorsement.

This knack for reflecting on our thoughts is often viewed as a hallmark of the human mind. It is also a vital survival skill. Metacognition is how we identify our limitations and compensate for them. A student who thinks she is unprepared for a chemistry exam, for example, can devote an extra evening to brushing up on atomic orbitals. When you set an alarm to remind yourself of something you suspect you will forget or make a to-do list to keep track of the day's activities, metacognition has stepped in to save you from your own deficiencies.

Metacognition is not only for spotting weakness. It also kicks in when you assess your strengths, such as when a new swimmer kicks off his floaties or a budding cyclist removes the training wheels. A person with accurate metacognition can move on to the next challenge as soon as she is ready, wasting no time in her journey to mastery.

Ultimately metacognition serves as a foundation

for learning and success. When it is impaired, however, performance in school or at work may suffer. You become less able to recognize a bad decision and correct course. Several psychiatric disorders include deficits in metacognition, which can prevent individuals from identifying their own problems. But with new techniques to quantify metacognition in the laboratory and relate it to brain function, researchers are beginning to understand how metacognition works and why it might go awry. Centuries after Socrates counseled the average Athenian to "know thyself," psychologists are discovering the tools to do a better job of it—to train metacognition and to improve our judgments of our own abilities.

## The Metacognitive Mind

Reflecting on our own thoughts is as old as human civilization. The scientific study of it gained its first big boost from Sigmund Freud and his notion that a person's self-knowledge can be inaccurate, with a large portion of the human mind inaccessible to consciousness. He believed that with enough excavation, we could unearth the hidden forces guiding our actions, thus bringing our true beliefs into the light of consciousness. Yet psychologists soon realized that such analysis was unreliable, and pure introspection as a method for gaining insight into our own minds was gradually discarded.

Flavell, a longtime observer of child development, proposed that one aspect of introspection—metacognition—was key to educational success. In a test of memory, for example, he found that "older subjects studied for a while, said they were ready, and usually were," whereas "younger children studied for a while, said they were ready, and usually were not."

His observation hinted that as the brain matures, certain areas or networks might need to firm up in young minds for children to become better judges of their own learning. To study this idea in

### FAST FACTS

#### THINKING ABOUT THINKING

- 1 Metacognition is the ability to make judgments about our own thoughts—for example, assessing whether a memory is accurate or a decision is appropriate.
- 2 People vary in the accuracy of their metacognition. Certain psychological disorders, including dementia and schizophrenia, can impair this ability.
- 3 Several strategies appear to shore up metacognition, including meditation and taking breaks while studying to reflect on one's own learning.



# The knack for reflecting on our own thoughts is often viewed as a hallmark of the human mind. It is also a vital survival skill.

the lab, however, researchers have had to grapple with a conundrum: how to test people's thoughts about their thoughts.

Because there are no obvious markers of metacognition, my colleagues and I use a shortcut. We measure individuals' confidence in a judgment and see whether their certitude was warranted. Examples of misplaced confidence crop up in daily life all the time. When an inexperienced cook decides it is a good idea to try out new recipes on a dozen friends but then burns the salmon, undercooks the paella and forgets to dress the salad, he or she might be demonstrating poor metacognition.

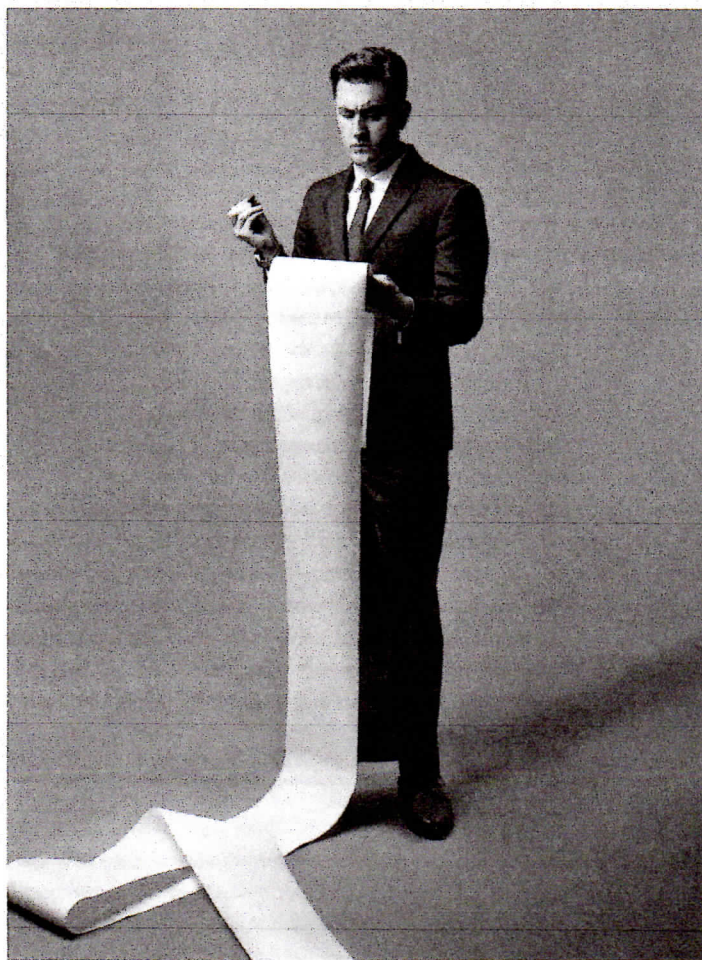
In my studies, the task is much simpler than pulling off a four-course meal. Participants sit in front of a computer screen and see two big circles flash briefly. The circles are filled with dots, and the goal is to decide which patch contains more dots. Most people find it pretty challenging. Getting the answer right is not what interests me here—I want to know how sure people are of their selections. Trial after trial, subjects pick circles and rate their faith in their answers, and eventually a pattern emerges. If your confidence is high only when you do well, and vice versa, your metacognition is in fine form. Similar test setups can quantify metacognition related to other aspects of behavior, such as learning and memory.

Using such experiments, my colleagues and I have found that metacognitive accuracy varies widely across the population. Some people have very poor insight into their own thinking, whereas others appear capable of excellent mental self-assessment. Yet it is important to note that a person's metacognitive prowess does not predict performance. You can have little concept of your own skill level but do a marvelous job of counting dots (or throwing dinner parties).

## The Anatomy of Insight

Using the tools of contemporary neuroscience, scientists are now beginning to identify the brain mechanisms that govern metacognition. The first

clues came from patients with a peculiar form of brain damage. In the mid-1980s neuroscientist Art Shimamura was a postdoctoral student working with Larry R. Squire of the University of California, San Diego. They were studying patients with amnesia, all of whom had damage to the hippocampus, a critical memory region, when they noticed an odd pattern in their data. Most of their patients had poor memory, as expected, but only some of them were aware of their problems with recall. The am-







nesic patients who were unaware of their deficits—they had poor metacognition—suffered from Korsakoff's syndrome, a disorder often associated with alcoholism. Patients with this condition not only become amnesic from an injured hippocampus, they also endure damage to the frontal lobe of the brain. That insight led Shimamura and Squire to suspect that metacognition is governed by the brain's frontal lobe.

To confirm their hunch, they needed to find patients with damaged frontal lobes but intact memory regions. Working with their colleague Jeri Janowsky, they found seven individuals with lesions to the frontal lobe and observed that their metacognition was indeed impaired: when the scientists showed them a list of sentences and asked them how likely

they were to recognize those sentences later, the subjects made inaccurate predictions. Their memory for the sentences remained intact, however. These studies were the first to show that metacognition is an independent function of the brain and not simply part and parcel of everyday abilities.

The frontal lobe of the brain covers a vast swath of neural real estate, and my colleagues and I wanted to pinpoint the hubs of metacognition more precisely. Research we published in 2010 set out to do just that. In a study carried out with Rimona S. Weil, Geraint Rees and other colleagues at University College London, we briefly showed volunteers two images and asked them which one looked brighter. Then they reported how confident they were in their answer. After a series of trials, we calculated a metacognition score for every subject.

To rule out any differences in visual perception, we made sure our subjects were equally capable of identifying the brighter patch, which they did approximately 70 percent of the time. After we scored them, we scanned their brains and found that the people with better metacognition had more gray matter in the anterior prefrontal cortex (aPFC), a brain region toward the front of the frontal lobe that is disproportionately enlarged in humans as compared with other primates. Gray matter consists mostly of neuron cell bodies, as opposed to the white matter of slender axons, which extend from the cell body and transmit electrical impulses to other neurons. More metacognitive individuals also had denser white matter tracts connecting the aPFC to the rest of the brain.

Other brain-imaging studies have suggested that neural activity in the aPFC is more tightly correlated with confidence in people with better metacognition. In addition, zapping that general area with magnetic pulses, which temporarily interferes with the activity of neurons, has been found to impair subjects' metacognition without affecting other aspects of their perception or decision making.

Many of these studies quantify metacognition in highly artificial scenarios, and naturally my colleagues and I were curious whether the principles and brain regions we were identifying for simple judgments also play a role in more complex decisions. Together with neuroscientists Benedetto De Martino, Ray Dolan and Neil Garrett, all then at University College London, we designed a more life-like experiment, albeit inside a brain scanner. We asked participants to decide which of two snacks they prefer, for example, Pringles versus KitKats. Then they told us how confident they felt that they had picked the superior food. After they climbed

#### THE AUTHOR

**STEPHEN M. FLEMING** is a cognitive neuroscientist who studies how the brain analyzes its own computations, or metacognition. He is currently a Sir Henry Wellcome Postdoctoral Fellow, based at both New York University and the University of Oxford.



out of the scanner, they reported how much they would be willing to pay for either snack, and again they rated their confidence, this time in the dollar amounts they had named.

This elaborate procedure helped us to tease apart the brain activity supporting our actions from the neural hubbub governing our thoughts about our actions. As it turned out, not everyone said they would pay more for the item they claimed to prefer—the seemingly logical response. Yet some individuals were more aware than others of their inconsistent behavior. As we reported in 2013, these individuals had stronger connectivity between a region of the brain involved in value computation and the aPFC. Although they did not always make optimal

choices, at least they knew they were floundering.

We still have much to learn. For example, we do not yet know how the aPFC contributes to metacognition or why greater brain volume in this region leads to changes in insight. Yet these findings are a crucial first step toward identifying ways to shore up metacognition, whose absence can produce devastating effects.

### Lack of Insight

When a person with a disorder is unaware of his or her impairments, clinicians use the term “anosognosia,” from Greek roots meaning “without knowledge of disease.” Patients with dementia, for instance, may not notice that their memory is

## Metacognition in Mammals—and Machines

Is metacognition unique to humans? We cannot ask animals for a verbal judgment about their behavior, but ingenious animal-friendly tests can nonetheless probe whether other creatures form thoughts about their own thoughts.

In a pioneering experiment, psychologist David Smith of the University at Buffalo trained a dolphin named Natua to swim



Dolphins can be trained to report their confidence in their own judgments.

toward one of two levers when he heard either a low- or high-pitched sound. When Natua answered correctly, he scored a fish. But some sounds were more difficult for him to distinguish. So Smith introduced a third lever, which triggered an easier trial and let Natua collect his fish. The dolphin learned to press this lever only on more difficult trials.

Smith reckoned that for Natua to choose the third lever, the dolphin must recognize the absence of knowledge and thus must be reflecting on how much he knows. Additional support for this conclusion came from observations that the longer the dolphin hesitated or wavered between the two response options, the more likely he was to choose the third lever. So his opt-out

choices appeared to be based on bona fide feelings of uncertainty. As later experiments demonstrated, macaques also show similar metacognitive behaviors, but another species of monkey, capuchins, do not.

A different metacognition test mimics the confidence judgments we ask of humans in the lab. As in the dolphin experiment, an animal decides which of two answers it thinks is correct. Then it is given a chance to commit to that choice or go for a separate, safe option that always delivers a small snack. Betting on the original selection is riskier—it garners a larger treat if correct but no food otherwise. Macaques pass with flying colors: they take the riskier bet when they are more likely to be correct. The activity of neurons in the monkeys' frontal cortex also tracks their confidence, providing a window on how metacognition is implemented at the level of neural circuits. Even rats can learn to pass a version of this test.

Yet the evidence is not enough to conclude that animals have metacognition. For one thing, the anterior prefrontal cortex, a key brain area for human metacognition, is larger in humans than in monkeys and does not exist in rats. This anatomical difference does not necessarily eliminate the possibility of introspective rodents, because metacognition might have evolved in more than one form. It might manifest both as an implicit feeling of uncertainty that animals share with us, as well as the conscious self-knowledge that might be unique to humans.

Even some computers may embody a form of metacognition. When Watson, IBM's *Jeopardy*-playing machine, beat two champion human players in 2011, it relied on a skill very similar to human self-knowledge. Watson not only came up with answers but also generated a confidence rating for them. The supercomputer then used the rating to decide whether to hit the buzzer. As it turned out, Watson knew that it knew—it calculated a rating—faster than the human *Jeopardy* champs knew that they knew, giving IBM's machine the edge it needed to win. —S.M.F.



slipping away. As a result, they may not seek help, remember to take medications or recognize that they can no longer safely drive a car. Schizophrenia, addiction and stroke can also harm metacognition. A parent's or sibling's lack of insight into his or her own illness can be heartbreaking to a family, causing the shared reality on which social relationships are built to crumble.

Psychiatrists have traditionally believed that such

I'll remember to pay the rent tomorrow, so I better make a note to myself") and metacognition of perception ("Did I really spot an endangered Henslow's sparrow—or just another humdrum song sparrow?"). My collaborators at New York University and I have similarly found that individuals with damage to the aPFC struggle with perceptual metacognition but seem to have no trouble making accurate judgments of their memories. Uncovering the

**B**ecause meditation involves consistent self-focus and the ability to zero in on your own mental states, it might also hone our self-appraisal abilities.

patients were simply in denial. In this view, patients recognized their deficits but were reluctant to admit them to physicians and family members. Now, however, metacognitive failure is seen as a consequence of certain disorders. For instance, most alcoholics do not see their own drinking as problematic—even though they also believe that excessive drinking is unhealthy. As psychiatrist Rita Z. Goldstein and her colleagues wrote in 2009, "one of the greatest challenges in drug addiction treatment is that the individuals who require treatment do not even recognize the need for therapeutic help."

Whether metacognition and anosognosia are two sides of the same coin is not yet clear, although we do know they are closely related. Patients with schizophrenia who lack awareness of their illness, for example, tend to have a smaller frontal lobe than those who recognize their disorder—the same pattern seen in the healthy individuals with impaired metacognition mentioned earlier. (Because psychiatric illness has multiple effects on the brain, dysfunction in a network of brain regions most likely underpins anosognosia.)

We may learn that anosognosia is simply one kind of metacognitive failure. Recent studies hint that a person's capacity for introspection can differ across various domains; perhaps anosognosia is one such category. In support of this view, scientists have documented differences in the brain activity associated with metacognition of memory ("I doubt

neural roots of different kinds of introspective failure will help researchers home in on therapies targeting anosognosia—and potentially help patients manage their ailments or seek treatment.

### A Metacognitive Boost

Efforts to restore metacognition first began in the late 1990s. A small-scale trial looked into the effects of clozapine, an antipsychotic drug, on people with schizophrenia. The study found that patients' insight into clinical symptoms improved after six months of treatment. The medication also lifted their schizophrenic symptoms, so the researchers could not surmise which aspect of their recovery aided metacognition.

More recently, psychologist Robert Hester of the University of Melbourne in Australia and his colleagues found that methylphenidate (Ritalin) could enhance metacognition in healthy volunteers. In these experiments, subjects performed a difficult color-detection task under time pressure and noted whenever they thought they made an error—a metacognitive judgment. The participants given Ritalin—but not subjects given other drugs, such as citalopram, a common antidepressant—could consciously recognize more of their own errors.

Electrical brain stimulation might also serve to enhance metacognition. Using the same task, a team at Trinity College Dublin found that passing a weak electric current through the frontal cortex



of elderly volunteers increased their awareness of their own errors. These weak currents temporarily excite neurons, which may place the frontal lobe in a “primed” state that improves metacognition. But we are still a long way from understanding how drugs or brain stimulation may enhance our ability to reason about our own thoughts.

A more readily available way to improve self-judgments is through meditation. In a 2014 study led by psychologists Benjamin Baird and Jonathan W. Schooler of U.C. Santa Barbara, engaging in two weeks of meditation training increased metacognition during a memory test (but not during a task involving visual discrimination). Because meditation involves consistent self-focus and the ability to zero in on your own mental states, it might also hone our self-appraisal abilities. Indeed, other studies have found that meditation leads to changes in the structure, function and connectivity of the aPFC, raising the tantalizing possibility that such training induces neuroplasticity in brain circuits involved in both meditation and metacognition. But this idea remains speculative: no one has yet documented neural changes that persist after improvements in metacognition.

Simple psychological strategies can shore up metacognition in the classroom. In the early 1990s the late psychologist Thomas O. Nelson and his student John Dunlosky, then at the University of Washington, reported an intriguing effect. When volunteers were asked to reflect on how well they had learned a list of word pairs after a short delay, they were more self-aware than if asked immediately. Many studies have since replicated this finding. Encouraging a student to take a break before deciding how well he or she has studied for an upcoming test could aid learning in a simple but effective way.

Learners could also trigger better insight by coming up with their own subject keywords. Educational psychologist Keith Thiede of Boise State University and his colleagues found that asking students to generate a few words summarizing a particular topic led to greater metacognitive accuracy. The students then allocated their study time better by focusing on material that was less well understood.

Yet we might not always want to increase insight. In some scenarios, it might prove traumatic. A patient with Alzheimer’s disease, for instance, may be troubled by an awareness of his failing memory. This and other ethical questions will need to be grappled with as the field of metacognitive neuroscience matures.

Through the lens of metacognition we experience our thoughts and feelings, but the focus of this lens is finely tuned and fragile. Overly distorted metacog-



nition can lead to failures of self-knowledge and poor decision making. In extreme cases, such as in psychiatric disorders, a person may fail to connect with the shared social reality that others enjoy. Refocusing the lens may be the key to ameliorating some mysterious and devastating aspects of psychiatric illness. Through the marriage of cognitive neuroscience, psychology and computational models, the tools for doing so may soon be within our grasp. **M**

#### FURTHER READING

- **Metacognition.** John Dunlosky and Janet Metcalfe. Sage, 2009.
- **Relating Introspective Accuracy to Individual Differences in Brain Structure.** Stephen M. Fleming et al. in *Science*, Vol. 329, pages 1541–1543; September 17, 2010.
- **The Neural Basis of Metacognitive Ability.** Stephen M. Fleming and Raymond J. Dolan in *Philosophical Transactions of the Royal Society B*, Vol. 367, pages 1338–1349; May 19, 2012.
- **Domain-Specific Enhancement of Metacognitive Ability following Meditation Training.** Benjamin Baird, Michael D. Mrazek, Dawa T. Phillips and Jonathan W. Schooler in *Journal of Experimental Psychology: General*. Published online May 12, 2014.

From Our Archives

- **Being in the Now.** Amishi P. Jha; March/April 2013.
- **What Works, What Doesn't.** John Dunlosky et al.; September/October 2013.