MICHIO KAKU

THE FUTURE OF THE MIND

'Fascinating, fizzes with effervescence' Independent

The Scientific Quest To Understand, Enhance and Empower the Mind

Contents

BY THE SAME AUTHOR

Physics of the Future

Physics of the Impossible

Parallel Worlds

Hyperspace

Visions

Einstein's Cosmos

Beyond Einstein

This book is dedicated to my loving wife, Shizue,
and my daughters, Michelle and Alyson

Introduction

The two greatest mysteries in all of nature are the mind and the universe. With our vast technology, we have been able to photograph galaxies billions of light-years away, manipulate the genes that control life, and probe the inner sanctum of the atom, but the mind and the universe still elude and tantalize us. They are the most mysterious and fascinating frontiers known to science.

If you want to appreciate the majesty of the universe, just turn your gaze to the heavens at night, ablaze with billions of stars. Ever since our ancestors first gasped at the splendor of the starry sky, we have puzzled over these eternal questions: Where did it all come from? What does it all mean?

To witness the mystery of our mind, all we have to do is stare at ourselves in the mirror and wonder, What lurks behind our eyes? This raises haunting questions like: Do we have a soul? What happens to us after we die? Who am "I" anyway? And most important, this brings us to the ultimate question: Where do we fit into this great cosmic scheme? As the great Victorian biologist Thomas Huxley once said, "The question of all questions for humanity, the problem which lies behind all others and is more interesting than any of them, is that of the determination of man's place in Nature and his relation to the Cosmos."

There are 100 billion stars in the Milky Way galaxy, roughly the same as the number of neurons in our brain. You may have to travel twenty-four trillion miles, to the first star outside our solar system, to find an object as complex as what is sitting on your shoulders. The mind and the universe pose the greatest scientific challenge of all, but they also share a curious relationship. On one hand they are polar opposites. One is concerned with the vastness of outer space, where we encounter strange denizens like black holes, exploding stars, and colliding galaxies. The other is concerned with inner space, where we find our most intimate and private hopes and desires. The mind is no farther than our next thought, yet we are often clueless when asked to articulate and explain it.

But although they may be opposites in this respect, they also have a common history and narrative. Both were shrouded in superstition and magic since time immemorial. Astrologers and phrenologists claimed to find the meaning of the universe in every constellation of the zodiac and in every bump on your head. Meanwhile, mind readers and seers have been alternately celebrated and vilified over the years.

The universe and the mind continue to intersect in a variety of ways, thanks in no small part to some of the eye-opening ideas we often encounter in science fiction. Reading these books as a child, I would daydream about being a member of the Slan, a race of telepaths created by A. E. van Vogt. I marveled at how a mutant called the Mule could unleash his vast telepathic powers and nearly seize control of the Galactic Empire in Isaac Asimov's *Foundation Trilogy.* And in the movie *Forbidden Planet,* I wondered how an advanced civilization millions of years beyond ours could channel its enormous telekinetic powers to reshape reality to its whims and wishes.

Then when I was about ten, "The Amazing Dunninger" appeared on TV. He would dazzle his audience with his spectacular magic tricks. His motto was "For those who believe, no explanation is necessary; for those who do not believe, no explanation will suffice." One day, he declared that he would send his thoughts to millions of people throughout the country. He closed his eyes and began to concentrate, stating that he was beaming the name of a president of the United States. He asked people to write down the name that popped into their heads on a postcard and mail it in. The next week, he announced triumphantly that thousands of postcards had come pouring in with the name "Roosevelt," the very same name he was "beaming" across the United States.

I wasn't impressed. Back then, the legacy of Roosevelt was strong among those who had lived through the Depression and World War II, so this came as no surprise. (I thought to myself that it would have been truly amazing if he had been thinking of President Millard Fillmore.)

Still, it stoked my imagination, and I couldn't resist experimenting with telepathy on my own, trying to read other people's minds by concentrating as hard as I could. Closing my eyes and focusing intently, I and a variety of advanced brain scans in the mid-1990s and 2000s has transformed neuroscience. We have learned more about the brain in the last fifteen years than in all prior human history, and the mind, once considered out of reach, is finally assuming center stage.

Nobel laureate Eric R. Kandel of the Max Planck Institute in Tübingen, Germany, writes, "The most valuable insights into the human mind to emerge during this period did not come from the disciplines traditionally concerned with the mind—philosophy, psychology, or psycho-analysis. Instead they came from a merger of these disciplines with the biology of the brain. …"

Physicists have played a pivotal role in this endeavor, providing a flood of new tools with acronyms like MRI, EEG, PET, CAT, TCM, TES, and DBS that have dramatically changed the study of the brain. Suddenly with these machines we could see thoughts moving within the living, thinking brain. As neurologist V. S. Ramachandran of the University of California, San Diego, says, "All of these questions that philosophers have been studying for millennia, we scientists can begin to explore by doing brain imaging and by studying patients and asking the right questions."

Looking back, some of my initial forays into the world of physics intersected with the very technologies that are now opening up the mind for science. In high school, for instance, I became aware of a new form of matter, called antimatter, and decided to conduct a science project on the topic. As it is one of the most exotic substances on Earth, I had to appeal to the old Atomic Energy Commission just to obtain a tiny quantity of sodium-22, a substance that naturally emits a positive electron (anti-electron, or positron). With my small sample in hand, I was able to build a cloud chamber and powerful magnetic field that allowed me to photograph the trails of vapor left by antimatter particles. I didn't know it at the time, but sodium-22 would soon become instrumental in a new technology, called PET (positron emission tomography), which has since given us startling new insights into the thinking brain.

Yet another technology I experimented with in high school was magnetic resonance. I attended a lecture by Felix Bloch of Stanford University, who shared the 1952 Nobel Prize for Physics with Edward Purcell for the discovery of nuclear magnetic resonance. Dr. Bloch explained to us high school kids that if you had a powerful magnetic field, the atoms would align vertically in that field like compass needles. Then if you applied a radio pulse to these atoms at a precise resonant frequency, you could make them flip over. When they eventually flipped back, they would emit another pulse, like an echo, which would allow you to determine the identity of these atoms. (Later, I used the principle of magnetic resonance to build a 2.3-million-electron-volt particle accelerator in my mom’s garage.)

Just a couple of years later, as a freshman at Harvard University, it was an honor to have Dr. Purcell teach me electrodynamics. Around that same time, I also had a summer job and got a chance to work with Dr. Richard Ernst, who was trying to generalize the work of Bloch and Purcell on magnetic resonance. He succeeded spectacularly and would eventually win the Nobel Prize for Physics in 1991 for laying the foundation for the modern MRI (magnetic resonance imaging) machine. The MRI machine, in turn, has given us detailed photographs of the living brain in even finer detail than PET scans.

EMPOWERING THE MIND

Eventually I became a professor of theoretical physics, but my fascination with the mind remained. It is thrilling to see that, just within the last decade, advances in physics have made possible some of the feats of mentalism that excited me when I was a child. Using MRI scans, scientists can now read thoughts circulating in our brains. Scientists can also insert a chip into the brain of a patient who is totally paralyzed and connect it to a computer, so that through thought alone that patient can surf the web, read and write e-mails, play video games, control their wheelchair, operate household appliances, and manipulate mechanical arms. In fact, such patients can do anything a normal person can do via a computer.

Scientists are now going even further, by connecting the brain directly to an exoskeleton that these patients can wear around their paralyzed limbs. Quadriplegics may one day lead near-normal lives. Such exoskeletons may also give us superpowers enabling us to handle deadly emergencies. One day, our astronauts may even explore the planets by mentally controlling mechanical surrogates from the comfort of their living rooms.

As in the movie *The Matrix*, we might one day be able to download memories and skills using computers. In animal studies, scientists have already been able to insert memories into the brain. Perhaps it's only a matter of time before we, too, can insert artificial memories into our brains to learn new subjects, vacation in new places, and master new hobbies. And if technical skills can be downloaded into the minds of workers and scientists, this may even affect the world economy. We might even be able to share these memories as well. One day, scientists might construct an "Internet of the mind," or a brain-net, where thoughts and emotions are sent electronically around the world. Even dreams will be videotaped and then "brain-mailed" across the Internet.

Technology may also give us the power to enhance our intelligence. Progress has been made in understanding the extraordinary powers of "savants" whose mental, artistic, and mathematical abilities are truly astonishing. Furthermore, the genes that separate us from the apes are now being sequenced, giving us an unparalleled glimpse into the evolutionary origins of the brain. Genes have already been isolated in animals that can increase their memory and mental performance.

The excitement and promise generated by these eye-opening advances are so enormous that they have also caught the attention of the politicians. In fact, brain science has suddenly become the source of a transatlantic rivalry between the greatest economic powers on the planet. In January 2013, both President Barack Obama and the European Union announced what could eventually become multibillion-dollar funding for two independent projects that would reverse engineer the brain. Deciphering the intricate neural circuitry of the brain, once considered hopelessly beyond the scope of modern science, is now the focus of two crash projects that, like the Human Genome Project, will change the scientific and medical landscape. Not only will this give us unparalleled insight into the mind, it will also generate new industries, spur economic activity, and open up new vistas for neuroscience.

Once the neural pathways of the brain are finally decoded, one can envision understanding the precise origins of mental illness, perhaps leading to a cure for this ancient affliction. This decoding also makes it possible to create a copy of the brain, which raises philosophical and ethical questions. Who are we, if our consciousness can be uploaded into a computer? We can also toy with the concept of immortality. Our bodies may eventually decay and die, but can our consciousness live forever?

And beyond that, perhaps one day in the distant future the mind will be freed of its bodily constraints and roam among the stars, as several scientists have speculated. Centuries from now, one can imagine placing our entire neural blueprint on laser beams, which will then be sent into deep space, perhaps the most convenient way for our consciousness to explore the stars.

A brilliant new scientific landscape that will reshape human destiny is now truly opening up. We are now entering a new golden age of neuroscience.

In making these predictions, I have had the invaluable assistance of scientists who graciously allowed me to interview them, broadcast their ideas on national radio, and even take a TV crew into their laboratories. These are the scientists who are laying the foundation for the future of the mind. For their ideas to be incorporated into this book, I made only two requirements: (1) their predictions must rigorously obey the laws of physics; and (2) prototypes must exist to show proof-of-principle for these far-reaching ideas.

TOUCHED BY MENTAL ILLNESS

I once wrote a biography of Albert Einstein, called *Einstein's Cosmos,* and had to delve into the minute details of his private life. I had known that Einstein's youngest son was afflicted with schizophrenia, but did not realize the enormous emotional toll that it had taken on the great scientist's life. Einstein was also touched by mental illness in another way; one of his closest colleagues was the physicist Paul Ehrenfest, who helped Einstein create the theory of general relativity. After suffering bouts of depression, Ehrenfest tragically killed his own son, who had Down's syndrome, and then committed suicide. Over the years, I have found that many of my colleagues and friends have struggled to manage mental illness in their families.

Mental illness has also deeply touched my own life. Several years ago, my mother died after a long battle with Alzheimer's disease. It was heartbreaking to see her gradually lose her memories of her loved ones, to gaze into her eyes and realize that she did not know who I was. I could see the glimmer of humanity slowly being extinguished. She had spent a lifetime struggling to raise a family, and instead of enjoying her golden years, she was robbed of all the memories she held dear.

As the baby boomers age, the sad experience that I and many others have had will be repeated across the world. My wish is that rapid advances in neuroscience will one day alleviate the suffering felt by those afflicted with mental illness and dementia.

WHAT IS DRIVING THIS REVOLUTION?

The data pouring in from brain scans are now being decoded, and the progress is stunning. Several times a year, headlines herald a fresh breakthrough. It took 350 years, since the invention of the telescope, to enter the space age, but it has taken only fifteen years since the introduction of the MRI and advanced brain scans to actively connect the brain to the outside world. *Why so quickly, and how much is there to come?*

Part of this rapid progress has occurred because physicists today have a good understanding of electromagnetism, which governs the electrical signals racing through our neurons. The mathematical equations of James Clerk Maxwell, which are used to calculate the physics of antennas, radar, radio receivers, and microwave towers, form the very cornerstone of MRI technology. It took centuries to finally solve the secret of electromagnetism, but neuroscience can enjoy the fruits of this grand endeavor. In Book I, I will survey the history of the brain and explain how a galaxy of new instruments has left the physics labs and given us glorious color pictures of the mechanics of thought. Because consciousness plays so central a role in any discussion of the mind, I also give a physicist's perspective, offering a definition of consciousness that includes the animal kingdom as well. In fact, I provide a ranking of consciousness, showing how it is possible to assign a number to various types of consciousness.

But to fully answer the question of how this technology will advance, we also have to look at Moore's law, which states that computer power doubles every two years. I often surprise people with the simple fact that your cell phone today has more computer power than *all* of NASA when it put two men on the moon in 1969. Computers are now powerful enough to record the electrical signals emanating from the brain and partially decode them into a familiar digital language. This makes it possible for the brain to directly interface with computers to control any object around it. The fast-growing field is called BMI (brain-machine

Book I

THE MIND AND CONSCIOUSNESS

My fundamental premise about the brain is that its workings—what we sometimes call "mind"—are a consequence of its anatomy and physiology, and nothing more.
—CARL SAGAN

1 UNLOCKING THE MIND

In 1848, Phineas Gage was working as a railroad foreman in Vermont, when dynamite accidentally went off, propelling a three-foot, seven-inch spike straight into his face, through the front part of his brain, and out the top of his skull, eventually landing eighty feet away. His fellow workers, shocked to see part of their foreman's brain blown off, immediately called for a doctor. To the workers' (and even the doctor's) amazement, Mr. Gage did not die on-site.

He was semiconscious for weeks, but eventually made what seemed like a full recovery. (A rare photograph of Gage surfaced in 2009, showing a handsome, confident man, with an injury to his head and left eye, holding the iron rod.) But after this incident, his coworkers began to notice a sharp change in his personality. A normally cheerful, helpful foreman, Gage became abusive, hostile, and selfish. Ladies were warned to stay clear of him. Dr. John Harlow, the doctor who treated him, observed that Gage was "capricious and vacillating, devising many plans of future operations, which are no sooner arranged than they are abandoned in turn for others appearing more feasible. A child in his intellectual capacity and manifestations, yet with the animal passions of a strong man." Dr. Harlow noted that he was "radically changed" and that his fellow workers said that "he was no longer Gage." After Gage's death in 1860, Dr. Harlow preserved both his skull and the rod that had smashed into it. Detailed X-ray scans of the skull have since confirmed that the iron rod caused massive destruction in the area of the brain behind the forehead known as the frontal lobe, in both the left and right cerebral hemispheres.

This incredible accident would not only change the life of Phineas Gage, it would alter the course of science as well. Previously, the dominant thinking was that the brain and the soul were two separate entities, a philosophy called dualism. But it became increasingly clear that damage to the frontal lobe of his brain had caused abrupt changes in Gage's personality. This, in turn, created a paradigm shift in scientific thinking: perhaps specific areas of the brain could be traced to certain behaviors.

BROCA'S BRAIN

In 1861, just a year after Gage's death, this view was further cemented through the work of Pierre Paul Broca, a physician in Paris who documented a patient who appeared normal except that he had a severe speech deficit. The patient could understand and comprehend speech perfectly, but he could utter only one sound, the word "tan." After the patient died, Dr. Broca confirmed during the autopsy that the patient suffered from a lesion in his left temporal lobe, a region of the brain near his left ear. Dr. Broca would later confirm twelve similar cases of patients with damage to this specific area of the brain. Today patients who have damage to the temporal lobe, usually in the left hemisphere, are said to suffer from Broca's aphasia. (In general, patients with this disorder can understand speech but cannot say anything, or else they drop many words when speaking.)

Soon afterward, in 1874, German physician Carl Wernicke described patients who suffered from the opposite problem. They could articulate clearly, but they could not understand written or spoken speech. Often these patients could speak fluently with correct grammar and syntax, but with nonsensical words and meaningless jargon. Sadly, these patients often didn't know they were spouting gibberish. Wernicke confirmed after performing autopsies that these patients had suffered damage to a slightly different area of the left temporal lobe.

The works of Broca and Wernicke were landmark studies in neuroscience, establishing a clear link between behavioral problems, such as speech and language impairment, and damage to specific regions of the brain.

Another breakthrough took place amid the chaos of war. Throughout history, there were many religious taboos prohibiting the dissection of the human body, which severely restricted progress in medicine. In warfare, however, with tens of thousands of bleeding soldiers dying on the battlefield, it became an urgent mission for doctors to develop any medical treatment that worked. During the Prusso-Danish War in 1864, German doctor Gustav Fritsch treated many soldiers with gaping wounds to the brain and happened to notice that when he touched one hemisphere of the brain, the opposite side of the body often twitched. Later Fritsch systematically showed that, when he electrically stimulated the brain, the left hemisphere controlled the right side of the body, and vice versa. This was a stunning discovery, demonstrating that the brain was basically electrical in nature and that a particular region of the brain controlled a part on the other side of the body. (Curiously, the use of electrical probes on the brain was first recorded a couple of thousand years earlier by the Romans. In the year A.D. 43, records show that the court doctor to the emperor Claudius used electrically charged torpedo fish, which were applied to the head of a patient suffering from severe headaches.)

The realization that there were electrical pathways connecting the brain to the body wasn't systematically analyzed until the 1930s, when Dr. Wilder Penfield began working with epilepsy patients, who often suffered from debilitating convulsions and seizures that were potentially life-threatening. For them, the last option was to have brain surgery, which involved removing parts of the skull and exposing the brain. (Since the brain has no pain sensors, a person can be conscious during this entire procedure, so Dr. Penfield used only a local anesthetic during the operation.)

Dr. Penfield noticed that when he stimulated certain parts of the cortex with an electrode, different parts of the body would respond. He suddenly realized that he could draw a rough one-to-one correspondence between specific regions of the cortex and the human body. His drawings were so accurate that they are still used today in almost unaltered form. They had an immediate impact on both the scientific community and the general public. In one diagram, you could see which region of the brain roughly controlled which function, and how important each function was. For example, because our hands and mouth are so vital for survival, a considerable amount of brain power is devoted to controlling them, while the sensors in our back hardly register at all.

Furthermore, Penfield found that by stimulating parts of the temporal lobe, his patients suddenly relived long-forgotten memories in a crystal-clear fashion. He was shocked when a patient, in the middle of brain surgery, suddenly blurted out, "It was like … standing in the doorway at

[my] high school. … I heard my mother talking on the phone, telling my aunt to come over that night.” Penfield realized that he was tapping into memories buried deep inside the brain. When he published his results in 1951, they created another transformation in our understanding of the brain.

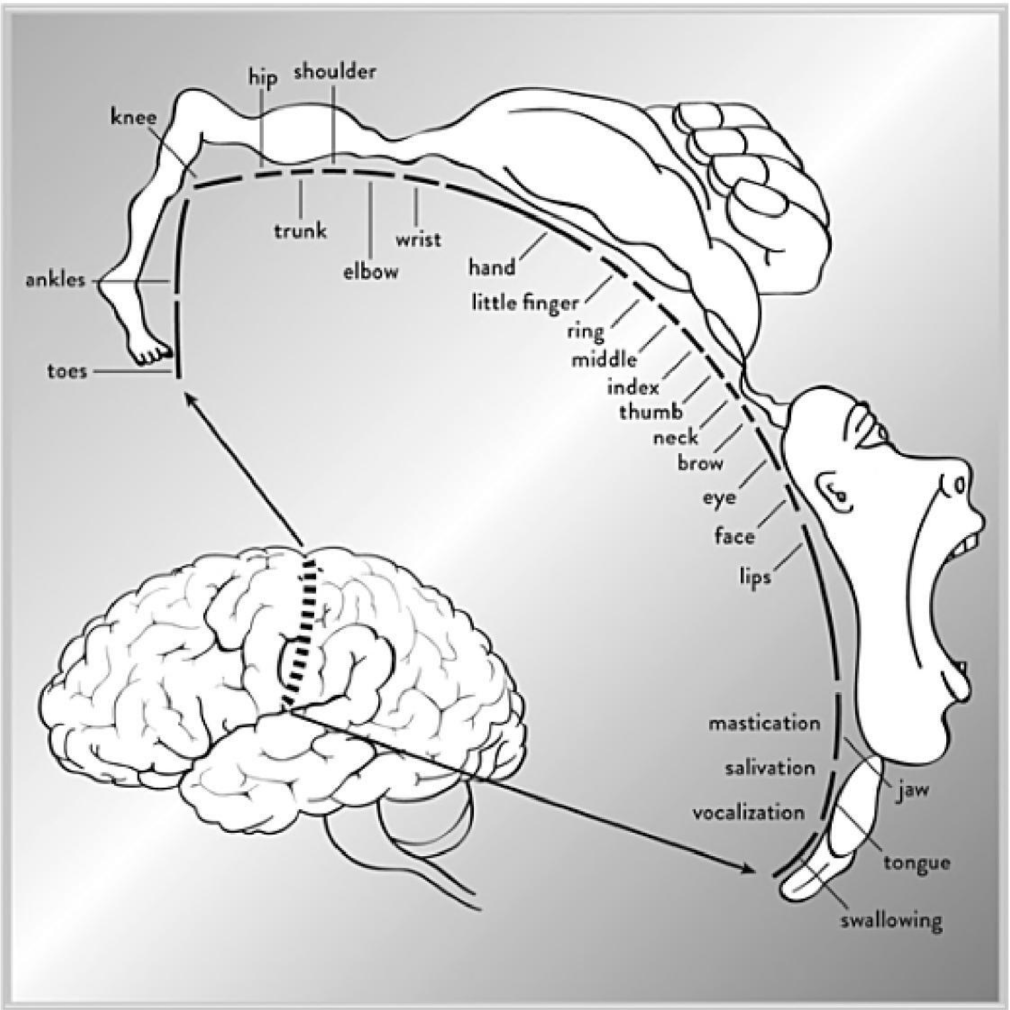


Figure 1. This is the map of the motor cortex that was created by Dr. Wilder Penfield, showing which region of the brain controls which part of the body.

A MAP OF THE BRAIN

By the 1950s and ’60s, it was possible to create a crude map of the brain, locating different regions and even identifying the functions of a few of

The temporal lobe controls language (on the left side only), as well as the visual recognition of faces and certain emotional feelings. Damage to this lobe can leave us speechless or without the ability to recognize familiar faces.

THE EVOLVING BRAIN

When you look at other organs of the body, such as our muscles, bones, and lungs, there seems to be an obvious rhyme and reason to them that we can immediately see. But the structure of the brain might seem slapped together in a rather chaotic fashion. In fact, trying to map the brain has often been called "cartography for fools."

To make sense of the seemingly random structure of the brain, in 1967 Dr. Paul MacLean of the National Institute of Mental Health applied Charles Darwin's theory of evolution to the brain. He divided the brain into three parts. (Since then, studies have shown that there are refinements to this model, but we will use it as a rough organizing principle to explain the overall structure of the brain.) First, he noticed that the back and center part of our brains, containing the brain stem, cerebellum, and basal ganglia, are almost identical to the brains of reptiles. Known as the "reptilian brain," these are the oldest structures of the brain, governing basic animal functions such as balance, breathing, digestion, heartbeat, and blood pressure. They also control behaviors such as fighting, hunting, mating, and territoriality, which are necessary for survival and reproduction. The reptilian brain can be traced back about 500 million years. (See Figure 3.)

But as we evolved from reptiles to mammals, the brain also became more complex, evolving outward and creating entirely new structures. Here we encounter the "mammalian brain," or the limbic system, which is located near the center of the brain, surrounding parts of the reptilian brain. The limbic system is prominent among animals living in social groups, such as the apes. It also contains structures that are involved in emotions. Since the dynamics of social groups can be quite complex, the limbic system is essential in sorting out potential enemies, allies, and rivals.

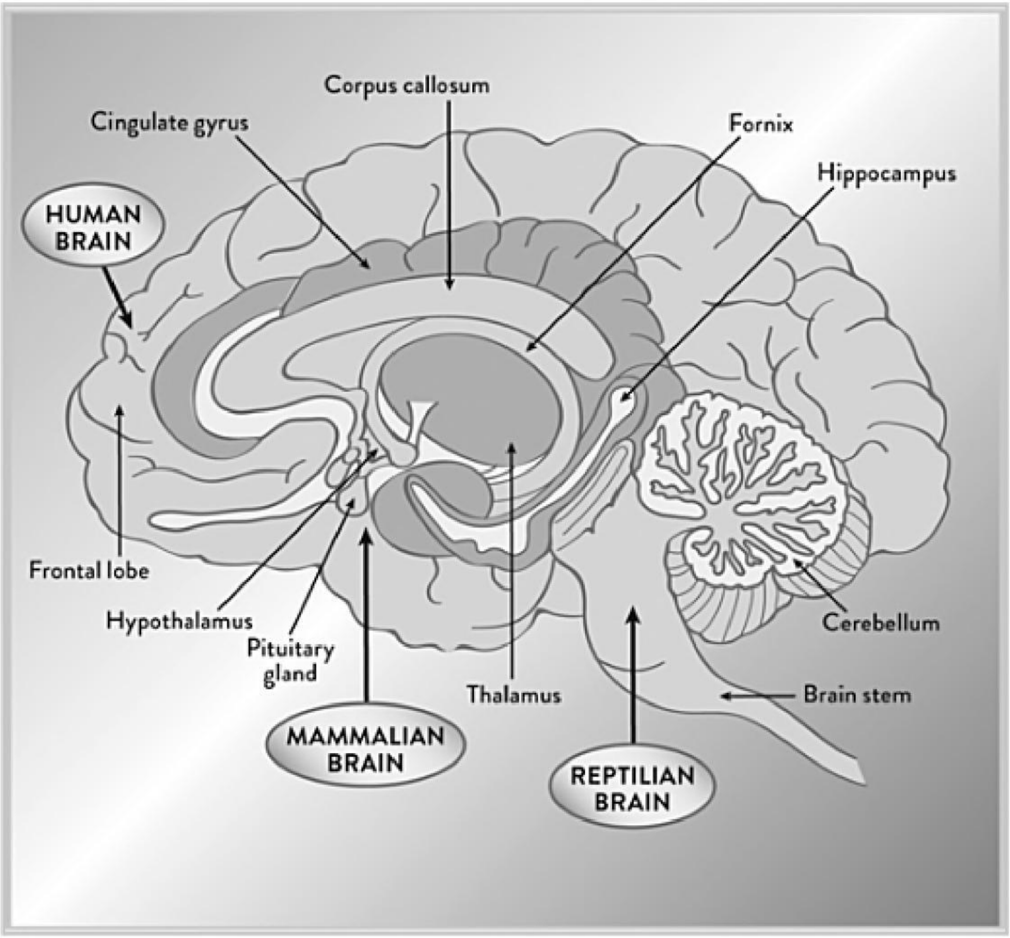


Figure 3. The evolutionary history of the brain, with the reptilian brain, the limbic system (the mammalian brain), and the neocortex (the human brain). Roughly speaking, one can argue that the path of our brain's evolution passed from the reptilian brain to the mammalian brain to the human brain.

The different parts of the limbic system that control behaviors crucial for social animals are:

- The hippocampus. This is the gateway to memory, where short-term memories are processed into long-term memories. Its name means "seahorse," which describes its strange shape. Damage here will destroy the ability to make new long-term memories. You are left a prisoner of the present.
- The amygdala. This is the seat of emotions, especially fear, where emotions are first registered and generated. Its name means "almond."
- The thalamus. This is like a relay station, gathering sensory signals from the

brain stem and then sending them out to the various cortices. Its name means "inner chamber."

- The hypothalamus. This regulates body temperature, our circadian rhythm, hunger, thirst, and aspects of reproduction and pleasure. It lies below the thalamus—hence its name.

Finally, we have the third and most recent region of the mammalian brain, the cerebral cortex, which is the outer layer of the brain. The latest evolutionary structure within the cerebral cortex is the neocortex (meaning "new bark"), which governs higher cognitive behavior. It is most highly developed in humans: it makes up 80 percent of our brain's mass, yet is only as thick as a napkin. In rats the neocortex is smooth, but it is highly convoluted in humans, which allows a large amount of surface area to be crammed into the human skull.

In some sense, the human brain is like a museum containing remnants of all the previous stages in our evolution over millions of years, exploding outward and forward in size and function. (This is also roughly the path taken when an infant is born. The infant brain expands outward and toward the front, perhaps mimicking the stages of our evolution.)

Although the neocortex seems unassuming, looks are deceiving. Under a microscope you can appreciate the intricate architecture of the brain. The gray matter of the brain consists of billions of tiny brain cells called neurons. Like a gigantic telephone network, they receive messages from other neurons via dendrites, which are like tendrils sprouting from one end of the neuron. At the other end of the neuron, there is a long fiber called the axon. Eventually the axon connects to as many as ten thousand other neurons via their dendrites. At the juncture between the two, there is a tiny gap called the synapse. These synapses act like gates, regulating the flow of information within the brain. Special chemicals called neurotransmitters can enter the synapse and alter the flow of signals. Because neurotransmitters like dopamine, serotonin, and noradrenaline help control the stream of information moving across the myriad pathways of the brain, they exert a powerful effect on our moods, emotions, thoughts, and state of mind. (See Figure 4.)

This description of the brain roughly represented the state of knowledge through the 1980s. In the 1990s, however, with the introduction of new technologies from the field of physics, the mechanics of thought began to be revealed in exquisite detail, unleashing the current explosion of scientific discovery. One of the workhorses of this revolution has been the MRI machine.

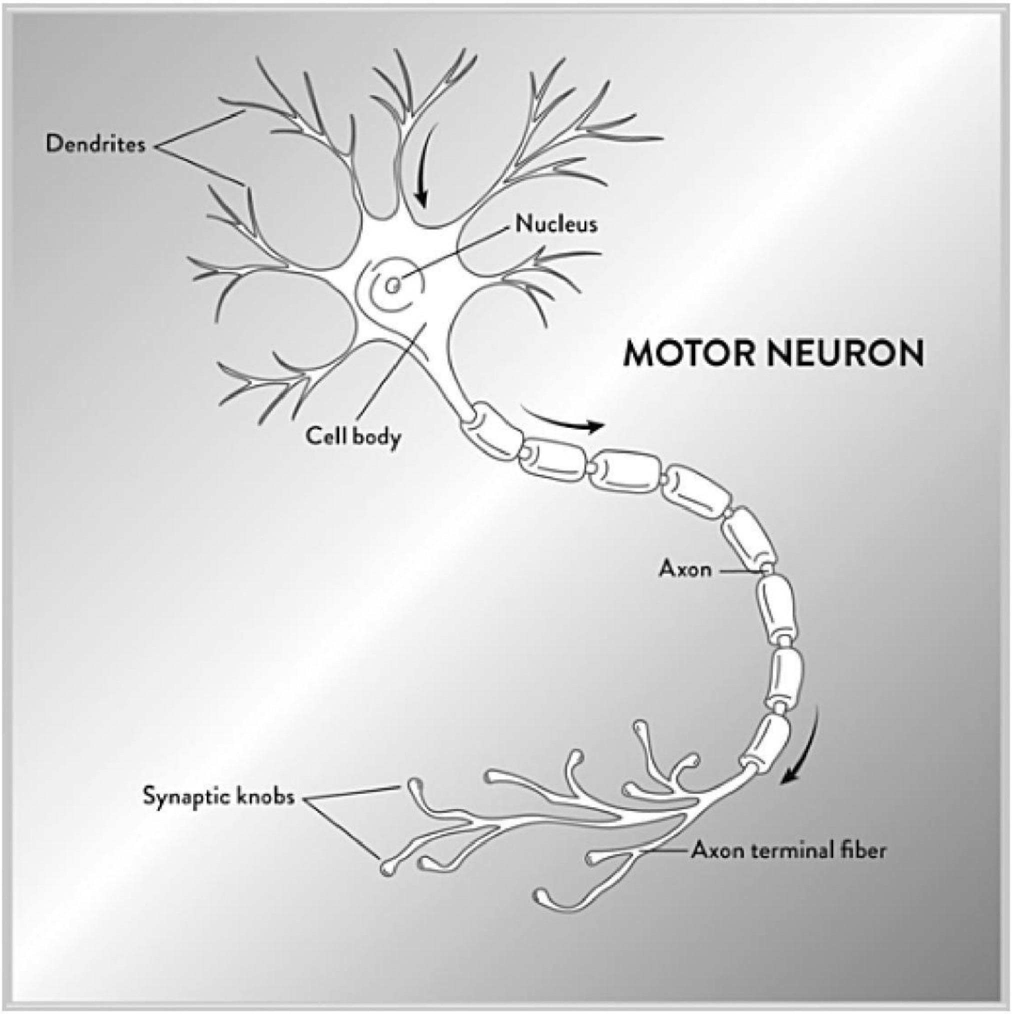


Figure 4. Diagram of a neuron. Electrical signals travel along the axon of the neuron until they hit the synapse. Neurotransmitters can regulate the flow of electrical signals past the synapse.

THE MRI: WINDOW INTO THE BRAIN

To understand the reason why this radical new technology has helped decode the thinking brain, we have to turn our attention to some basic principles of physics.

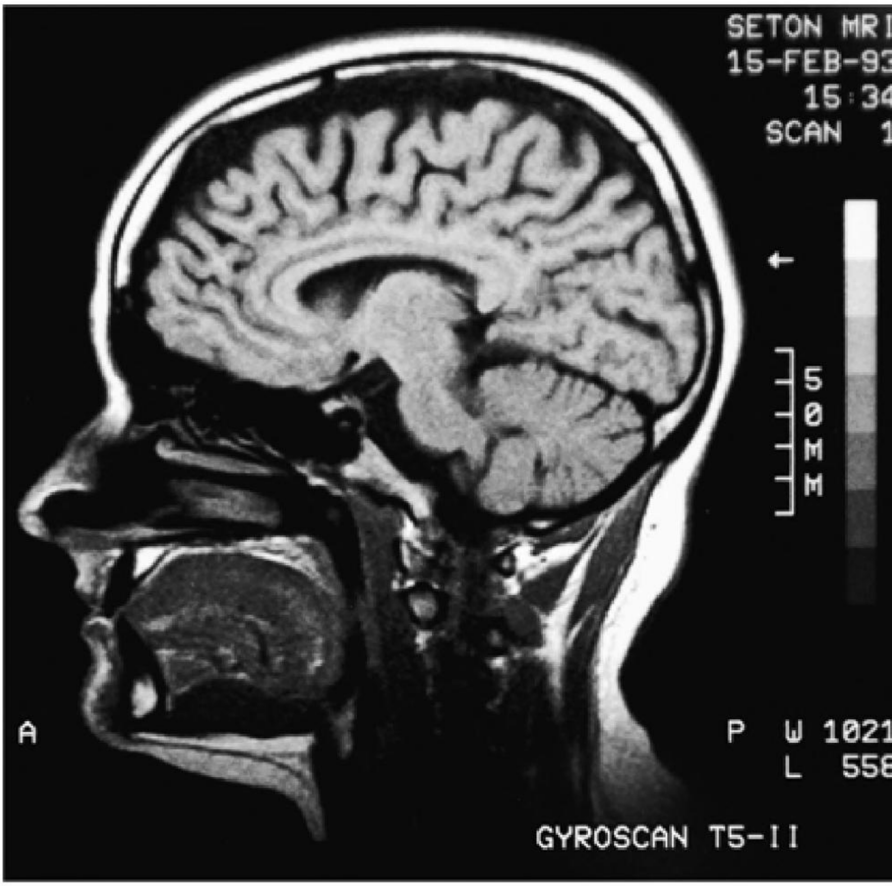
Radio waves, a type of electromagnetic radiation, can pass right through tissue without doing damage. MRI machines take advantage of this fact, allowing electromagnetic waves to freely penetrate the skull. In the process, this technology has given us glorious photographs of something once thought to be impossible to capture: the inner workings of the brain as it experiences sensations and emotions. Watching the dance of lights flickering in a MRI machine, one can trace out the thoughts moving within the brain. It's like being able to see the inside of a clock as it ticks.

The first thing you notice about an MRI machine is the huge, cylindrical magnetic coils, which can produce a magnetic field twenty to sixty thousand times greater than the strength of Earth's. The giant magnet is one of the principal reasons why an MRI machine can weigh a ton, fill up an entire room, and cost several million dollars. (MRI machines are safer than X-ray machines because they don't create harmful ions. CT scans, which can also create 3-D pictures, flood the body with many times the dosage from an ordinary X-ray, and hence have to be carefully regulated. By contrast, MRI machines are safe when used properly. One problem, however, is the carelessness of workers. The magnetic field is powerful enough to send tools hurling through the air at high velocity when turned on at the wrong time. People have been injured and even killed in this way.)

MRI machines work as follows: Patients lie flat and are inserted into a cylinder containing two large coils, which create the magnetic field. When the magnetic field is turned on, the nuclei of the atoms inside your body act very much like a compass needle: they align horizontally along the direction of the field. Then a small pulse of radio energy is generated, which causes some of the nuclei in our body to flip upside down. When the nuclei later revert back to their normal position, they emit a secondary pulse of radio energy, which is then analyzed by the MRI machine. By analyzing these tiny "echoes," one can then reconstruct the location and nature of these atoms. Like a bat, which uses echoes to determine the position of objects in its path, the echoes created by the MRI machine allow scientists to re-create a remarkable image of the inside of the brain. Computers then reconstruct the position of the atoms, giving us beautiful diagrams in three dimensions.

1924, but only recently has it been possible to employ computers to make sense out of all the data pouring in from each electrode.

To use the EEG machine, the patient usually puts on a futuristic-looking helmet with scores of electrodes on the surface. (More advanced versions place a hairnet over the head containing a series of tiny electrodes.) These electrodes detect the tiny electrical signals that are circulating in the brain.



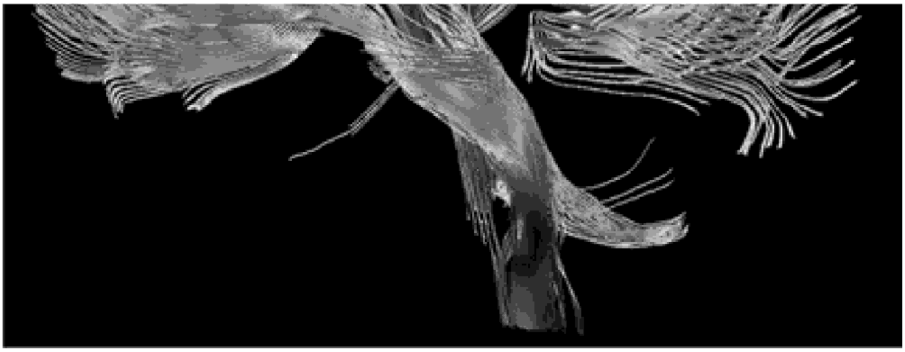


Figure 5. At the top, we see an image taken by a functional MRI machine, showing regions of high mental activity. In the bottom image, we see the flowerlike pattern created by a diffusion MRI machine, which can follow the neural pathways and connections of the brain.

An EEG scan differs from an MRI scan in several crucial ways. The MRI scan, as we have seen, shoots radio pulses into the brain and then analyzes the "echoes" that come back. This means you can vary the radio pulse to select different atoms for analysis, making it quite versatile. The EEG machine, however, is strictly passive; that is, it analyzes the tiny electromagnetic signals the brain naturally emits. The EEG excels at recording the broad electromagnetic signals that surge across the entire brain, which allows scientists to measure the overall activity of the brain as it sleeps, concentrates, relaxes, dreams, etc. Different states of consciousness vibrate at different frequencies. For example, deep sleep corresponds to delta waves, which vibrate at .1 to 4 cycles per second. Active mental states, such as problem solving, correspond to beta waves, vibrating from 12 to 30 cycles per second. These vibrations allow various parts of the brain to share information and communicate with one another, even if they are located on opposite sides of the brain. And while MRI scans measuring blood flow can be taken only several times a second, EEG scans measure electrical activity instantly.

The greatest advantage of the EEG scan, though, is its convenience and cost. Even high school students have done experiments in their living rooms with EEG sensors placed over their heads.

However, the main drawback to the EEG, which has held up its development for decades, is its very poor spatial resolution. The EEG picks up electrical signals that have already been diffused after passing through the skull, making it difficult to detect abnormal activity when it originates deep in the brain. Looking at the output of the muddled EEG signals, it is almost impossible to say for sure which part of the brain created it. Furthermore, slight motions, like moving a finger, can distort the signal, sometimes rendering it useless.

PET SCANS

Yet another useful tool from the world of physics is the positron emission topography (PET) scan, which calculates the flow of energy in the brain by locating the presence of glucose, the sugar molecule that fuels cells. Like the cloud chamber I made as a high school student, PET scans make use of the subatomic particles emitted from sodium-22 within the glucose. To start the PET scan, a special solution containing slightly radioactive sugar is injected into the patient. The sodium atoms inside the sugar molecules have been replaced by radioactive sodium-22 atoms. Every time a sodium atom decays, it emits a positive electron, or positron, which is easily detected by sensors. By following the path of the radioactive sodium atoms in sugar, one can then trace out the energy flow within the living brain.

The PET scan shares many of the same advantages of MRI scans but does not have the fine spatial resolution of an MRI photo. However, instead of measuring blood flow, which is only an indirect indicator of energy consumption in the body, PET scans measure energy consumption, so it is more closely related to neural activity.

There is another drawback to PET scans, however. Unlike MRI and EEG scans, PET scans are slightly radioactive, so patients cannot continually take them. In general, a person is not allowed to have a PET scan more than once a year because of the risk from radiation.

MAGNETISM IN THE BRAIN

Within the last decade, many new high-tech devices have entered the tool kit of neuroscientists, including the transcranial electromagnetic scanner (TES), magnetoencephalography (MEG), near-infrared spectroscopy (NIRS), and optogenetics, among others.

In particular, magnetism has been used to systematically shut down specific parts of the brain without cutting it open. The basic physics behind these new tools is that a rapidly changing electric field can create a magnetic field, and vice versa. MEGs passively measure the magnetic fields produced by the changing electric fields of the brain. These magnetic fields are weak and extremely tiny, only a billionth of Earth’s magnetic field. Like the EEG, the MEG is extremely good at time resolution, down to a thousandth of a second. Its spatial resolution, however, is only a cubic centimeter.

Unlike the passive measurement of the MEG, the TES generates a large pulse of electricity, which in turn creates a burst of magnetic energy. The TES is placed next to the brain, so the magnetic pulse penetrates the skull and creates yet another electric pulse inside the brain. This secondary electrical pulse, in turn, is sufficient to turn off or dampen the activity of selected areas of the brain.

Historically, scientists had to rely on strokes or tumors to silence certain parts of the brain and hence determine what they do. But with the TES, one can harmlessly turn off or dampen parts of the brain at will. By shooting magnetic energy at a particular spot in the brain, one can determine its function by simply watching how a person’s behavior has changed. (For example, by shooting magnetic pulses into the left temporal lobe, one can see that this adversely affects our ability to talk.)

One potential drawback of the TES is that these magnetic fields do not penetrate very far into the interior of the brain (because magnetic fields decrease much faster than the usual inverse square law for electricity). TES is quite useful in turning off parts of the brain near the skull, but the magnetic field cannot reach important centers located deep in the brain, such as the limbic system. But future generations of TES devices may overcome this technical problem by increasing the intensity and precision of the magnetic field.

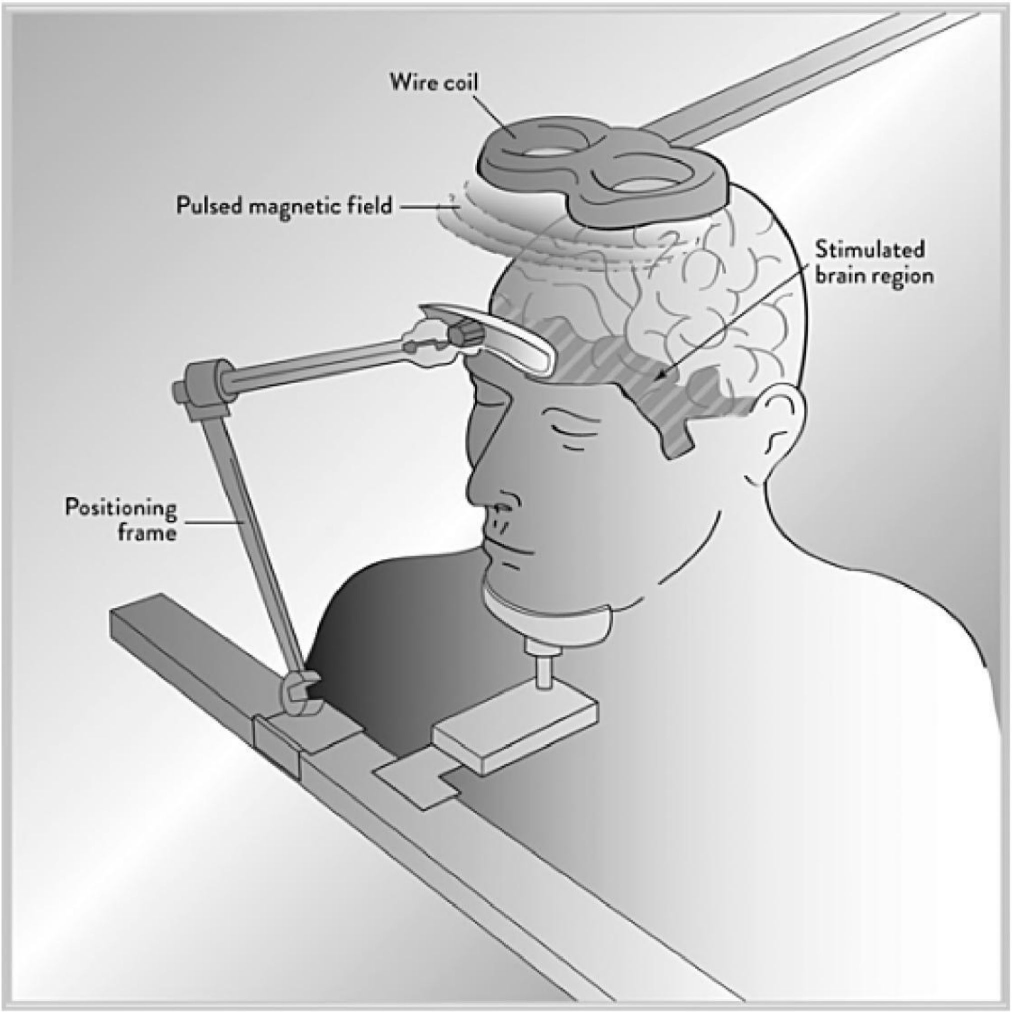


Figure 6. We see the transcranial electromagnetic scanner and the magnetoencephalograph, which uses magnetism rather than radio waves to penetrate the skull and determine the nature of thoughts within the brain. Magnetism can temporarily silence parts of the brain, allowing scientists to safely determine how these regions perform without relying on stroke victims.

DEEP BRAIN STIMULATION

Yet another tool that has proven vital to neurologists is deep brain stimulation (DBS). The probes originally used by Dr. Penfield were relatively crude. Today these electrodes can be hairlike and reach specific areas of the brain deep within its interior. Not only has DBS

—gravitational, electromagnetic, weak nuclear, and strong nuclear—that rule the universe. (Physicists have tried to find evidence for a fifth force, but so far all such attempts have failed.)

The electromagnetic force, which lights up our cities and represents the energy of electricity and magnetism, is the source of almost all the new scanning technologies (with the exception of the PET scan, which is governed by the weak nuclear force). Because physicists have had over 150 years of experience working with the electromagnetic force, there is no mystery in creating new electric and magnetic fields, so any new brain-scanning technology will most likely be a novel modification of existing technologies, rather than being something entirely new. As with most technology, the size and cost of these machines will drop, vastly increasing the widespread use of these sophisticated instruments. Already physicists are doing the basic calculations necessary to make an MRI machine fit into a cell phone. At the same time, the fundamental challenge facing these brain scans is resolution, both spatial and temporal. The spatial resolution of MRI scans will increase as the magnetic field becomes more uniform and as the electronics become more sensitive. At present, MRI scans can see only dots or voxels within a fraction of a millimeter. But each dot may contain hundreds of thousands of neurons. New scanning technology should reduce this even further. The holy grail of this approach would be to create an MRI-like machine that could identify individual neurons and their connections.

The temporal resolution of MRI machines is also limited because they analyze the flow of oxygenated blood in the brain. The machine itself has very good temporal resolution, but tracing the flow of blood slows it down. In the future, other MRI machines will be able to locate different substances that are more directly connected to the firing of neurons, thereby allowing real-time analysis of mental processes. No matter how spectacular the successes of the past fifteen years, then, they were just a taste of the future.

NEW MODELS OF THE BRAIN

Historically, with each new scientific discovery, a new model of the brain has emerged. One of the earliest models of the brain was the “homunculus,” a little man who lived inside the brain and made all the decisions. This picture was not very helpful, since it did not explain what was happening in the brain of the homunculus. Perhaps there was a homunculus hiding inside the homunculus.

With the arrival of simple mechanical devices, another model of the brain was proposed: that of a machine, such as a clock, with mechanical wheels and gears. This analogy was useful for scientists and inventors like Leonardo da Vinci, who actually designed a mechanical man.

During the late 1800s, when steam power was carving out new empires, another analogy emerged, that of a steam engine, with flows of energy competing with one another. This hydraulic model, historians have conjectured, affected Sigmund Freud's picture of the brain, in which there was a continual struggle between three forces: the ego (representing the self and rational thought), the id (representing repressed desires), and the superego (representing our conscience). In this model, if too much pressure built up because of a conflict among these three, there could be a regression or general breakdown of the entire system. This model was ingenious, but as even Freud himself admitted, it required detailed studies of the brain at the neuronal level, which would take another century.

Early in the last century, with the rise of the telephone, another analogy surfaced—that of a giant switchboard. The brain was a mesh of telephone lines connected into a vast network. Consciousness was a long row of telephone operators sitting in front of a large panel of switches, constantly plugging and unplugging wires. Unfortunately, this model said nothing about how these messages were wired together to form the brain.

With the rise of the transistor, yet another model became fashionable: the computer. The old-fashioned switching stations were replaced by microchips containing hundreds of millions of transistors. Perhaps the "mind" was just a software program running on "wetware" (i.e., brain tissue rather than transistors). This model is an enduring one, even today, but it has limitations. The transistor model cannot explain how the brain performs computations that would require a computer the size of New York City. Plus the brain has no programming, no Windows operating system or Pentium chip. (Also, a PC with a Pentium chip is extremely fast, but it has a bottleneck. All calculations must pass through this single processor. The brain is the opposite. The firing of each neuron is relatively slow, but it more than makes up for this by having 100 billion neurons processing data simultaneously. Therefore a slow parallel processor can trump a very fast single processor.)

The most recent analogy is that of the Internet, which lashes together billions of computers. Consciousness, in this picture, is an "emergent" phenomenon, miraculously arising out of the collective action of billions of neurons. (The problem with this picture is that it says absolutely nothing about how this miracle occurs. It brushes all the complexity of the brain under the rug of chaos theory.)

No doubt each of these analogies has kernels of truth, but none of them truly captures the complexity of the brain. However, one analogy for the brain that I have found useful (albeit still imperfect) is that of a large corporation. In this analogy, there is a huge bureaucracy and lines of authority, with vast flows of information channeled between different offices. But the important information eventually winds up at the command center with the CEO. There the final decisions are made.

If this analogy of the brain to a large corporation is valid, then it should be able to explain certain peculiar features of the brain:

- **Most information is "subconscious"**—that is, the CEO is blissfully unaware of the vast, complex information that is constantly flowing inside the bureaucracy. In fact, only a tiny amount of information finally reaches the desk of the CEO, who can be compared to the prefrontal cortex. The CEO just has to know information important enough to get his attention; otherwise, he would be paralyzed by an avalanche of extraneous information.

  This arrangement is probably a by-product of evolution, since our ancestors would have been overwhelmed with superfluous, subconscious information flooding their brains when facing an emergency. We are all mercifully unaware of the trillions of calculations being processed in our brains. Upon encountering a tiger in the forest, one does not have to be bothered with the status of our stomach, toes, hair, etc. All one has to know is how to run.
- **"Emotions" are rapid decisions made independently at a lower level.** Since rational thought takes many seconds, this means that it is often impossible to make a reasoned response to an emergency; hence lower-level brain regions must rapidly assess the situation and make a decision, an emotion, without permission from the top.

  So emotions (fear, anger, horror, etc.) are instantaneous red flags made at a lower level, generated by evolution, to warn the command center of possibly dangerous or serious situations. We have little conscious control over emotions. For example, no matter how much we practice giving a speech to a large audience, we still feel nervous.

Rita Carter, author of *Mapping the Mind*, writes, “Emotions are not feelings at all but a set of body-rooted survival mechanisms that have evolved to turn us away from danger and propel us forward to things that may be of benefit.”

- **There is a constant clamoring for the attention of the CEO.** There is no single homunculus, CPU, or Pentium chip making decisions; instead, the various subcenters within the command center are in constant competition with one another, vying for the attention of the CEO. So there is no smooth, steady continuity of thought, but the cacophony of different feedback loops competing with one another. The concept of “I,” as a single, unified whole making all decisions continuously, is an illusion created by our own subconscious minds.

  Mentally we feel that our mind is a single entity, continuously and smoothly processing information, totally in charge of our decisions. But the picture emerging from brain scans is quite different from the perception we have of our own mind.

  MIT professor Marvin Minsky, one of the founding fathers of artificial intelligence, told me that the mind is more like a “society of minds,” with different submodules, each trying to compete with the others.

  When I interviewed Steven Pinker, a psychologist at Harvard University, I asked him how consciousness emerges out of this mess. He said that consciousness was like a storm raging in our brain. He elaborated on this when he wrote that “the intuitive feeling we have that there’s an executive ‘I’ that sits in a control room of our brain, scanning the screens of the senses and pushing the buttons of our muscles, is an illusion. Consciousness turns out to consist of a maelstrom of events distributed across the brain. These events compete for attention, and as one process outshouts the others, the brain rationalizes the outcome after the fact and concocts the impression that a single self was in charge all along.”
- **Final decisions are made by the CEO in the command center.** Almost all the bureaucracy is devoted to accumulating and assembling information for the CEO, who meets only with the directors of each division. The CEO tries to mediate all the conflicting information pouring into the command center. The buck stops here. The CEO, located in the prefrontal cortex, has to make the final decision. While most decisions are made by instinct in animals, humans make higher-level decisions after sifting through different bodies of information from our senses.
- **Information flows are hierarchical.** Because of the vast amount of information that must flow upward toward the CEO’s office, or downward to the support staff, information must be arranged in complex arrays of nested networks, with many branches. Think of a pine tree, with the

command center on top and a pyramid of branches flowing downward, branching out into many subcenters.

There are, of course, differences between a bureaucracy and the structure of thought. The first rule of any bureaucracy is that "it expands to fill the space allotted to it." But wasting energy is a luxury the brain cannot afford. The brain consumes only about twenty watts of power (the power of a dim lightbulb), but that is probably the maximum energy it can consume before the body becomes dysfunctional. If it generates more heat, it will cause tissue damage. Therefore the brain is constantly using shortcuts to conserve energy. We will see throughout this book the clever and ingenious devices that evolution has crafted, without our knowledge, to cut corners.

IS "REALITY" REALLY REAL?

Everyone knows the expression "seeing is believing." Yet much of what we see is actually an illusion. For example, when we see a typical landscape, it seems like a smooth, movielike panorama. In reality, there is a gaping hole in our field of vision, corresponding to the location of the optic nerve in the retina. We should see this large ugly black spot wherever we look. But our brains fill in that hole by papering it over, by averaging it out. This means that part of our vision is actually fake, generated by our subconscious minds to deceive us.

Also, we see only the center of our field of vision, called the fovea, with clarity. The peripheral part is blurry, in order to save energy. But the fovea is very small. To capture as much information as possible with the tiny fovea, the eye darts around constantly. This rapid, jiggling motion of our eyes is called saccades. All this is done subconsciously, giving us the false impression that our field of vision is clear and focused.

When I was a child and first saw a diagram showing the electromagnetic spectrum in its true glory, I was shocked. I had been totally unaware that huge parts of the EM spectrum (e.g., infrared light, UV light, X-rays, gamma rays) were totally invisible to us. I began to realize that what I saw with my eyes was only a tiny, crude approximation of reality. (There is an old saying: "If appearance and essence were the same thing, there would be no need for science.") We have sensors in the retina that can detect only red, green, and blue. This means that we've never actually seen yellow, brown, orange, and a host of other colors. These colors do exist, but our brain can approximate each right hemisphere may be conscious simultaneously in different, even in mutually conflicting, mental experiences that run along in parallel."

When I interviewed Dr. Michael Gazzaniga of the University of California, Santa Barbara, an authority on split-brain patients, I asked him how experiments can be done to test this theory. There are a variety of ways to communicate separately to each hemisphere without the knowledge of the other hemisphere. One can, for example, have the subject wear special glasses on which questions can be shown to each eye separately, so that directing questions to each hemisphere is easy. The hard part is trying to get an answer from each hemisphere. Since the right brain cannot speak (the speech centers are located only in the left brain), it is difficult to get answers from the right brain. Dr. Gazzaniga told me that to find out what the right brain was thinking, he created an experiment in which the (mute) right brain could "talk" by using Scrabble letters.

He began by asking the patient's left brain what he would do after graduation. The patient replied that he wanted to become a draftsman. But things got interesting when the (mute) right brain was asked the same question. The right brain spelled out the words: "automobile racer." Unknown to the dominant left brain, the right brain secretly had a completely different agenda for the future. The right brain literally had a mind of its own.

Rita Carter writes, "The possible implications of this are mind-boggling. It suggests that we might all be carrying around in our skulls a mute prisoner with a personality, ambition, and self-awareness quite different from the day-to-day entity we believe ourselves to be."

Perhaps there is truth to the oft-heard statement that "inside him, there is someone yearning to be free." This means that the two hemispheres may even have different beliefs. For example, the neurologist V. S. Ramanchandran describes one split-brain patient who, when asked if he was a believer or not, said he was an atheist, but his right brain declared he was a believer. Apparently, it is possible to have two opposing religious beliefs residing in the same brain. Ramachandran continues: "If that person dies, what happens? Does one hemisphere go to heaven and the other go to hell? I don't know the answer to that."

(It is conceivable, therefore, that a person with a split-brain personality might be both Republican and Democrat at the same time. If you ask him whom he will vote for, he will give you the candidate of the left brain, since the right brain cannot speak. But you can imagine the chaos in the voting booth when he has to pull the lever with one hand.)

WHO IS IN CHARGE?

One person who has spent considerable time and done much research to understand the problem of the subconscious mind is Dr. David Eagleman, a neuroscientist at the Baylor College of Medicine. When I interviewed him, I asked him, If most of our mental processes are subconscious, then why are we ignorant of this important fact? He gave an example of a young king who inherits the throne and takes credit for everything in the kingdom, but hasn't the slightest clue about the thousands of staff, soldiers, and peasants necessary to maintain the throne.

Our choice of politicians, marriage partners, friends, and future occupations are all influenced by things that we are not conscious of. (For example, it is an odd result, he says, that "people named Denise or Dennis are disproportionately likely to become dentists, while people named Laura or Lawrence are more likely to become lawyers, and people with names like George or Georgina to become geologists.") This also means that what we consider to be "reality" is only an approximation that the brain makes to fill in the gaps. Each of us sees reality in a slightly different way. For example, he pointed out, "at least 15 percent of human females possess a genetic mutation that gives them an extra (fourth) type of color photoreceptor—and this allows them to discriminate between colors that look identical to the majority of us with a mere three types of color photoreceptors."

Clearly, the more we understand the mechanics of thought, the more questions arise. Precisely what happens in the command center of the mind when confronted with a rebellious shadow command center? What do we mean by "consciousness" anyway, if it can be split in half? And what is the relationship between consciousness and "self" and "self-awareness"?

If we can answer these difficult questions, then perhaps it will pave the way for understanding nonhuman consciousness, the consciousness of robots and aliens from outer space, for example, which may be entirely different from ours.

So let us now propose a clear answer to this deceptively complex question: What is consciousness?

> The mind of man is capable of anything … because everything is in it, all the past as well as all the future.
> —JOSEPH CONRAD

> Consciousness can reduce even the most fastidious thinker to blabbering incoherence.
> —COLIN MCGINN

2 CONSCIOUSNESS—A PHYSICIST'S VIEWPOINT

The idea of consciousness has intrigued philosophers for centuries, but it has resisted a simple definition, even to this day. The philosopher David Chalmers has cataloged more than twenty thousand papers written on the subject; nowhere in science have so many devoted so much to create so little consensus. The seventeenth-century thinker Gottfried Leibniz once wrote, "If you could blow the brain up to the size of a mill and walk about inside, you would not find consciousness."

Some philosophers doubt that a theory of consciousness is even possible. They claim that consciousness can never be explained since an object can never understand itself, so we don't even have the mental firepower to solve this perplexing question. Harvard psychologist Steven Pinker writes, "We cannot see ultraviolet light. We cannot mentally rotate an object in the fourth dimension. And perhaps we cannot solve conundrums like free will and sentience."

In fact, for most of the twentieth century, one of the dominant theories of psychology, behaviorism, denied the importance of consciousness entirely. Behaviorism is based on the idea that only the objective behavior of animals and people is worthy of study, not the subjective, internal states of the mind.

Others have given up trying to define consciousness, and try simply to describe it. Psychiatrist Giulio Tononi has said, "Everybody knows what consciousness is: it is what abandons you every night when you fall into dreamless sleep and returns the next morning when you wake up."

Although the nature of consciousness has been debated for centuries, there has been little resolution. Given that physicists created many of the inventions that have made the explosive advancements in brain science possible, perhaps it will be useful to follow an example from physics in reexamining this ancient question.

HOW PHYSICISTS UNDERSTAND THE UNIVERSE

When a physicist tries to understand something, first he collects data and then he proposes a "model," a simplified version of the object he is studying that captures its essential features. In physics, the model is described by a series of parameters (e.g., temperature, energy, time). Then the physicist uses the model to predict its future evolution by simulating its motions. In fact, some of the world's largest supercomputers are used to simulate the evolution of models, which can describe protons, nuclear explosions, weather patterns, the big bang, and the center of black holes. Then you create a better model, using more sophisticated parameters, and simulate it in time as well.

For example, when Isaac Newton was puzzling over the motion of the moon, he created a simple model that would eventually change the course of human history: he envisioned throwing an apple in the air. The faster you threw the apple, he reasoned, the farther it would travel. If you threw it fast enough, in fact, it would encircle the Earth entirely, and might even return to its original point. Then, Newton claimed, this model represented the path of the moon, so the forces that guided the motion of the apple circling the Earth were identical to the forces guiding the moon.

But the model, by itself, was still useless. The key breakthrough came when Newton was able to use his new theory to simulate the future, to calculate the future position of moving objects. This was a difficult problem, requiring him to create an entirely new branch of mathematics, called calculus. Using this new mathematics, Newton was then able to predict the trajectory of not just the moon, but also Halley's Comet and the planets. Since then, scientists have used Newton's laws to simulate the future path of moving objects, from cannonballs, machines, automobiles, and rockets to asteroids and meteors, and even stars and galaxies.

The success or failure of a model depends on how faithfully it very complex behaviors requiring a vastly expanded brain, so Level II consciousness coincides with the formation of new structures of the brain in the form of the limbic system. As noted earlier, the limbic system includes the hippocampus (for memories), amygdala (for emotions), and the thalamus (for sensory information), all of which provide new parameters for creating models in relation to others. The number and type of feedback loops therefore change.

We define the degree of Level II consciousness as the total number of distinct feedback loops required for an animal to interact socially with members of its grouping. Unfortunately, studies of animal consciousness are extremely limited, so little work has been done to catalog all the ways in which animals communicate socially with one another. But to a crude first approximation, we can estimate Level II consciousness by counting the number of fellow animals in its pack or tribe and then listing the total number of ways in which the animal interacts emotionally with each one. This would include recognizing rivals and friends, forming bonds with others, reciprocating favors, building coalitions, understanding your status and the social ranking of others, respecting the status of your superiors, displaying your power over your inferiors, plotting to rise on the social ladder, etc. (We exclude insects from Level II, because although they have social relations with members of their hive or group, they have no emotions as far as we can tell.)

Despite the lack of empirical studies of animal behaviors, we can give a very rough numerical rank to Level II consciousness by listing the total number of distinct emotions and social behaviors that the animal can exhibit. For example, if a wolf pack consists of ten wolves, and each wolf interacts with all the others with fifteen different emotions and gestures, then its level of consciousness, to a first approximation, is given by the product of the two, or 150, so it would have Level II:150 consciousness. This number takes into account both the number of other animals it has to interact with as well as the number of ways it can communicate with each one. This number only approximates the total number of social interactions that the animal can display, and will undoubtedly change as we learn more about its behavior.

(Of course, because evolution is never clean and precise, there are caveats that we have to explain, such as the level of consciousness of social animals that are solitary hunters. We will do so in the notes.)

With this framework for consciousness, we see that humans are not unique, and that there is a continuum of consciousness. As Charles Darwin once commented, “The difference between man and the higher animals, great as it is, is certainly one of degree and not of kind.” But what separates human consciousness from the consciousness of animals? Humans are alone in the animal kingdom in understanding the concept of tomorrow. Unlike animals, we constantly ask ourselves “What if?” weeks, months, and even years into the future, so I believe that Level III consciousness creates a model of its place in the world and then simulates it into the future, by making rough predictions. We can summarize this as follows:

> **Human consciousness is a specific form of consciousness that creates a model of the world and then simulates it in time, by evaluating the past to simulate the future. This requires mediating and evaluating many feedback loops in order to make a decision to achieve a goal.**

By the time we reach Level III consciousness, there are so many feedback loops that we need a CEO to sift through them in order to simulate the future and make a final decision. Accordingly, our brains differ from those of other animals, especially in the expanded prefrontal cortex, located just behind the forehead, which allows us to “see” into the future.

Dr. Daniel Gilbert, a Harvard psychologist, has written, “The greatest achievement of the human brain is its ability to imagine objects and episodes that do not exist in the realm of the real, and it is this ability that allows us to think about the future. As one philosopher noted, the human brain is an ‘anticipation machine,’ and ‘making the future’ is the most important thing it does.”

Using brain scans, we can even propose a candidate for the precise area of the brain where simulation of the future takes place. Neurologist Michael Gazzaniga notes that “area 10 (the internal granular layer IV), in the lateral prefrontal cortex, is almost twice as large in humans as in apes. Area 10 is involved with memory and planning¸ cognitive flexibility, abstract thinking¸ initiating appropriate behavior, and inhibiting inappropriate behavior, learning rules, and picking out relevant information from what is perceived through the senses.” (For this book, we will refer to this area, in which decision making is concentrated, as the dorsolateral prefrontal cortex, although there is some overlap with other areas of the brain.)

Although animals may have a well-defined understanding of their place in space and some have a degree of awareness of others, it is not clear if they systematically plan for the future and have an understanding of “tomorrow.” Most animals, even social animals with well-developed limbic systems, react to situations (e.g., the presence of predators or potential mates) by relying mainly on instinct, rather than systematically planning into the future.

For instance, mammals do not plan for the winter by preparing to hibernate, but largely follow instinct as the temperature drops. There is a feedback loop that regulates their hibernation. Their consciousness is dominated by messages coming in from their senses. There is no evidence that they systemically sift through various plans and schemes as they prepare to hibernate. Predators, when they use cunning and disguise to stalk an unsuspecting prey, do anticipate future events, but this planning is limited only to instinct and the duration of the hunt. Primates are adept at devising short-term plans (e.g., finding food), but there is no indication that they plan more than a few hours ahead.

Humans are different. Although we do rely on instinct and emotions in many situations, we also constantly analyze and evaluate information from many feedback loops. We do this by running simulations sometimes even beyond our own life span and even thousands of years into the future. The point of running simulations is to evaluate various possibilities to make the best decision to fulfill a goal. This occurs in the prefrontal cortex, which allows us to simulate the future and evaluate the possibilities in order to chart the best course of action.

This ability evolved for several reasons. First, having the ability to peer into the future has enormous evolutionary benefits, such as evading predators and finding food and mates. Second, it allows us to choose among several different outcomes and to select the best one.

Third, the number of feedback loops explodes exponentially as we go from Level 0 to Level I to Level II, so we need a “CEO” to evaluate all these conflicting, competing messages. Instinct is no longer enough. There has to be a central body that evaluates each of these feedback loops. This distinguishes human consciousness from that of the animals. These feedback loops are evaluated, in turn, by simulating them into the future to obtain the best outcome. If we didn't have a CEO, chaos would ensue and we would have sensory overload.

A simple experiment can demonstrate this. David Eagleman describes how you can take a male stickleback fish and have a female fish trespass on its territory. The male gets confused, because it wants to mate with the female, but it also wants to defend its territory. As a result, the male stickleback fish will simultaneously attack the female while initiating courtship behavior. The male is driven into a frenzy, trying to woo and kill the female at the same time.

This works for mice as well. Put an electrode in front of a piece of cheese. If the mouse gets too close, the electrode will shock it. One feedback loop tells the mouse to eat the cheese, but another one tells the mouse to stay away and avoid being shocked. By adjusting the location of the electrode, you can get the mouse to oscillate, torn between two conflicting feedback loops. While a human has a CEO in its brain to evaluate the pros and cons of the situation, the mouse, governed by two conflicting feedback loops, goes back and forth. (This is like the proverb about the donkey that starves to death because it is placed between two equal bales of hay.)

Precisely how does the brain simulate the future? The human brain is flooded by a large amount of sensory and emotional data. But the key is to simulate the future by making causal links between events—that is, if A happens, then B happens. But if B happens, then C and D might result. This sets off a chain reaction of events, eventually creating a tree of possible cascading futures with many branches. The CEO in the prefrontal cortex evaluates the results of these causal trees in order to make the ultimate decision.

Let's say you want to rob a bank. How many realistic simulations of this event can you make? To do this, you have to think of the various causal links involving the police, bystanders, alarm systems, relations with fellow criminals, traffic conditions, the DA's office, etc. For a successful simulation of the robbery, hundreds of causal links may have to be evaluated.

It is also possible to measure this level of consciousness numerically. Let's say that a person is given a series of different situations like the one above and is asked to simulate the future of each. The sum total number of causal links that the person can make for all these situations can be

ALLEN LANE

Published by the Penguin Group
Penguin Books Ltd, 80 Strand, London WC2R 0RL, England
Penguin Group (USA) Inc., 375 Hudson Street, New York, New York 10014, USA
Penguin Group (Canada), 90 Eglinton Avenue East, Suite 700, Toronto, Ontario, Canada M4P 2Y3 (a division of Pearson Penguin Canada Inc.)
Penguin Ireland, 25 St Stephen's Green, Dublin 2, Ireland (a division of Penguin Books Ltd)
Penguin Group (Australia), 707 Collins Street, Melbourne, Victoria 3008, Australia (a division of Pearson Australia Group Pty Ltd)
Penguin Books India Pvt Ltd, 11 Community Centre, Panchsheel Park, New Delhi – 110 017, India
Penguin Group (NZ), 67 Apollo Drive, Rosedale, Auckland 0632, New Zealand (a division of Pearson New Zealand Ltd)
Penguin Books (South Africa) (Pty) Ltd, Block D, Rosebank Office Park, 181 Jan Smuts Avenue, Parktown North, Gauteng 2193, South Africa

Penguin Books Ltd, Registered Offices: 80 Strand, London WC2R 0RL, England

www.penguin.com

First published in the United States of America by Doubleday, a division of Random House, LLC 2014
First published in Great Britain by Allen Lane 2014

ISBN: 978-0-141-97588-7