

TREN

"BRAINWASHED"

BY

SATEL
&
LITIGENT

INTRODUCTION

Losing Our Minds in the Age of Brain Science

YOU'VE SEEN THE HEADLINES: This is your brain on love. Or God. Or envy. Or happiness. And they're reliably accompanied by articles boasting pictures of color-drenched brains—scans capturing Buddhist monks meditating, addicts craving cocaine, and college sophomores choosing Coke over Pepsi. The media—and even some neuroscientists, it seems—love to invoke the neural foundations of human behavior to explain everything from the Bernie Madoff financial fiasco to slavish devotion to our iPhones, the sexual indiscretions of politicians, conservatives' dismissal of global warming, and even an obsession with self-tanning.¹

Brains are big on campus, too. Take a map of any major university, and you can trace the march of neuroscience from research labs and medical centers into schools of law and business and departments of economics and philosophy. In recent years, neuroscience has merged with a host of other disciplines, spawning such new areas of study as neurolaw, neuroeconomics, neurophilosophy, neuromarketing, and neurofinance. Add to this the birth of neuroaesthetics, neurohistory, neuroliterature, neuromusicology, neuropolitics, and neurotheology. The brain has even wandered into such unlikely

redoubts as English departments, where professors debate whether scanning subjects' brains as they read passages from Jane Austen novels represents (a) a fertile inquiry into the power of literature or (b) a desperate attempt to inject novelty into a field that has exhausted its romance with psychoanalysis and postmodernism.²

Clearly, brains are hot. Once the largely exclusive province of neuroscientists and neurologists, the brain has now entered the popular mainstream. As a newly minted cultural artifact, the brain is portrayed in paintings, sculptures, and tapestries and put on display in museums and galleries. One science pundit noted, "If Warhol were around today, he'd have a series of silkscreens dedicated to the cortex; the amygdala would hang alongside Marilyn Monroe."³

The prospect of solving the deepest riddle humanity has ever contemplated—*itself*—by studying the brain has captivated scholars and scientists for centuries. But never before has the brain so vigorously engaged the public imagination. The prime impetus behind this enthusiasm is a form of brain imaging called functional magnetic resonance imaging (fMRI), an instrument that came of age a mere two decades ago, which measures brain activity and converts it into the now-iconic vibrant images one sees in the science pages of the daily newspaper.

As a tool for exploring the biology of the mind, neuroimaging has given brain science a strong cultural presence. As one scientist remarked, brain images are now "replacing Bohr's planetary atom as the symbol of science."⁴ With its implied promise of decoding the brain, it is easy to see why brain imaging would beguile almost anyone interested in pulling back the curtain on the mental lives of others: politicians hoping to manipulate voter attitudes, marketers tapping the brain to learn what consumers really want to buy, agents of the law seeking an infallible lie detector, addiction researchers trying to gauge the pull of temptations, psychologists and psychiatrists seeking the causes of mental illness, and defense attorneys fighting to prove that their clients lack malign intent or even free will.

The problem is that brain imaging cannot do any of these things—at least not yet.

AUTHOR Tom Wolfe was characteristically prescient when he wrote of fMRI in 1996, just a few years after its introduction, "Anyone who cares to get up early and catch a truly blinding twenty-first-century dawn will want to keep an eye on it."⁵ Now we can't look away.

Why the fixation? First, of course, there is the very subject of the scans: the brain itself. More complex than any structure in the known cosmos, the brain is a masterwork of nature endowed with cognitive powers that far outstrip the capacity of any silicon machine built to emulate it. Containing roughly 80 billion brain cells, or neurons, each of which communicates with thousands of other neurons, the three-pound universe cradled between our ears has more connections than there are stars in the Milky Way.⁶ How this enormous neural edifice gives rise to subjective feelings is one of the greatest mysteries of science and philosophy.

Now combine this mystique with the simple fact that pictures—in this case, brain scans—are powerful. Of all our senses, vision is the most developed. There are good evolutionary reasons for this arrangement: The major threats to our ancestors were apprehended visually; so were their sources of food. Plausibly, the survival advantage of vision gave rise to our reflexive bias for believing that the world is as we perceive it to be, an error that psychologists and philosophers call naive realism. This misplaced faith in the trustworthiness of our perceptions is the wellspring of two of history's most famously misguided theories: that the world is flat and that the sun revolves around the earth. For thousands of years, people trusted their raw impressions of the heavens. Yet, as Galileo understood all too well, our eyes can deceive us. He wrote in his *Dialogues* of 1632 that the Copernican model of the heliocentric universe commits a "rape upon the senses"—it violates everything our eyes tell us.⁷

Brain scan images are not what they seem either—or at least not how the media often depict them. Nor are brain-scan images what they seem. They are not photographs of the brain in action in real time. Scientists can't just look "in" the brain and see what it does. Those beautiful color-dappled images are actually representations of particular areas in the brain that are working the hardest—as measured by increased oxygen consumption—when a subject performs a task such as reading a passage or reacting to stimuli, such as pictures of faces. The powerful computer located within the scanning machine transforms changes in oxygen levels into the familiar candy-colored splorches indicating the brain regions that become especially active during the subject's performance. Despite well-informed inferences, the greatest challenge of imaging is that it is very difficult for scientists to look at a fiery spot on a brain scan and conclude with certainty what is going on in the mind of the person.⁸

Neuroimaging is a young science, barely out of its infancy, really. In such a fledgling enterprise, the half-life of facts can be especially brief. To regard research findings as settled wisdom is folly, especially when they emanate from a technology whose implications are still poorly understood. As any good scientist knows, there will always be questions to hone, theories to refine, and techniques to perfect. Nonetheless, scientific humility can readily give way to exuberance. When it does, the media often seem to have a ringside seat at the spectacle.

Several years ago, as the 2008 presidential election season was gearing up, a team of neuroscientists from UCLA sought to solve the riddle of the undecided, or swing, voter. They scanned the brains of swing voters as they reacted to photos and video footage of the candidates. The researchers translated the resultant brain activity into the voters' unspoken attitudes and, together with three political consultants from a Washington, D.C.-based firm called FKF Applied Research, presented their findings in the *New York Times* in an op-ed titled "This Is Your Brain on Politics."⁹ There, readers could view scans dotted with tangerine and neon yellow hot spots indicating

regions that "lit up" when the subjects were exposed to images of Hillary Clinton, Mitt Romney, John Edwards, and other candidates. Revealed in these activity patterns, the authors claimed, were "some voter impressions on which this election may well turn." Among those impressions was that two candidates had utterly failed to "engage" with swing voters. Who were these unpopular politicians? John McCain and Barack Obama, the two eventual nominees for president.

Another much-circulated study, published in 2008, "The Neural Correlates of Hate" came from neuroscientists at University College London. The researchers asked subjects to bring in photos of people they hated—generally ex-lovers, work rivals, or reviled politicians—as well as people about whom subjects felt neutrally. By comparing their responses—that is, patterns of brain activation elicited by the hated face—with their reaction to the neutral photos, the team claimed to identify the neurological correlates of intense hatred. Not surprisingly, much of the media coverage attracted by the study flew under the headline: "Hate Circuit' Found in Brain."

One of the researchers, Semir Zeki, told the press that brain scans could one day be used in court—for example, to assess whether a murder suspect felt a strong hatred toward the victim.¹⁰ Not so fast. True, these data do reveal that certain parts of the brain become more active when people look at images of people they hate and presumably feel contempt for them as they do so. The problem is that the illuminated areas on the scan are activated by many other emotions, not just hate. There is no newly discovered collection of brain regions that are wired together in such a way that they comprise the identifiable neural counterpart of hatred.

University press offices, too, are notorious for touting sensational details in their media-friendly releases: Here's a spot that lights up when subjects think of God ("Religion center found!"), or researchers find a region for love ("Love found in the brain"). Neuroscientists sometimes refer disparagingly to these studies as "blobology," their tongue-in-cheek label for studies that show which brain areas

become activated as subjects experience X or perform task Y. To repeat: It's all too easy for the nonexpert to lose sight of the fact that fMRI and other brain-imaging techniques do not literally read thoughts or feelings. By obtaining measures of brain oxygen levels, they show which regions of the brain are more active when a person is thinking, feeling, or, say, reading or calculating. But it is a rather daring leap to go from these patterns to drawing confident inferences about how people feel about political candidates or paying taxes, or what they experience in the throes of love.¹¹

Pop neuroscience makes an easy target, we know. Yet we invoke it because these studies garner a disproportionate amount of media coverage and shape public perception of what brain imaging can tell us. Skilled science journalists cringe when they read accounts claiming that scans can capture the mind itself in action. Serious science writers take pains to describe quality neuroscience research accurately. Indeed, an eddy of discontent is already forming. "Neuromania," "neurohubris," and "neurohype"—"neurobollocks," if you're a Brit—are just some of the labels that have been brandished, sometimes by frustrated neuroscientists themselves. But in a world where university press releases elbow one another for media attention, it's often the study with a buzzy storyline ("Men See Bikini-Clad Women as Objects, Psychologists Say") that gets picked up and dumbed down.¹²

The problem with such mindless neuroscience is not neuroscience itself. The field is one of the great intellectual achievements of modern science. Its instruments are remarkable. The goal of brain imaging is enormously important and fascinating: to bridge the explanatory gap between the intangible mind and the corporeal brain. But that relationship is extremely complex and incompletely understood. Therefore, it is vulnerable to being oversold by the media, some overzealous scientists, and neuroentrepreneurs who tout facile conclusions that reach far beyond what the current evidence warrants—fits of "premature extrapolation," as British neuroskeptic Steven

Poole calls them.¹³ When it comes to brain scans, seeing may be believing, but it isn't necessarily understanding.

Some of the misapplications of neuroscience are amusing and essentially harmless. Take, for instance, the new trend of neuromanagement books such as *Your Brain and Business: The Neuroscience of Great Leaders*, which advises nervous CEOs "to be aware that anxiety centers in the brain connect to thinking centers, including the PFC [prefrontal cortex] and ACC [anterior cingulate cortex]." The fad has, perhaps not surprisingly, infiltrated the parenting and education markets, too. Parents and teachers are easy marks for "brain gyms," "brain-compatible education," and "brain-based parenting," not to mention dozens of other unsubstantiated techniques. For the most part, these slick enterprises merely dress up or repack-age good advice with neuroscientific findings that add nothing to the overall program. As one cognitive psychologist quipped, "Unable to persuade others about your viewpoint? Take a Neuro-Prefix—influence grows or your money back."¹⁴

But reading too much into brain scans matters when real-world concerns hang in the balance. Consider the law. When a person commits a crime, who is at fault: the perpetrator or his or her brain? Of course, this is a false choice. If biology has taught us anything, it is that "my brain" versus "me" is a false distinction. Still, if biological roots can be identified—and better yet, captured on a brain scan as juicy blotches of color—it is too easy for nonprofessionals to assume that the behavior under scrutiny must be "biological" and therefore "hardwired," involuntary or uncontrollable. Criminal lawyers, not surprisingly, are increasingly drawing on brain images supposedly showing a biological defect that "made" their clients commit murder.

Looking to the future, some neuroscientists envision a dramatic transformation of criminal law. David Eagleman, for one, welcomes a time when "we may someday find that many types of bad behavior have a basic biological explanation [and] eventually think about bad decision making in the same way we think about any physical

process, such as diabetes or lung disease.”¹⁵ As this comes to pass, he predicts, “more juries will place defendants on the not-blameworthy side of the line.”¹⁶ But is this the correct conclusion to draw from neuroscientific data? After all, if every behavior is eventually traced to detectable correlates of brain activity, does this mean we can one day write off all troublesome behavior on a don’t-blame-me-blame-my-brain theory of crime? Will no one ever be judged responsible? Thinking through these profoundly important questions turns on how we understand the relationship between the brain and the mind.

THE mind cannot exist without the brain. Virtually all modern scientists, ourselves included, are “mind-body monists”: they believe that mind and brain are composed of the same material “stuff.” All subjective experience, from a frisson of fear to the sweetness of nostalgia, corresponds to physical events in the brain. Decapitation proves this point handily: no functioning brain, no mind. But even though the mind is produced by the action of neurons and brain circuits, the mind is not identical with the matter that produces it. There is nothing mystical or spooky about this statement, nor does it imply an endorsement of mind-body “dualism,” the dubious assertion that mind and brain are composed of different physical material. Instead, it means simply that one cannot use the physical rules from the cellular level to completely predict activity at the psychological level. By way of analogy, if you wanted to understand the text on this page, you could analyze the words by submitting their contents to an inorganic chemist, who could ascertain the precise molecular composition of the ink. Yet no amount of chemical analysis could help you understand what these words mean, let alone what they mean in the context of other words on the page.

Scientists have made great strides in reducing the organizational complexity of the brain from the intact organ to its constituent neurons, the proteins they contain, genes, and so on. Using this template,

we can see how human thought and action unfold at a number of explanatory levels, working upward from the most basic elements. At one of the lower tiers in this hierarchy is the neurobiological level, which comprises the brain and its constituent cells.¹⁷ Genes direct neuronal development; neurons assemble into brain circuits. Information processing, or computation, and neural network dynamics hover above. At the middle level are conscious mental states, such as thoughts, feelings, perceptions, knowledge, and intentions. Social and cultural contexts, which play a powerful role in shaping our thoughts, feelings, and behavior, occupy the highest landings of the hierarchy.

Problems arise, however, when we ascribe too much importance to the brain-based explanations and not enough to psychological or social ones. Just as one obtains differing perspectives on the layout of a sprawling city while ascending in a skyscraper’s glass elevator, we can gather different insights into human behavior at different levels of analysis.¹⁸

The key to this approach is recognizing that some levels of explanation are more informative for certain purposes than others. This principle is profoundly important in therapeutic intervention. A scientist trying to develop a medication for Alzheimer’s disease will toil on the lower levels of the explanatory ladder, perhaps developing compounds aimed at preventing the formation of the amyloid plaques and neurofibrillary tangles endemic to the disease. A marriage counselor helping a distraught couple, though, must work on the psychological level. Efforts by this counselor to understand the couple’s problems by subjecting their brains to fMRIs could be worse than useless because doing so would draw attention away from their thoughts, feelings, and actions toward each other—the level at which intervention would be most helpful.

This discussion brings us back to brain scans and other representations of brain-derived data. What can we infer from this information about what people are thinking and feeling or how their social world is influencing them? In a way, imaging rekindles the age-old

debate over whether brain equals mind. Can we ever fully comprehend the psychological by referring to the neural? This “hard problem,” as philosophers call it, is one of the most daunting puzzles in all of scientific inquiry. What would the solution even look like? Will the parallel languages of neurobiology and mental life ever converge on a common vernacular?¹⁹

Many believe it will. According to neuroscientist Sam Harris, inquiry into the brain will eventually and exhaustively explain the mind and, hence, human nature. Ultimately, he says, neuroscience will—and should—dictate human values. Semir Zeki, the British neuroscientist, and legal scholar Oliver Goodenough hail a “millennial” future, perhaps only decades away, [when] a good knowledge of the brain’s system of justice and of how the brain reacts to conflicts may provide critical tools in resolving international political and economic conflicts.” No less towering a figure than neuroscientist Michael Gazzaniga hopes for a “brain-based philosophy of life” based on an ethics that is “built into our brains. A lot of suffering, war, and conflict could be eliminated if we could agree to live by them more consciously.”²⁰

It’s no wonder, then, that some see neuroscientists as the “new high priests of the secrets of the psyche and explainers of human behavior in general.”²¹ Will we one day replace government bureaucrats with neurocrats? Though short on details—neuroscientists don’t say *how* brain science is supposed to determine human values or achieve world peace—their predictions are long on ambition. In fact, some experts talk of neuroscience as if it is the new genetics, that is, just the latest overarching narrative commandeered to explain and predict virtually all human behavior. And before genetic determinism there was the radical behaviorism of B. F. Skinner, who sought to explain human behavior in terms of rewards and punishments. Earlier in the late nineteenth and twentieth centuries, Freudianism posited that people were the products of unconscious conflicts and drives. Each of these movements suggested that the causes of our

actions are not what we think they are. Is neurodeterminism poised to become the next grand narrative of human behavior?

As a psychiatrist and a psychologist, we have followed the rise of popular neuroscience with mixed feelings. We’re delighted to see laypeople so interested in brain science, and we are excited by the promise of new neurophysiological discoveries. Yet we’re dismayed that much of the media diet consists of “vulgarized neuroscience,” as the science watchdog Neuroskeptic puts it, that offers facile and overly mechanistic explanations for complicated behaviors. We were both in training when modern neuroimaging techniques made their debut. The earliest major functional imaging technique (PET, or positron emission tomography) appeared in the mid-1980s. Less than a decade later, the near wizardry of fMRI was unveiled and soon became a prominent instrument of research in psychology and psychiatry. Indeed, expertise in imaging technology is becoming a sine qua non for graduate students in many psychology programs, increasing their odds of obtaining federal research grants and teaching posts and boosting the acceptance rates of their papers by top-flight journals. Many psychology departments now make expertise in brain imaging a requirement for their new hires.²²

The brain is said to be the final scientific frontier, and rightly so, in our view. Yet in many quarters brain-based explanations appear to be granted a kind of inherent superiority over all other ways of accounting for human behavior. We call this assumption “neurocentrism”—the view that human experience and behavior can be best explained from the predominant or even exclusive perspective of the brain.²³ From this popular vantage point, the study of the brain is somehow more “scientific” than the study of human motives, thoughts, feelings, and actions. By making the hidden visible, brain imaging has been a spectacular boon to neurocentrism.

Consider addiction. “Understanding the biological basis of pleasure leads us to fundamentally rethink the moral and legal aspects of

addiction,” writes neuroscientist David Linden.²⁴ This is popular logic among addiction experts, but to us, it makes little sense. Granted, there may be good reasons to reform the way the criminal justice system deals with addicts, but the biology of addiction is not one of them. Why? Because the fact that addiction is associated with neurobiological changes is not, in itself, proof that the addict is unable to choose. Just look at American actor Robert Downey Jr. He was once a poster boy for drug excess. “It’s like I have a loaded gun in my mouth and my finger’s on the trigger, and I like the taste of gun-metal,” he said. It seemed only a matter of time before he would meet a horrible end. But Downey entered rehab and decided to change his life. Why did Downey use drugs? Why did he decide to stop and to remain clean and sober? An examination of his brain, no matter how sophisticated the probe, could not tell us why and perhaps never will. The key problem with neurocentrism is that it devalues the importance of psychological explanations and environmental factors, such as familial chaos, stress, and widespread access to drugs, in sustaining addiction.

Our goal in this book is to bring some perspective to the bold speculations surrounding the promise of neuroscience. The chapters follow the migration of brain imaging (and occasionally brain-wave technologies, such as EEG, or electroencephalography) outside the lab and medical center and into marketing suites, drug-treatment clinics, and courtrooms.

We begin in Chapter 1 with a basic overview of fMRI. We review the principles of brain organization, how scans are constructed, and how simple studies are designed. We also examine some of the potential pitfalls of interpretation introduced by brain imaging. One of our major aims is to convey an appreciation for the staggering complexity of the brain and the implications of attempts to infer mental contents, such as thoughts, desires, intentions, and feelings, from brain-derived information.

In Chapter 2 we turn to neuromarketing. The impetus behind neuromarketing is the notion that consumers are inaccurate reporters of what they truly like and plan to purchase. If their brains can be tapped to measure their immediate responses to products or other stimuli, such as commercials or movie trailers, neuromarketers, who advise many Fortune 500 companies, believe that they can guide corporations in designing the most compelling ads and sales campaigns.

The biology of pathological desire figures prominently in Chapter 3 on addiction. Indeed, within research circles and some clinical venues, the idea that addiction is a “brain disease” is the dominant conceptual framework. The mechanical simplicity of this neurocentric view carries a seductive appeal that obscures the myriad other factors that drive addiction. A broader understanding of addiction that goes beyond its biological dimension is imperative if treatment is to be successful and recovery sustained.

The remaining chapters focus on the implications of the age of neuroscience for the law. Chapter 4 examines brain-based deception detection. Like neuromarketing, it is an arena animated by a major entrepreneurial spirit. Commercial outfits, such as No Lie MRI, claim to provide security firms, employers, and suspicious spouses with “unbiased methods for the detection of deception and other information stored in the brain.” Several times, No Lie and its competitor Cephus have tried to bring their evidence to court. We assess the scientific justification for using these techniques in forensic situations, where stakes are high. We also ask whether citizens are about to be confronted by the chilling words “We have a warrant to search your brain” anytime soon.

Chapter 5 on neurolaw puts neuroscience before the judge and jury. As the triers of fact consider the neurobiological facts of cases, neuroscientists such as David Eagleman and Sam Harris hope to see a general attitude “shift from blame to biology.” Yet the relationship between brain-based explanations of a defendant’s crime and what

they mean for holding that person responsible is by no means straightforward.²⁵

In Chapter 6 we explore a momentous question: What are the implications of neuroscience for individuals' freedom of choice? We generally think of ourselves as free agents who have the power to alter our destinies and earn praise or blame for our deeds, good and bad. But a number of prominent scholars claim that we are mistaken. "Our growing knowledge about the brain makes the notions of volition, culpability, and, ultimately, the very premise of the criminal justice system, deeply suspect," contends biologist Robert Sapolsky.²⁶ Will our coming to understand how the brain works necessitate a radically new way of thinking about human beings as moral agents worthy of blame and praise? As we will see, there is ample reason to doubt that it will.

Finally, in the Epilogue, we reprise what we have learned, and examine the crucial question of what neuroscience can—and cannot—tell us about human behavior. Brain imaging tools hold enormous potential for elucidating the neural correlates of everyday decisions, addiction, and mental illness. Yet these promising new technologies must not detract from the importance of levels of analysis other than the brain in explaining human behavior. Ours is an age in which brain research is flourishing—a time of truly great expectations. Yet it is also a time of mindless neuroscience that leads us to overestimate how much neuroscience can improve legal, clinical, and marketing practices, let alone inform social policy. Naive media, slick neuroentrepreneurs, and even an occasional overzealous neuroscientist exaggerate the capacity of scans to reveal the contents of our minds, exalt brain physiology as inherently the most valuable level of explanation for understanding behavior, and rush to apply underdeveloped, if dazzling, science for commercial and forensic use.²⁷

Granted, it is only natural that advances in knowledge about the brain make us think more mechanistically about ourselves. But if we become too carried away with this view, we may impede one of the

most challenging cultural projects looming in the years ahead: how to reconcile advances in brain science with personal, legal, and civic notions of freedom.

The neurobiological domain is one of brains and physical causes. The psychological domain, the domain of the mind, is one of people and their motives. Both are essential to a full understanding of why we act as we do and to the alleviation of human suffering. The brain and the mind are different frameworks for explaining experience. And the distinction between them is hardly an academic matter; it bears crucial implications for how we think about human nature, personal responsibility, and moral action.

THIS IS YOUR BRAIN ON

AHMADINEJAD

Or What Is Brain Imaging?

IN THE SPRING OF 2008, the folks at FKF Applied Research—the political consultants and neuroscientists behind the aforementioned swing-voter study—were at it again. This time they took journalist Jeffrey Goldberg of the *Atlantic* on a guided tour of his brain. The idea had been hatched at a family Passover seder at which Goldberg had spent the evening “issuing a series of ideologically contradictory, Manischewitz-fueled political pronouncements.” Fortunately for science, Bill Knapp, a political consultant and cofounder of FKF, was one of the seder guests. He suggested that if Goldberg wanted to get to the bottom of his confusion, he should submit to a brain scan to learn whether he was “neurologically wired for liberalism or conservatism.” As Goldberg understood the process, researchers would measure his brain’s responses to a series of images of famous politicians to uncover the truth about his “actual inclinations and predispositions by sidestepping the usual inhibition controls that can make focus-group testing unreliable.”¹

When Goldberg arrived at the facility, he was slid faceup into the mouth of a sleek MRI machine and asked to lie as still as a cadaver lest movement disrupt the readings. Despite noise-dampening

headphones, Goldberg could still hear the magnet in the state-of-the-art fMRI machine as it scanned his brain, a racket that's been likened to the sound of metal-cleated golf shoes tumbling in a clothes dryer followed by a long period of high-pitched ping-pong.² The researchers had fitted him with video goggles through which they flashed scores of photographs and film clips of cultural and political celebrities, including John McCain, Edie Falco, Golda Meir, Barack Obama, Yasser Arafat, Bruce Springsteen, George W. Bush, and Iranian president Mahmoud Ahmadinejad. A lesser man might have been daunted by the fusillade of images, but Goldberg's trials as a war correspondent in the Middle East seemed to have prepared him for a full hour inside the machine. Goldberg emerged with a clanging headache, but with his sense of humor intact. "If you haven't lain supine in a claustrophobia-inducing magnetized tunnel while watching Hillary Clinton talk about health care one inch from your eyeballs, well, you just haven't lived," he quipped.

Goldberg's brain, which the fMRI declared to be nonpartisan, displayed the same ambivalent reaction to Hillary Clinton as did the swing voters.³ The team's neuroscientist, Marco Iacoboni, speculated that enhanced activity in Goldberg's dorsolateral prefrontal cortex, an area linked to inhibition of ones spontaneous responses, indicated that he might be trying to "suppress unwanted emotions" about Clinton. Scanning also revealed that Goldberg loved Edie Falco by virtue of a strong response in the ventral striatum, an area of the brain that revs up at the prospect of reward. "I didn't need a million-dollar machine to tell me that," wrote Goldberg, an avowed fan of *The Sopranos*.

Goldberg's reaction to Ahmadinejad, however, took him by surprise. The sight of the Iranian leader also stimulated Goldberg's ventral striatum. "Reward!" Iacoboni exclaimed. "You'll have to explain this one." Although Goldberg couldn't fathom why Ahmadinejad would stimulate pleasurable thoughts, Joshua Freedman, a psychiatrist working with Iacoboni, offered a conjecture: "You seem to

believe that the Jewish people endure [and] that people who try to hurt the Jewish people ultimately fail. Therefore, you derive pleasure from believing that Ahmadinejad will also eventually fail." Freedman paused. "Or it means that you're a Shiite."

Goldberg reflected on his adventure in "vanity scanning," as he called it, and questioned the analytic rigor of the procedure. "I wondered to what degree this was truly scientific and to what degree it was 21st-century phrenology." Goldberg isn't the first to express such doubts. Frustrated experts have also dubbed overreager readings of fMRI images "neophrenology" in reference to the long-discredited method of revealing a person's personality traits and talents by "reading" his or her skull's bumps and depressions.³ Yet in many respects this analogy is unfair. Unlike phrenology, brain imaging is a technological marvel that does reveal something about the relationship between brain and mind. But exactly what can a "lit" brain region really tell us about an individual's thoughts and feelings?

This question sits at the leading edge of a large and time-honored body of inquiry: What can the workings of the brain tell us about the mind? Approaching this project through fMRI, one of the most up-to-the-minute and surely most mediate of neurotechnologies, hinges on scientists' ability to translate brain activity (mechanism) into accounts of what a person is thinking or feeling (meaning). Scientists, of course, cannot "read" specific thoughts with fMRI; they can only tell that brain regions already known to be associated with certain thoughts or feelings have demonstrated an increase in activity—hence the proper term "neural correlates" for the colorful dabs on brain scans. The value of brain scans in the courtroom and other venues rests on how accurately scientists can infer thoughts and feelings from these correlates. This challenging task began over a century ago, using far more primitive technologies.

NEUROIMAGING has come a long way from its earliest ancestor, the X-ray technique, invented in 1895 by German physicist William

Conrad Roentgen. His now-famous first X-ray showed the five bones of his wife's left hand, with the fourth bone encircled by a thick wedding ring. Roentgen's transformation of the previously hidden into the visible triggered a craze on both sides of the Atlantic. Department stores in Chicago, New York, and Paris installed X-ray slot machines so customers could view the skeletal anatomy of their hands, with the occasional customer fainting at the sight of his or her bones. A Parisian physician, Hippolyte-Ferdinand Baraduc, even claimed that he could use X-rays to photograph his own ideas and feelings. He called the resulting pictures "psychicons," or images of the mind. The X-ray, of course, is mute when it comes to the brain, let alone the mind, because the rays cannot pass readily through the skull's thick walls.⁴

At the turn of the twentieth century, scientists developed ventriculography, a method of pumping air into the brain's ventricles—hollow spaces that drain fluid from the brain—to increase the pressure inside and exaggerate density differentials across regions. In the early 1970s, computerized axial tomography (CT or CAT) scans enabled neuroradiologists to distinguish the white and gray matter of the brain from the ventricles that run through it. The technique uses high-density X-rays to capture images in slices and produce a three-dimensional model of the brain. A decade later, structural MRIs (magnetic resonance images) came on the radiographic scene, yielding an increasingly precise representation of brain anatomy. Structural MRI can detect static problems, such as tumors, blood clots, and deformed blood vessels. Taken together, MRIs and CT scans provide valuable information about fixed anatomy but leave us largely in the dark about the brain's functioning.⁵

That limitation began to change with the development of positron emission tomography (PET), one of the earliest three-dimensional functional imaging techniques. In contrast to structural techniques, PET and other functional methods allow neuroscientists to image the brain in action. Introduced in the 1980s, PET measures

brain metabolism or brain blood flow by deploying radioactive tracer molecules. The underlying principle is that when brain cells are active, they need more energy in the form of glucose or oxygen. The tracer, typically low-dose glucose labeled with a radioisotope, is either injected directly into a vein or inhaled. The glucose travels to the most active brain cells, where it emits energy (positrons) that are detected and displayed as a glowing "hot spot" on a PET scan. Although PET can also be used to examine the brain while subjects respond to stimuli or perform tasks, neuroscientists tend to prefer fMRI for that purpose because it has higher spatial and temporal resolution and does not involve radioactive material.⁶

Functional MRI leverages the fact that everything the brain enables us to do—feel, think, perceive, and act—is linked, or correlated, with changes in oxygen consumption and regional blood flow in the brain. When a person responds to a task, such as looking at photos or solving a math problem, specific regions of the brain are typically engaged and receive more oxygen-laden, or oxygenized, blood. The increased blood flow and the boost in oxygen associated with it are proxies for increased activation of neurons. We say "increased" because the entire living brain is always on; blood is always circulating and oxygen is always being consumed. The only truly silent brain is a dead brain.

Measuring the concentration of oxygen dissolved in the blood, therefore, is the key to detecting brain activity. The large and immensely powerful magnet within the fMRI machine can measure the influx of blood to areas of the brain because blood that is carrying more oxygen has different magnetic properties than blood that has already given up its oxygen to supply neurons. The relative concentrations of oxygenated and oxygen-depleted, or deoxygenated, blood in a small area of brain tissue creates a signal known technically as the BOLD (blood-oxygen-level-dependent) response. The higher the ratio of oxygenated to deoxygenated blood in a particular area, the higher the energy consumption in that region.⁷

During experiments, researchers do not simply ask subjects to perform a task and then measure their brain activity. They measure brain activity while subjects are engaged in a task, perhaps reacting to faces, and compare it with a baseline signal present while a subject is sitting with eyes closed and mind as blank as possible. Imagine an experiment designed to identify the neural regions associated with reading aloud. Researchers ask subjects to read letters silently to themselves as they appear on a screen and then ask the subjects to read them aloud. Presumably, these two tasks engage all of the same mental processes except one. If one “subtracts” the signal generated when the subject reads silently from the signal produced when he or she reads aloud, the remaining nonoverlapping region will ostensibly be associated with speech. Brain regions engaged in common functions—such as attention and visual and mental processing of the letters, which both tasks require—will probably cancel each other out so that they appear dark in the final brain scan.

During such an experiment, the scanner’s computer acquires BOLD data and collects them in tiny three-dimensional units called voxels, a portmanteau term derived from the words “volume” and “pixel.” A typical brain holds about 50,000 such voxels, each measuring about three cubic millimeters. The subtraction step we described takes place on a voxel-by-voxel level. Each voxel is then assigned a color depending on the strength of the difference in activation of that individual voxel between the control and experimental conditions. The computer then generates an image highlighting the regions that become more active in one condition relative to the other. By convention, researchers use color gradations to reflect the likelihood that the subtraction—the difference in activation between the resting and the stimulated state (or between two stimulated states)—was not due to chance. The brighter the color of a region, the greater the confidence the investigator has in the difference. Thus a bright color like yellow might mean that there is only one chance in a thousand that the difference between brain activations in a given area is

due to luck, whereas a darker color like purple might mean that the chances are higher, and that the brain differences are more likely to be attributable to random fluctuations in the data.

Finally, the computer filters out background noise and prepares the data to be mapped onto a three-dimensional template of a human brain. The final brain scan that we see in a magazine or on television rarely portrays the brain activity of a single person. Instead, it almost always represents the averaged results of all participants in the study. As we noted in the Introduction, any resemblance between brain scans and photographs is illusory. Photos capture images in real time and space. Functional imaging scans are constructed from information derived from the magnetic properties of blood flowing in the brain. If we removed half of the skull to observe the surface of the living brain in action, we wouldn’t see a multicolored light show as various areas become active during thinking, feeling, and behaving. As striking as they are, scans are far less immediate; at their most accurate, they simply represent local activation based on statistical differences in BOLD signals.

FUNCTIONAL brain imaging is the most recent chapter in the centuries-long quest to map and comprehend the connection between the brain and the mind. Classically, the mind was regarded as the thinking part of the soul, but unlike the soul, which is by definition immaterial and is believed to survive after death, there is nothing spooky or nonmaterial about the mind. The brain enables the mind, and when the brain dies, so does the mind. The Greek physician Hippocrates, who lived around 400 BCE, is believed to have been the first to posit that the brain creates the mind. Observations of individuals with head injuries led him to conclude that “from the brain, and from the brain alone, arise our pleasures, joys, laughter and jests, as well as our sorrows, pain, grief, and tears. . . . It is the brain which makes us mad or delirious.” The Epicureans of 300 BCE also believed that the human soul does not survive the death of the body. This materialist

view was to be overshadowed for centuries by the dualist doctrine put forward by Hippocrates's contemporary, Plato.⁸

Plato believed that the mind, or soul, as he called it, was immortal. It floated in parallel with a person's corporeal brain, which controlled perception and movement. Components of the Platonic mind—reason, will, and desire—somehow preexisted the individual and survived after death. Plato's version of dualism prevailed more or less intact over the next five centuries until it was succeeded by the ideas of the famed Roman physician Galen, circa 200 CE. Galen posited that such faculties as memory, intellect, and imagination—that is, the rational soul—swirled within the brain's ventricles. His view was adopted by the early church fathers, who were devout dualists.

After lying dormant during a medieval interlude of several centuries, the materialist-dualist contest reawakened during the seventeenth-century French Enlightenment, when the great mathematician and philosopher René Descartes introduced another variant of dualism. Descartes was the first to advance the idea, correctly, as it turned out, that emotions, memories, and sensory perceptions are functions of the material brain. But distinct from the mechanical brain, he maintained, was a nonmaterial mind, or rational soul, that was capable of language, mathematics, consciousness, will, doubt, and understanding. The mind and the brain were connected, Descartes believed famously, through a small nub of tissue near the center of the brain called the pineal gland.⁹

Throughout the eighteenth and nineteenth centuries, anatomists and physiologists began to forge a clear association between the brain and abstract thought, emotion, and behavior. By the close of the nineteenth century, scientists, physicians, and psychologists mostly agreed that the phenomenological mind arose from the physical brain. Still, they were mystified by how the brain's chemical and electrical actions could give rise to the experience of emotional states, a problem known as the mind-brain (or mind-body) problem. Solving that problem, according to William James, the founding father of American

psychology, would be "the scientific achievement before which past achievements fail." James built his own magisterial science of mental life on self-reports of his patients. Through their introspection, he developed theories of emotion, perception, imagination, and memory.¹⁰

As Paul Bloom notes, data suggest that even today, most adults are implicit mind-body dualists: they see the mind as largely or entirely distinct from the workings of the brain. This implicit dualism may help to explain why brain imaging studies garner so much media attention. Their results seem surprising, even fascinating, to many people. ("Wow, you mean that depression is actually in the brain? And love too?") "We intuitively think of ourselves as non-physical, and so it is a shock, and endlessly interesting, to see our brains at work in the act of thinking," remarks Bloom.¹¹

Most nineteenth-century investigators relied on crude experiments to gain better insight into the human brain. Eager to apply a scientific approach, these scientists and neurologists, physicians who treated diseases of the brain and nerves, resorted to surgically destroying or deactivating parts of the brains of animals. After the operation, they observed how rabbits, pigeons, and cats moved and responded to stimuli. Similarly, to identify areas involved in sensory perception and movement control, researchers applied electric current directly to specific regions of animal brains. Research on humans, however, required either less invasive measures or nonliving subjects. By dissecting the brains of individuals who had died of head injuries, tumors, infections, or strokes, neurologists and anatomists acquired considerable insights into the relationships between anatomy, on the one hand, and emotion, intellectual function, and behavior, on the other.¹²

Possibly the most famous brain-injury case was that of Phineas Gage. A railroad foreman in Vermont, Gage lost much of his left prefrontal cortex in a grisly 1848 accident in which a long iron rod shot upward through his left cheek and flew out the top of his skull. Miraculously, Gage survived, but his former even-keeled temperament

gave way to profanity, grandiosity, and belligerence. Gage's accident, later buttressed by more systematic research, helped demonstrate that the frontal lobes are the primary place, or node, at which a vast amount of neural processing comes together to regulate impulse control and social judgment.¹³

PHRENOLOGY was one of the first major brain-related theories of human behavior. During the 1800s it spread throughout Europe and the United States. Developed by Franz Joseph Gall, an esteemed Austrian-German anatomist, phrenology attempted to construct a science of brain function and human behavior. Gall believed that the mind was wholly situated within the brain. Phrenologists "read" personality by examining bumps and depressions on the skull that were supposedly responsible for dozens of traits, such as wit, inquisitiveness, and benevolence. Better-developed organs, Gall believed, pushed out areas of the overlying skull and formed bumps on the outer surface. Indentations in the cranium, in contrast, marked the weakest organs; although they had failed to grow to normal size, they could be developed through exercise, like a muscle. Accordingly, people regularly consulted phrenologists to learn their natural talents and receive advice on the type of career and life partner that would best match their brains.¹⁴

During a triumphant European tour between 1805 and 1807, Gall lectured to crowned heads of state, universities, and scientific societies and even received a commemorative medal from the king of Prussia—"He found a way to espy the workshop of the soul," read the inscription. Most of Gall's scientific contemporaries, however, were less enchanted. The predictive value of phrenology was dismal. Also, different examiners reading the same person's head routinely came to different conclusions about his or her personality.¹⁵

This is exactly what happened to Mark Twain. In the early 1870s, the great American humorist (and phrenology skeptic) had his head examined in London by the famous phrenologist Lorenzo Fowler. As

Twain described the visit in his autobiography, he was "glad of an opportunity to personally test [Fowler's] art" and disguised himself by using a fictitious name. Twain said that he was "startled" when Fowler told him that his skull bore a cavity that "represented the total absence of a sense of humor! . . . I was hurt, humiliated, resentful." Three months later, after he was certain that Fowler had forgotten him, Twain visited again—this time as himself, the famous author—and voilâ, "the cavity was gone, and in its place was a Mount Everest—figuratively speaking—31,000 feet high, the loftiest bump of humor he had ever encountered in his life-long experience!"¹⁶

As a scheme for linking personality traits to brain anatomy, phrenology failed spectacularly, but its foundational notion—that certain types of mental phenomena are localized in the brain—is broadly correct and informs several important clinical practices today. It has become increasingly common in presurgical planning for neurosurgeons to use fMRI to map the language and motor regions of the brain to minimize damage to these functionally important areas while removing a tumor, blood clot, or epileptic tissue. Brain mapping has also been invaluable in pinpointing some of the central sites of defective activity in patients with severe, chronic depression or obsessive-compulsive disorder, allowing for optimal placement of therapeutic electrodes to stimulate the affected areas, a technique called deep brain stimulation. It is also used to ascertain stroke damage, to follow the course of Alzheimer's disease and epilepsy, and to determine brain maturity. Scientists hope that fMRI will improve the treatment of comatose patients by allowing doctors to directly measure levels of consciousness.¹⁷

THE idea that a specific area in the brain is solely responsible for enabling a given mental function may be intuitively appealing, but in reality it is rarely the case. Mental activities do not map neatly onto discrete brain regions. For example, Broca's area—once believed to be the brain's one and only language-production center—has been

discovered not to have exclusive rights over this capacity. More precisely, it can be thought of as one of the key nodes, or convergence centers, for pathways that process language. Nor is there one designated site in charge of speech comprehension; it too relies on patterns of connectivity across multiple brain regions. Although neuroscientists regard a few cortical regions as being highly specialized for particular operations—such as the perception of faces, places, and body parts, ascribing mental states to others (“theory of mind”), and processing visually presented words—most neural real estate is zoned for mixed-use development. Furthermore, the brain can sometimes reorganize itself after injury so that other areas take over the functions of damaged regions, especially when the injury occurs early in life. For example, the “visual cortex” in blind people can be used to perceive touch, such as the feel of Braille letters.¹⁸

Take the variety of functions performed by the amygdala. This is a small region, one on each side of the brain, located within the temporal lobe at the point of intersection of a line that goes through the eye and another through the ear. In media reports, the amygdala has become almost synonymous with the emotional state of fearfulness. As it turns out, however, the amygdala handles much more than fear. “If I put you in a state of fear, your amygdala lights up,” says imaging expert Russell Poldrack. “But that doesn’t mean that every time your amygdala lights up you are experiencing fear. Almost every brain area lights up under lots of different states.” Indeed, the amygdala becomes more activated during feelings of happiness, anger, and even sexual arousal—at least in women. (Perhaps, then, the women whose amygdalae lit up while looking at images of Mitt Romney circa 2007 in the swing-voter study found him appealing rather than threatening.)¹⁹

The amygdala also mediates responses to things that are unexpected, novel, unfamiliar, or exciting. This probably explains its increased activation when men look at pictures of a Ferrari 360 Modena. The amygdala reacts to photos of faces with menacing ex-

pressions, but also to photos of friendly, unfamiliar faces. If fearful faces are expected and happy faces are unexpected, the amygdala will respond more strongly to the happy faces. The amygdala also helps register the personal relevance of a stimulus at any given moment. One study, for example, revealed that hungry subjects manifested more robust amygdala responses to pictures of food than did their nonhungry counterparts.²⁰

This example illustrates the knotty problem of reverse inference, a common practice wherein investigators reason backward from neural activation to subjective experience.²¹ The difficulty with reverse inference is that specific brain structures rarely perform single tasks, so one-to-one mapping between a given region and a particular mental state is nearly impossible. In short, we can’t glibly reason backward from brain activations to mental functions. When Jeffrey Goldberg views a picture of Mahmoud Ahmadinejad and his ventral striatum lights up like a menorah, some investigators might think, “Well, we know that the ventral striatum is involved with processing reward, so this subject, with his activated ventral striatum, is experiencing positive feelings for the dictator.” This interpretation works only if the ventral striatum’s assignment is exclusively to process the experience of pleasure. But that’s not the case; novelty can also stimulate the ventral striatum.

To be fair, there is nothing wrong with the reverse-inference approach as long as the investigative buck doesn’t stop there. Indeed, the approach frequently offers a valuable starting point for generating fruitful hypotheses that can later be tested in systematic experiments. Unfortunately, the studies that tend to attract media attention are the ones trafficking in conclusions based solely on reverse inference. Thus, in the swing-voter study, researchers concluded that candidate John Edwards provoked feelings of disgust among some of the undecided. Why? Because his picture spurred activity in the insula, a prune-sized area of the cortex nestled under the juncture of the temporal and frontal lobes. True, the region plays a role in

mediating the experience of visceral disgust, but it too does much more. The insula has at least ten anatomical subunits, each with its own population of neurons and specialized functions that play roles in a broad range of experiences, such as trust, sudden insight, empathy, uncertainty, aversion, and disbelief. Researchers have also linked the left hemispheric insula to the quality of orgasm in women (like most brain areas, the insula is a paired structure, with one in each hemisphere). Most striking, the insula helps mediate awareness of bodily sensations. It integrates visceral states, such as pain, hunger, thirst, and temperature, and thereby contributes to the conscious experience of emotion.²²

So, did the amygdala-energized swing voters viewing Mitt Romney experience anxiety or a sense of novelty? Or something else? Were the insulae of undecided voters signaling an attraction to John Edwards or a repulsion? Is Jeffrey Goldberg pro-Israel or a closet Shiite? The most daring neuropundits, as journalist Daniel Engber has labeled them, are not humbled by these complexities. They seem to regard a brain image almost as a kind of high-tech Rorschach inkblot test. But reading into largely ambiguous patterns what one wants to see is a serious breach of the fundamental test of a good theory: falsifiability, the ability of a hypothesis to be disproved by a test or observation.²³

★ WHENEVER a newspaper headline proclaims, "Brain Scans Show . . .," the reader should entertain some healthy skepticism. There are several reasons why.

First, brain scans rarely allow investigators to conclude that structure X "causes" function Y. This is not what fMRI alone can demonstrate. Instead, it at best indicates only correlation—that is, which parts of the brain are active when a person participates in a particular task—not which brain area is causing a particular psychological operation or behavior. For example, some teenagers' brains show increased activity in regions associated with aggressive tendencies

when they play violent video games. But we cannot conclude from this observation alone that violent videos cause violent behavior. That inference would be unfounded. Perhaps teens with known aggressive tendencies also enjoy playing those games.²⁴ Or maybe teens whose parents are generally inattentive to their doings—including playing violent games—set the stage for all manner of misbehaving by their kids. And what about teens who are otherwise well behaved but just happen to be stimulated by these videos when they are playing them?

Second, the subtraction technique used in most fMRI experiments is not necessarily well suited to the question being asked. Recall that the assumption behind subtraction is that two mental tasks' conditions differ in only a single cognitive process. However, most mental operations that seem like a unified task are composed of many smaller components. Consider what is involved in performing a simple arithmetic problem. First, subjects must recognize a visually presented number. Second, they must register its numerical magnitude. Third, they must compute the correct solution to the problem at hand. Because these operations are not performed by the same brain region, researchers must "decompose" the math task into the neural correlates associated with each step.²⁵

Now, if decomposing an arithmetic task is complicated, imagine the difficulty of breaking down even more complex states, such as attitudes and emotions. Is it even possible to translate a complex pattern of neural activity into a simple interpretation, as the swing-voter team did? Was it reasonable to infer that swing voters were "trying to suppress unwanted emotions" while they looked at Hillary Clinton? There is good reason to be dubious.

Third, although neuroimaging has deepened our knowledge of brain anatomy and function, its popular application tends to reinforce the misbegotten notion of the brain as a repository of discrete modules that control distinct capacities to think and feel. This is not the case, of course. Studies that suggest a "brain spot for X" are

typically misleading because mental functions are rarely localized to one place in the brain. There is a babel of crosstalk among numerous brain regions as they are strung together in specialized neural circuits that work in parallel to process thoughts and feelings. Almost nothing is static in the brain. The organ continuously rewires itself in response to experience and learning by altering the strength of its connections countless times every second. Neuroscientists now think of the brain as an ever-changing ecosystem crackling with electrochemical energy from which our thoughts, emotions, and intentions arise, rather than a collection of blinking neural islands.

The heavily interconnected aspect of the brain explains why researchers are increasingly moving away from the kind of regional brain mapping used in the swing-voter study and embracing an fMRI technique called pattern analysis. Pattern analysis, also called decoding, mathematically examines the brain's extensive interconnections. Investigators first gather data on the "correct" brain response—say, fear—as it is observed in subjects asked to view frightening things. Once the computer program has been "taught" to determine what the subject is looking at, researchers can later infer what he or she is watching just by analyzing activity across the brain. Thus, instead of inferring that a photo of Mitt Romney induces anxiety, researchers might collect patterns of brain activity evoked by known anxiety inducers (photos of spiders, snakes, and hypodermic needles, perhaps) and then see whether the pattern that Romney elicits is a statistical match with the pattern for anxiety derived from the inducers.²⁶

A fourth caveat to keep in mind when one is interpreting brain scans is the importance of experimental design. The way in which investigators design their task can exert a big impact on the responses they obtain. This lesson is powerfully illustrated by a set of experiments designed to explore how teenagers process emotional information compared with adults. In a 1999 fMRI study, Harvard University researchers asked normal teens to look at a series of black-and-white photos depicting frightened faces. They mischaracterized

one in four photos, sometimes seeing anger, surprise, confusion, or even happiness. Regardless of whether they correctly identified the emotion as fear, subjects displayed significant amygdala activation. A subsequent study showed that adults made few errors when identifying fear. "I think this has important implications in terms of . . . trying to inhibit their own gut responses," said one of the researchers.²⁷ The studies were taken to mean that teens were innately deficient in their ability to interpret the emotions of others in social situations and, therefore, more prone to impulsive violence than adults. In Chapter 5, we show how defense lawyers have employed these kinds of findings to argue that teens are less criminally responsible for murder than are adults.

It turned out, however, that teens might not be so bad at detecting fear in others after all. When one of the original Harvard investigators, Abigail Baird, conducted additional trials with new photographs of faces, she obtained different results. She switched the photos from the dated black-and-white shots of people who resembled bad actors in a B-grade horror film (the ones used in the first two studies) to color photos of more contemporary-looking people. When she did this, Baird's adolescents provided correct responses almost 100 percent of the time. "They were simply more engaged by more contemporary peer photos in color," Baird concluded. "They did well if they cared."²⁸ Whatever elements of the new stimuli lead to the differences in teens' responses, the point is that a seemingly trivial aspect of the experimental task—aspects of the photo unrelated to fear—gave rise to an entirely different conclusion about the ability of teens to identify the facial expression of fear.

A fifth caveat stems from the fact that fMRI is an indirect method. Contrary to popular belief, imaging does not measure action of brain cells *per se*. True, most neuroscientists regard BOLD as a reasonable proxy for changes in neuronal activation, but the link between blood flow and neural activity is not straightforward. For example, there is a delay of at least two to five seconds between

activation of neurons and the increase in oxygen-rich blood flowing to them. Thus information about mental processes occurring in the brain may be out of sync with the neural activity actually producing it, and therefore, rapid fluctuations of neural activity may go undetected. To compensate for lost data, researchers use electroencephalography (EEG), which detects electrical activity on the surface of the brain very rapidly, producing a data point every four milliseconds (thousandths of a second) or less, thousands of times faster than fMRI creates a single brain image.²⁹

But even when neural activity is detected, it is not always clear what is going on in the hot spot. When brain cells, or neurons, fire, they send electrical impulses down a long, tapered filament called an axon. When the impulse reaches the end of the axon, neurotransmitters, or chemical messengers, are released into a synapse—a tiny gap between the axon and another neuron—and influence the action of the receiving neuron. The chemical message released by the firing neuron often stimulates the receiving neuron, making it fire. But although some neurons are excitatory, ramping up the activity of certain brain regions, others are inhibitory, tamping them down. As a consequence, regions of inhibitory neurons may “light up” brightly in the final scan, even though they are working to depress rather than stimulate activity elsewhere in the brain.

Alternatively, spots may be dark when they should show activity. This might happen when the voxel, at three cubic millimeters, is still too big a unit of spatial resolution to capture activity occurring on a smaller scale, such as in a tiny cluster of neurons that, despite its size, happens to perform a critical function. These small clusters may or may not appear on the final scan. What’s more, a given region may appear deceptively less engaged with a task, on the basis of activation levels, than it really is. In fact, the region could be very important to enabling the task but appear less active because the brain becomes more efficient at tasks it performs repeatedly or automatically. Such a “practice-suppression” effect means that the blood oxy-

gen level required to perform the task is lower than it would be for someone who has never before performed it. It’s essential, therefore, to take practice effects into account when one is gauging the relative contribution of various regions.³⁰

Last, it is important to keep in mind that before the final data even “reach” the voxel, analysts must deploy statistical approaches to extract meaningful information from the noise. This is where, as imaging expert Hal Pashler puts it, this “hellishly complicated [process] creates great opportunity for inadvertent mischief.” The mischief is not intentional, of course. It is due partly to the fact that analytic methods are constantly evolving and can vary from lab to lab. Such lack of standardization, not unexpected in a rapidly growing field, has implications for reproducing the work of others, collaborations across labs, and building on the work of other teams.³¹

Additional mischief involves statistical error, not the neuroimaging process itself. When researchers run large numbers of statistical tests simultaneously on the BOLD signals, some of those tests are bound to turn up “statistically significant” just as a result of chance. In other words, these results will suggest, by mistake, that the brain is more active when a subject performs a task when, in fact, that part of the brain was actually not recruited. To make this point in a dramatic way, neuroscientist Craig Bennett set out to demonstrate how brain scans can produce (literally) fishy results. Bennett and his team purchased a dead Atlantic salmon from a store, placed their cooperative subject in a brain scanner, “showed” it photographs of people in various social situations, and “asked” the salmon to guess what the people were feeling. Bennett’s team found what they were looking for: a tiny area in the salmon’s brain flared to life in response to the task. Of course, this island of brain activation was merely a statistical artifact. Bennett and his fellow researchers had deliberately computed so many subtractions that chance alone caused a few of the results to become statistically significant despite their being entirely spurious.³² The salmon “study,” which won a 2012 Ig Nobel

Prize (for work “that makes people laugh, then think”), illustrates that decisions in data analysis can impact the reliability of fMRI results.

It is relatively easy to correct for the problem of false positives by using standard statistical tests. But there are plenty of other pitfalls. In a “bombshell” paper, as a fellow neuroscientist put it, MIT graduate student Edward Vul concluded that something was deeply wrong with how many brain-imaging researchers were analyzing their data.³³ Vul had become suspicious when he came across what he described as “impossibly high” estimated relationships between psychological states and the activation of various brain regions. Vul was skeptical, for example, of a 2005 study that purported to find a near-perfect relationship of .96 (with 1.0 being the maximum) between proneness to anxiety in response to angry speech and activity in the right cuneus, an area at the back of the brain believed to be involved in impulse control. Another implausible finding came from a 2006 study that reported a correlation of .88 between self-reported partner jealousy over emotional infidelity and activation in the insula.

In poring over the original articles, Vul and his collaborator Hal Pashler realized that these researchers were drawing inferences from a biased sample of results. When investigators look for correlations between stimuli and brain activations, they often cast a wide net. This leads them, first, to tiny regions of highest activation. Once they have homed in on those small regions, researchers compute the correlation between the psychological state in question and brain activation. In doing so, they are inadvertently capitalizing on chance fluctuations in the data set that are unlikely to hold up in later studies.³⁴

Many aspects of Vul’s critique are technical, but his basic point is easy to grasp: If you search a huge set of data—in this case, tens of thousands of voxels—for associations that are statistically significant and then do more analyses on only those associations, you are almost guaranteed to find something “good.” (To avoid this mistake, the second analysis must be truly independent of the first one.) This error is

known variously as the “circular analysis problem,” the “nonindependence problem,” or, more colloquially, “double-dipping.”

These are not arcane matters. The strength of the correlation tells future researchers how to design their investigations; it alerts them not only to where they ought to look, but also to where they shouldn’t. When the article of Vul and his colleagues detonated in the research community, some of the criticized authors pushed back. Countercharges, rebuttals, and rebutted rebuttals flew around the web.³⁵ Yet, in the end, most scientists agreed that the statistical problem Vul identified had troubling implications, and that caution was indeed warranted in moving forward.

FUNCTIONAL brain scanning plays a vital role in the nascent study of human brain-behavior relationships. To fully appreciate its virtues and limits, one must keep three general points in mind. First, even the most superficial insight into how brain scans are generated should dispel the assumption of naive realism—the commonsense theory of perception we discussed in the Introduction. Naive realism, as philosophers define it, is an intuition held by most people that the view of the world that we derive from our senses is to be taken at face value. Scans show how profoundly we will be misled if we view them through the lens of naive realism. They are not raw snapshots of the brain’s real-time functioning. They are highly processed representations of the brain’s activity.³⁶

University of Montreal researcher Eric Racine refers to a corollary assumption termed “neurorealism.” A first cousin of naive realism, neurorealism denotes the misbegotten propensity to regard brain images as inherently more “real” or valid than other types of behavioral data. Neuroeconomist Paul Zak has described his work on the neurobiology of trust: A brain scan “lets me embrace words like ‘morality’ or ‘love’ or ‘compassion’ in a non-squishy way. These are real things.” So is the fact that we like to talk about ourselves—but it is not really news. In discussing the psychological impact of combat, a

researcher quoted as saying that brain imaging tells us that post-traumatic stress disorder (PTSD) is a "real disorder" (as if we did not know this). We coined the term "neuroredundancy" as a "neurologism" to denote things we already knew without brain scanning.³⁷

A second crucial consideration is the design of the experiments. The kind of task that researchers put before their subjects—whether it involves showing them photos of loved ones or fails to account for color in photos of scared people—can have a profound influence on the neural correlates that appear in the final scan. Even seemingly trivial differences in the nature of the experimental setup can generate huge differences in brain-imaging findings. When it comes to interpreting the results of imaging studies, context is everything.

Third, a healthy sense of caution should be aroused when one hears of a study in which particular psychological explanations were derived from brain activity. There is a big difference between an investigator reporting that "region A exhibited enhanced activation when subjects viewed candidate B but not candidates C and D"—the accurate interpretation—and the less cautious conclusion that "the activity in region A means that voters prefer candidate B over others," or, worse still, "the activity in region A means that voters prefer candidate B over others because they find him to be sexier [or friendlier, more appealing, smarter, or whatever]."

BEARING these caveats in mind may help to restrain premature enthusiasm regarding the promise of brain imaging. When fMRI, PET, and other brain-imaging technologies first gained widespread currency in the heady days of the 1980s and 1990s, scores of scientists—perhaps not heeding these cautions sufficiently—confidently predicted a revolution in our understanding of mental illnesses, addiction, emotions, and personality. The scientific potential of what former president George H. W. Bush anointed "The Decade of the Brain" on July 17, 1990, seemed almost limitless. The fields of neuroscience, psychology, and psychiatry, many sensed, were on the brink of a new paradigm.³⁸

Forecasts from the field's most esteemed leaders were expansive. In her 1984 book *The Broken Brain—The Biological Revolution in Psychiatry*, psychiatrist Nancy C. Andreasen, who went on to receive the National Medal of Science from President Clinton, predicted that "as they improve and become more accurate, these imaging techniques and other laboratory tests for mental illness will become part of standard medical practice during the coming years, thereby improving the precision of diagnosis and assisting in the search for causes."³⁹ Two years later, Herbert Pardes, then head of the National Institute of Mental Health, wrote that "neuroscience is offering not only new information but startling new technologies and approaches. . . . While much in the way of clinical implications from brain research is promise, there is an expectation of great change over the next ten to twenty years."⁴⁰

Comparing the views of the current head of the National Institute of Mental Health, psychiatrist Thomas Insel, with those of Herbert Pardes over two decades ago, is instructive. As Insel observed in a sobering 2009 article, there is no evidence that the past two decades of advances in neuroscience have born witness to decreases in mental disorders' prevalence or to any impact on patient life span.⁴¹ The failure of brain-imaging techniques to have yet made major inroads into the causes and treatment of mental illness offers a necessary reminder for modesty in our expectations.

Yet such modesty is not evident in all quarters. In a worrisome recent development, a popular nationwide outfit called Amen Clinics promises patients that it can diagnose and treat depression, anxiety, and attention-deficit hyperactivity disorder using brain scans. Its founder, psychiatrist Daniel Amen, oversees an empire that includes book publishing, television shows, and a line of nutritional supplements. Single photon emission computed tomography, SPECT, a nuclear-imaging technique that measures blood flow, is the type of scan favored by Amen. His clinics charge over three thousand dollars for an assessment and, according to the *Washington Post*, grossed about

\$20 million in 2011. Despite near-universal agreement among psychiatrists and psychologists that scans cannot presently be used to diagnose mental illness, Amen insists that “it will soon be malpractice to not use imaging in complicated cases,” as he told a symposium at an American Psychiatric Association meeting.⁴²

Brain-imaging experts, however, shy away from such expansive claims. They are well versed in the conceptual limits of inferring mental states from biological indicators. They readily understand that the prowess of fMRI, a remarkable technology that is still relatively young and destined to evolve, is best demonstrated in the cognitive or affective neuroscience lab.⁴³ The danger comes when scanning leaves the experimental realm for socially consequential domains such as law and business, where much-needed interpretive restraint often gives way to extravagant claims about what brain scans can tell us about the mind.

Nowhere is this truer than in the nascent field of neuromarketing, our next topic, where brain science meets advertising hype. Savvy neuroentrepreneurs are selling brain-scanning and other techniques to corporate clients with the promise of unlocking the mysteries of consumer buying behavior. Using brain science to enlarge on valid insights into consumer behavior is one thing—indeed, it is a serious pursuit by a growing cadre of academics—but the most brazen promoters of neuromarketing may be pulling off little more than a brain scam.

2

THE BUYOLOGIST IS IN

The Rise of Neuromarketing

“THE ULTIMATE NO-BULLSHIT ZONE.” This is how globe-trotting Danish branding expert Martin Lindstrom refers to the human brain. “Our truest selves react to stimuli at a level far deeper than conscious thought,” he writes, estimating that a whopping 90 percent of our buying decisions take place at this level. As a result, “we can’t actually explain our preferences, or likely buying decisions, with any accuracy.” The author of the 2008 business bestseller *Buyology: Truth and Lies About Why We Buy* and one of *Time* magazine’s top 100 “Scientists and Thinkers,” Lindstrom advises marketers to cut out the middlemen—the buyers themselves—and ask their brains directly: Will you buy our product? Forget focus groups and questionnaires. The brain is the route to the heart’s desire.¹

Lindstrom is a high-profile member of an upstart generation of Mad Men known as neuromarketers. They apply the tools of neuroscience, such as fMRI and brain-wave technologies, to learn how consumers’ brains instantly react to ads and products. It’s all in the service of answering elusive questions as old as advertising itself: What do customers want? What motivates them to buy? And how can I get them to buy *my* product? “Half the money I spend on

Epilogue

MIND OVER GRAY MATTER

BRAIN IMAGING, the iconic tool of neuroscience, finds itself at the eye of a perfect storm of seduction. Riding one current is the glamour of a sophisticated and exciting new technology. Borne aloft on another is the brain itself, an organ of great moment and mystery. On a third front floats an overly simplified brain-to-behavior narrative, all rendered in stunning biological portraiture. It is easy to see how nonprofessionals, and an occasional expert, tossed by these powerful swells, can get swept away.

We wrote this book to serve as an anchor. Our project is not a critique of neuroscience or of its signature instrument, brain imaging. It is foremost an exposé of mindless neuroscience: the oversimplification, interpretive license, and premature application of brain science in the legal, commercial, clinical, and philosophical domains.¹ Secondarily but importantly, it is also a critique of the increasingly fashionable assumption that the brain is the most important level of analysis for understanding human behavior, and that the mind—the psychological products of brain activity—is more or less expendable.

We are unreserved champions of neurotechnological progress. We are certain that brain imaging techniques and other exciting

developments in neuroscience will further elucidate the relationship between the brain and the mind. We deeply admire the neuroscientists whose inquiries are yielding new discoveries and, perhaps soon, much-needed treatments. In the preceding chapters, however, we have tried to bring a circumspect view to real-world applications of neuroscience and to speculations about where insights gleaned from brain science may take our society. As we've seen, the illuminated brain cannot be trusted to offer an unfiltered view of the mind. Nor is it logical to regard behavior as beyond an individual's control simply because the associated neural mechanisms can be shown to be "in the brain."

Scans alone cannot tell us whether a person is a shameless liar, loyal to a product brand, compelled to use cocaine, or incapable of resisting an urge to kill. In fact, brain-derived data currently add little or nothing to the more ordinary sources of information we rely on to make those determinations; mostly, they are neuroredundant. At worst, neuroscientific information sometimes distort our ability to distinguish good explanations of psychological phenomena from bad ones.

We don't foresee neuroscience prompting a legal revolution. We agree with Stephen Morse that neuroscience will take its place along with other sciences that had their moment in the courtroom: Freudian analysis, behavioral psychology, the Chicago school of sociology, and the promise of genetic explanations. "The only thing different about neuroscience," according to Morse, "is that we have prettier pictures and it appears more scientific."² With the probable exception of Freudianism, the other disciplines have indeed made courtroom contributions to understanding why people act as they do. But they have hardly supplanted the bread-and-butter tools of the law, such as witness reports and cross-examination.

Neuroscientists cannot yet forge tight causal links between brain data and behavior. Until they can shed light on the measurable attributes that the law regards as important for culpability—who is

and who isn't responsive to reason—the rhetorical value of brain images will greatly outstrip their legal relevance. Within the law, ascriptions of criminal and moral responsibility do not hinge on what caused the bad behavior, but on whether wrongdoers possessed sufficient rational capacity to have been influenced by foreseeable consequences and to alter their behavior accordingly. This is why it has been said that "actions speak louder than images" in today's courtrooms, as well they ought to.³

Brain-based explanations for excessive appetites and for social behaviors that elide the crucial psychological, social, and cultural levels of analysis fall into the trap of neurocentrism. Therefore, they are virtually guaranteed to be impoverished explanations. Although scientists can describe human behavior on a number of different levels—the neuronal, the mental, the behavioral, the social—they are not close to bridging the yawning gap between the physical and psychological. The brain enables the mind and thus the person. But neuroscience cannot yet, if ever, fully explain how this happens.

As brain science continues to permeate the culture, neuroliteracy becomes ever more important. Neuroscience is one of the most important intellectual achievements of the past half century, but it is young and still getting its bearings. To demand the wrong things of brain science, to overpromise on what it can deliver, and to apply its technology prematurely will not only tarnish its credibility, it will also risk diverting crucial and limited resources, including federal funding for research, into less profitable ventures and blind alleys.

Skilled science journalists and bloggers, as well as neuroscientists and philosophers who write for the public and neuroethicists (a hybrid sort of scholar with training in both practical philosophy and science), now see part of their jobs as protecting the integrity of neuroscience from the growing legion of brain overclaimers.⁴ Responsible translators of neuroscience encourage a healthy skepticism, cautioning judges and policy makers in particular that brain activity elicited under narrow experimental conditions cannot currently yield

enough information to explain or predict human behavior in the real world, let alone inform the design of social policy.

Crucial lessons in neuroliteracy must also inculcate the importance of distinguishing the questions that neuroscience is equipped to answer from those that it is not. The job of neuroscience is to elucidate the brain mechanisms associated with mental phenomena, and when technical prowess is applied to the questions it can usefully address, the prospects for conceptual breakthroughs and clinical advances are bountiful. Asking the wrong questions of the brain, however, is at best a dead end and at worst a misappropriation of the mantle of science.

Recall neuroscientist Sam Harris, whom we cited early in this book. "The more we understand ourselves at the level of the brain," he wrote, "the more we will see that there are right and wrong answers to questions of human values."⁵ How so? Neuroscience can help answer questions about the neural processes involved in moral decision making, but it is not at all evident how such discoverable facts could ever constitute a prescription for how things *should* be. Surely, empirical facts can help us act more effectively on our values—if we want to rehabilitate prisoners more effectively, data on new therapies are essential. And neuroscience may be able to offer guidance in this regard. But whether we should jettison the practice of retribution on moral grounds is not a question that science, neuroscience included, can answer. Indeed, history is replete with feckless and at times bloody attempts at social engineering through biology. Then and now, it is a serious mistake to think that one can erect an ethical system based on science alone; philosophers call this confusion between "ought" and "is" the naturalistic fallacy.

Nonetheless, the great cultural authority of brain science renders it vulnerable to conscription in the service of one or another political or social agenda. The framing of addiction as a brain disease to attract more funding for research and better services for drug abusers might seem benign; in most cases it is surely well intentioned. But that per-

spective sorely misrepresents the multilayered nature of addiction and risks distracting clinicians from the most promising kinds of interventions. The same is true to some extent for many other psychological maladies (including psychopathy, the condition that likely afflicted murderer Brian Dugan), which, although surely rooted in brain dysfunction at some level, can be fully understood only by also accommodating the idiom of motives, feelings, thoughts, and decisions.

Likewise, invoking brain science as a rationale for negating blame and abandoning punishment practices is misguided. Neuroscience itself is not a threat to personhood. It will help explain how human agency works, but it will not explain it away. A strictly utilitarian model of justice—one in which we punish people solely because aversive stimuli make society work better, not because blame is truly deserved—has its merits and its shortcomings, depending on your view. But whether human beings who live in a material world can also be moral agents is not a question that brain science can resolve. Not unless, that is, investigators can show something truly spectacular: that people are *not* conscious beings whose actions flow from reasons and who are responsive to reason. True, we do not exert as much conscious control over our actions as we think we do, but this doesn't mean that we are powerless.

In 1996, author Tom Wolfe penned a widely cited essay, "Sorry, but Your Soul Just Died." Neuroscience, he wrote, was on "the threshold of a unified theory that will have an impact as powerful as that of Darwinism a hundred years ago."⁶ Almost two decades later, the excitement surrounding neuroscience continues to grow, as well it should. But the promise of a unified theory in the foreseeable future is an illusion. As with sociobiology and the genomic revolution—two valuable conceptual legacies of Darwinism—we should extract the wisdom neuroscience has to offer without asking it to explain all of human nature.

In 2011, science writer David Dobbs recounted a sobering encounter at a gathering of neuroscientists. He asked them, "Of what

we need to know to fully understand the brain, what percentage do we know now?"⁷ They all responded with figures in the single digits. This humbling estimate will improve with time, of course. Brain imaging will become more precise; new technologies are yet to be unveiled or even envisioned. Yet no matter how dazzling the fruits of inquiry or how clever the means by which they are obtained, it is our values that will guide us in implementing them for good or for bad. The danger lies in muddling those values under the pretense of following where neuroscience supposedly leads us.

To some neuroscientists and philosophers, you may be nothing more than your brain—and of course, without a brain there is no consciousness at all. But to you, you are a "self," and to others you are a person—a person whose brain affords, at once, the capacity for decisions, the ability to study how decisions happen, and the wisdom to weigh the responsibilities and freedoms that these decisions make possible.

ACKNOWLEDGMENTS

We are indebted to many colleagues for their knowledge and insight. Stephen Morse and Hank Greely were invaluable to us in writing chapters on legal matters. Peter Bandettini was the most patient tutor on fMRI that two nonexperts could ever have. Tamler Sommers helped guide us through the rocky shoals of philosophy. Mark Kleiman was a tireless reader of several drafts of the chapter on addiction.

We are fortunate to have superb colleagues who graciously shared their wisdom. They included Craig Bennett, Marc Blitz, Paul Bloom, Nancy Campbell, Christopher Chabris, David Courtwright, Franz Dill, Roger Dooley, Robert DuPont, Steven Erickson, Martha Farah, Nita Farahany, Nathan Greenslit, Stephan Hamann, Wray Herbert, Bryce Huebner, Steven Hyman, Jerome Jaffe, Adam Keiper, Adam Kolber, Annie Lang, Carl Marci, Lori Marino, Richard McNally, Barbara Mellers, Jonathan Moreno, Emily Murphy, Eric Nestler, Joshua Penrod, Steven Pinker, David Pizarro, Russell Poldrack, Anthony Pratkanis, Eric Racine, Richard Redding, Kevin Sabet, Charles Schuster, Roger Scruton, Francis X. Shen, Raymond Tallis, Carol