CHAPTER 2

HOW WE THINK ABOUT SEEING

Seeing is something we do, and we continually learn how to do it. It is now clear that modern visual technology is a part of that learning process. Seeing is changing. A widely cited 2006 study from the University of Rochester showed that playing video games improved both peripheral and central visual perception. In other words, playing visual games makes you see better. There are many such reports of improved hand-eye coordination. In 2010, another Rochester study showed that gamers make faster and more accurate decisions based on sensory perceptions. Lead author Daphne Bavelier (now at the University of Geneva) describes this as "probabilistic inference," meaning the kinds of decisions we make based on incomplete information, such as choices made while driving (Bavelier Lab). The point here is that we do not actually "see" with our eyes but with our brain. And we have learned that in turn by becoming able to see how the brain operates. What we see with the eyes, it turns out, is less like a photograph than it is like a rapidly drawn sketch. Seeing the world is not about how we see but about what we make of what we see. We put together an understanding of the world that makes sense from what we already know or think we know.

It has long been realized that we do not see exactly what there is to be seen. The ancient Greek architects of the Parthenon in Athens designed the sides of their columns with a slight outward curve (entasis) as they rose in order to convey the appearance of being perfectly straight. In the seventeenth century, Western science began to distinguish between biological sight, which sees what there is to see, and cultural judgment, which makes sense of it. The philosopher and natural scientist René Descartes pointed out that when we look at a work of art drawn in perspective, we perceive what is actually an oval as a circle. He interpreted this as evidence that judgment corrects the perception of sight. This understanding was the basis of modern observational science. Descartes moved the knowledge of the world from being derived from the classical thought of ancient Greece and Rome to what each person observes in his famous aphorism "I think, therefore I am" (Descartes 1637). Only the fact that we think indicates that we exist. Everything else must be doubted and tested.

Descartes used vision as his example. The ancient Greeks and Romans had two contradictory theories to explain vision. One said that the eye threw out rays to "touch" the things we see. The problem with this idea is that we can see very distant objects immediately: so, how does vision throw its rays so fast? Another theory said that objects emitted little copies of themselves that got smaller and smaller until they entered the eye. The problem here is that large objects can be seen close up and enormous objects like mountains can also be seen: how did the copies get small enough, quickly enough, to enter the eye? No one could solve these problems and they did not really try to do so because light was held to be divine and so not subject to human understanding.

Descartes believed that the existence of God was the only way to guarantee that our observations are not simply delusions or the ravings of the insane. So, he tested everything. In 1637, he produced a famous diagram showing how vision was mathematically possible; it is still shown in many art and visual culture classes today.

He showed light entering the eye as a set of geometric lines. He solved the question of how large objects can be seen by showing that the rays are refracted by the eye's lens and converge on the retina at the back of the eye. However, this is not seeing. The image produced on the retina was interpreted by what Descartes called the sense of judgment. The drawing represents judgment as an elderly judge, assessing what there is to be seen and coming to a decision about it. Vision was understood as a courtroom, in which the eye presents evidence for the judge to decide. (There was no jury, as in the French courts of the time.) Descartes's breakthrough not only helped us for the first time to understand how vision was possible. It also raised the importance of seeing to a new level as the key sense in modern science, which centers on the observed experiment.

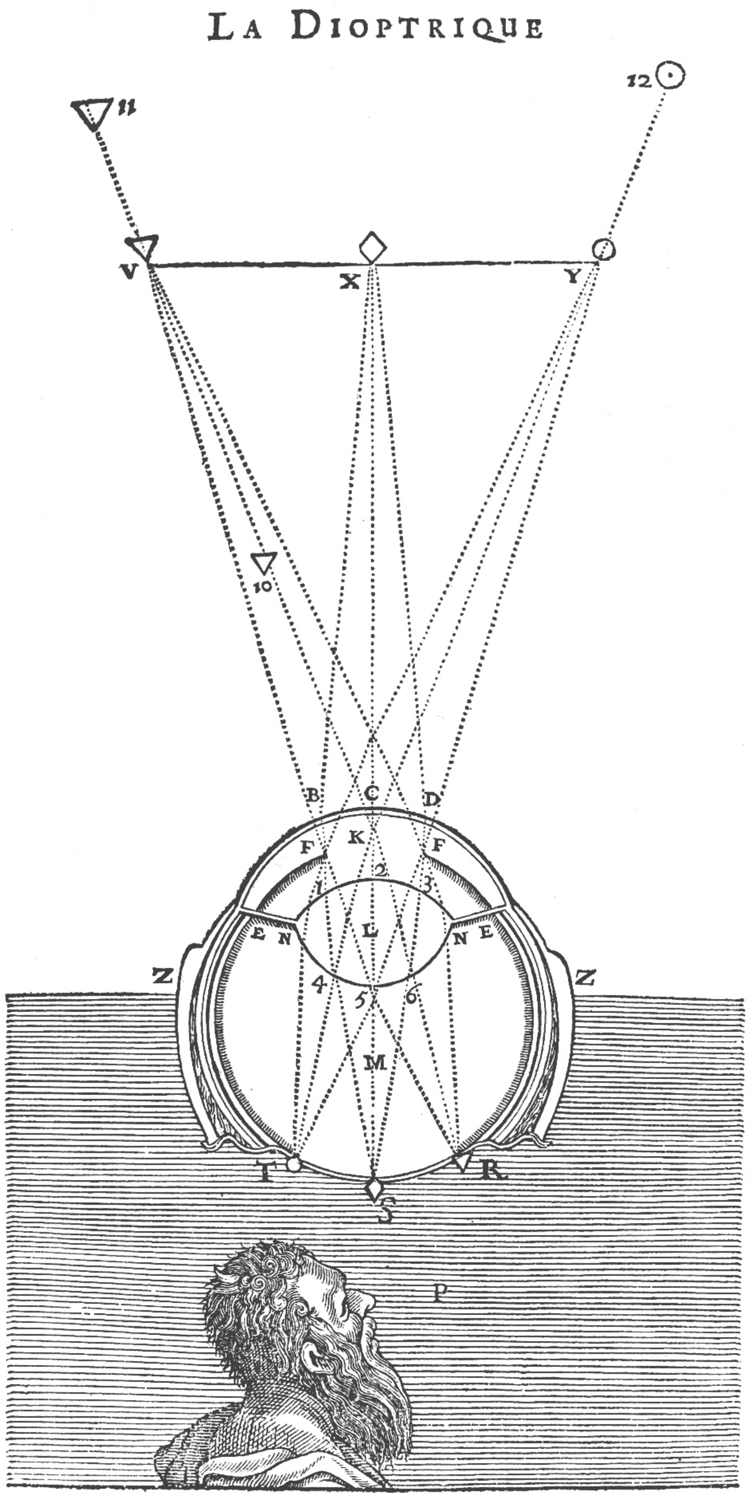

Figure 21. Descartes, "Vision," from *La Dioptrique*

In our own time, we are witnessing how neurology, a fast-developing part of biological science, sees the body and mind as integrated systems and people as communal, social beings connected by empathy. The metaphors here are not taken from the courtroom but from computer networks. It is a very different way to see ourselves and to think about seeing. According to this perspective, we learn how to become individuals as part of a wider community. This outcome is the intriguing result of the revolution in studying the brain, which many would consider to be the most individual organ of all, and in particular how humans and other primates see. My point is not that modern neuroscience is the final version of the "truth" and all other previous understandings have been shown to be wrong (although some neuroscience boosters do come close to saying this). Rather, as we shall see, neuroscience and its ways of visualizing the mind and human thought are becoming the vital visual metaphors of our time. It is our version of the truth, for better or worse.

VISUALIZING VISION

In the late 1990s, the psychologist Daniel Simons and his student Christopher Chabris devised what would become a famous experiment: a video test known as the "Invisible Gorilla" (Chabris and Simons 2010). Those who participated in the study were asked to watch a video and count the number of times the team wearing white passes a basketball while they play a team wearing black shirts. As this

Figure 22. Simons and Chabris, still from "Invisible Gorilla" video (1999)

simple action unfolded, a person wearing a gorilla suit walks across the court.

Roughly half of the people watching did not even notice the gorilla. They were concentrating on counting. Simons attributes this to what he calls "inattentional blindness," the inability to perceive outside information when concentrating on a task. Researchers had been aware of this phenomenon since the 1970s, as had magicians from time immemorial—"the sleight of the hand deceives the eye" because the magician distracts your attention. But it was the video that made the test so dramatic. You could test yourself and then watch the video again to see how obvious the gorilla then appears. Some people get very upset when they realize their failure.

This experiment built on the research of neuroscientists like Humberto Maturana from the 1970s. Maturana demonstrated that a frog, for example, sees very differently from the way we do. It perceives small, fast-moving objects, like the insects it eats, very clearly, while ignoring large, slow-moving things. Birds can perceive ultraviolet light invisible to humans, which allows them to see their own plumage differently than we do. However, even this seeing is not vision. Maturana stressed that living things change themselves because of their awareness of their interactions with the outside world, not just in the very long run described by evolution, but as a condition of day-to-day existence (Maturana 1980).

That is exactly what has happened in response to new media. When I show the "Invisible Gorilla" test video to students and others today, nearly everyone sees the gorilla. A population that has grown up with video games and touch screens sees things differently. Simons himself has found that when you show the video to experienced basketball players, the number seeing the gorilla jumps to about 70 percent. Simons carried out a more recent study demonstrating that some people today do not see the gorilla, based on a small sample of 64 people. Of these only 41 were previously unaware of his video. In this group, 18 did not see the gorilla, well under 50 percent. My sample group is larger in size, and compiled over several years, although not conducted as a scientific study. Perhaps my countersample, drawn as it is from participants in visual culture classes, is just more visually aware.

The capacities of the human body obviously cannot have evolved in such a short space of time. Rather, the change comes in the way we make use of visual information. In the age of industrial work, concentrating on a specific activity and ignoring distractions was highly desirable. From academic research by a student in a library to the adjustment of a machine by a factory worker, attention needed to be focused. Today, we prioritize the ability to keep in touch with multiple channels of information—*multitasking* is the popular term. As I write this book, people are sending me e-mails and text messages to which they expect prompt replies, regardless of what I am doing. Formerly, we were trained to concentrate on one task, meaning we might not see the gorilla, and mostly, though not exclusively, we did not. Now we are trained to pay attention to distractions and mostly, though not exclusively, we do. Neuroscience has changed the way that vision is understood. However, there's still noticeable room for interpreting that change.

SEEING THE BRAIN

Let's begin with how we can now "see" the brain in action. With the invention of new forms of medical imaging, especially magnetic resonance imaging (MRI) in 1977, it became possible to make "pictures" of the brain at work. Of course, no light is involved, and no drawing or other representational work is done. The magnetic field created by the machine excites hydrogen atoms in the brain (or whatever body part is being examined). As a result, they emit a radio frequency that is detected by the machine and converted into images. It is possible to imagine a species that could hear those frequencies and detect what is wrong (or not) with the person being scanned. Humans need to see something.

An MRI scan is actually an exercise in the history of media. Magnetism and its relation to electricity was a fascination of nineteenth-century science. The Scottish scientist James Clerk Maxwell demonstrated that light itself is a form of electromagnetism and he calculated its speed. He was also able to take the first partially successful color photograph in 1861. It was the study of electromagnetism that later led Albert Einstein to his theory of relativity and also made possible the invention of radio, the first mass medium. First used to communicate with ships, radio became a popular format in the 1920s. Now, however, it is our own bodies that are the transmitters. As we cannot interpret these waves unaided, the MRI machine converts them into visual form. These are not images, however, in the sense that they are representations of something seen. The scanned organ remains inside the body, unseen by anyone or anything. Like any other picture produced by a computer, MRI scans are computations, not images. Neuroscientists use a particular process called functional magnetic resonance imaging (fMRI) that allows them to "see" where the blood flows in the brain in response to specific stimuli. These pictures have developed a new mapping of the brain's functions, showing which structures in the brain "light up" when an action is carried out.

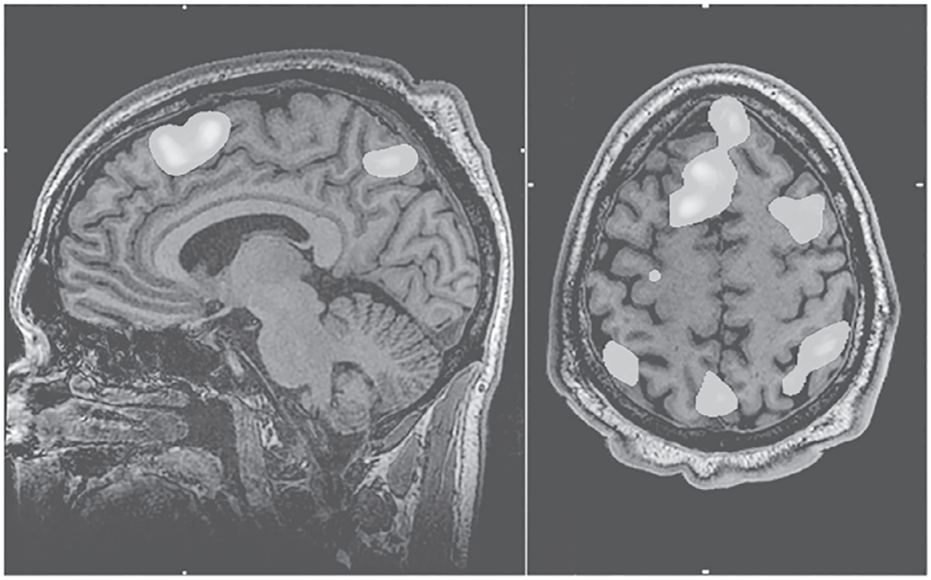

Figure 23. fMRI scan

These dramatic images seemed to make it clear that the brain has local specialties. The scan in Figure 23 shows a person using their memory, and it is clear that certain areas of the brain are in use and others are not.

At the same time that academics, artists, and activists were creating visual culture studies, scientists were transforming our understanding of vision itself by using these new techniques. In 1991, Daniel J. Felleman and David C. Van Essen published a now-classic analysis of visual function in primates, based on a study of macaque monkeys, because of their similarity to humans. In their summary, Felleman and Van Essen concluded that they had found

> 25 neocortical areas that are predominantly or exclusively visual in function, plus an additional 7 areas that we regard as visual-association areas on the basis of

> their extensive visual inputs. A total of 305 connections among these 32 visual and visual-association areas have been reported.[1]

In short, seeing is a very complex and interactive process. It does not, in fact, happen at a single place in the brain, as the first "lighting-up" images had suggested, but all over it in a rapid series of back-and-forth exchanges. Further, this interactivity between the visual zones of the brain and their associated areas happens at a series of ten to fourteen hierarchical levels. That is to say, seeing is not a definitive judgment, as we had once assumed, but a process of mental analysis that goes backward and forward between different areas of the brain. It takes a brain to see, not just a pair of eyes.

Felleman and Van Essen created a diagram of vision in the era of digital computing. In this mapping, the neural pathways for each sense are distinct but are processed in parallel, as in a computer. Their understanding of vision shows it as a set of feedback loops. Their map looks quite unlike any earlier model of seeing.

The one point of overlap with Descartes's earlier drawing is that the retina is still included. It's right at the bottom, labelled RGC. What we call vision happens in the set of feedback and parallel processing that takes place between this point and the hippocampus (HC) at the top. Clearly, vision is not just a case of light entering the eyes and being judged, as it was for Descartes. It is a back-and-forth shuffle with twists and turns, creating a vibrant sense of rhythm in the image.

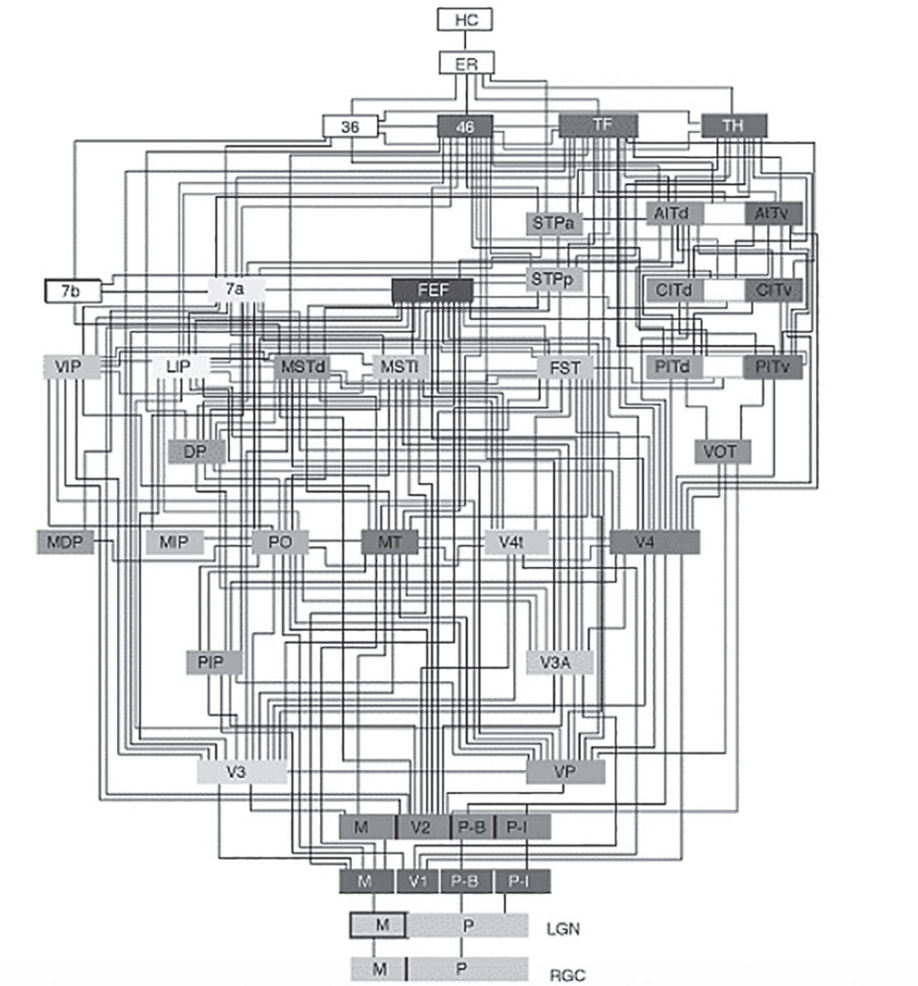

Figure 24. Felleman and Van Essen, "Hierarchy of Visual Areas"[2]

The diagram suggests a visual parallel with one of the great modernist paintings: Piet Mondrian's 1942 classic *Broadway Boogie Woogie*. Here the Dutch artist adapted his neoplastic aesthetic to express his sense of jazz-era New York. The painting conveys the dynamism of the grid city of Midtown, as well as its unexpected vibrations and drama that make the urban experience anything but

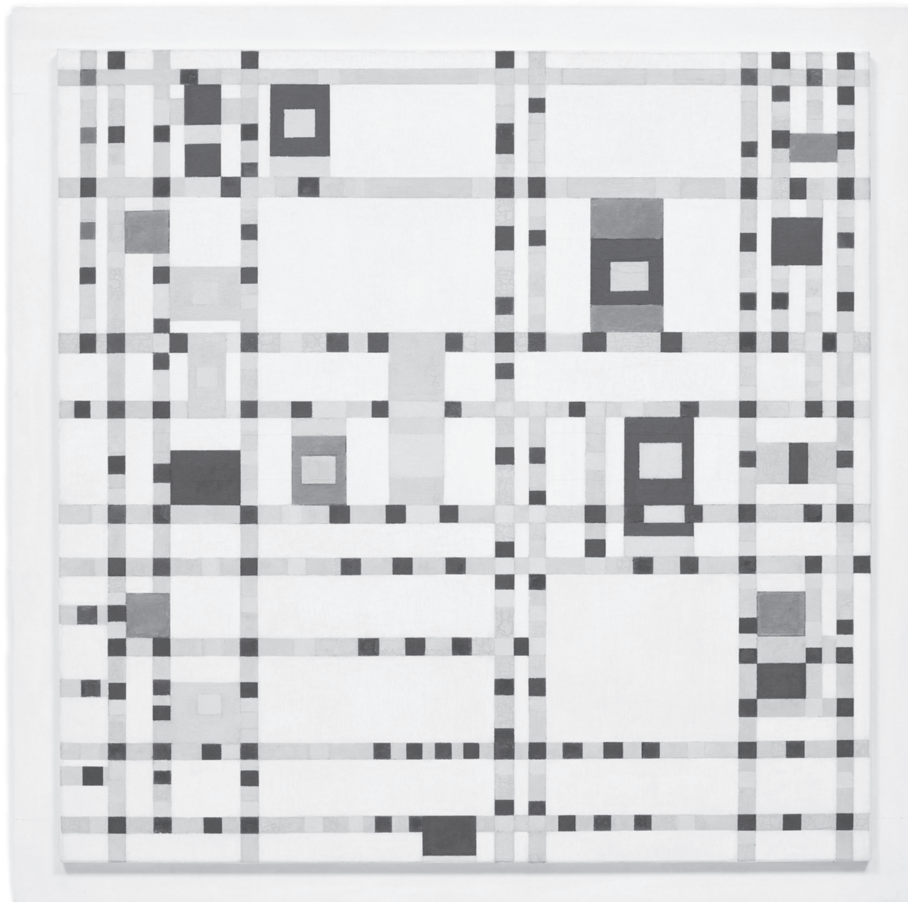

Figure 25. Mondrian, *Broadway Boogie Woogie*

regular. Boogie-woogie was an up-tempo piano form of the blues, often accompanied by a spectacular form of jazz dance that centers on a back-and-forth between the dance partners. In the painting, the combination of repeated bass rhythm and staccato dance movements evocatively conveyed the affect of the machine age. If we take the liberty of making a formal comparison between this painting and the 1991 depiction of vision (recognizing that they come from very different contexts), we can see in both cases how vision has gone from being the single decision imagined by Descartes to the dynamic experience of the modern, machine-based city, conveyed by its flickering lights, back-and-forth journeys, and infectious music.

In his analysis of this depiction of visual processing, the celebrated neuroscientist V. S. Ramachandran highlights the importance of the feedback between stages:

> Note especially that there are at least as many fibers (actually many more!) coming back from each stage of processing to an earlier stage as there are fibers going forward from each area into the next area higher up the hierarchy. The classical notion of vision as a stage-by-stage sequential analysis of the image, with increasing sophistication as you go along, is demolished by the existence of so much feedback.[3]

It is not just up to the judge. To extend my metaphor, it's the interplay between city residents on the street, the way the dancer responds to the piano, the feeling of being part of something. The information goes back and forth from one level to the next, filling in pieces as it goes. So, it is (for now) settled: seeing is something we do, rather than something that just happens naturally. More precisely, we need to set aside the persistent notion that an image gets relayed from the retina to the brain—the retina is itself part of the brain where, in Ramachandran's words, "the rays of light are converted into neural impulses."[4] From that point, information is distributed and processed in a series of parallel steps that continually reinforce the other layers. What we used to call an image is now known to be a computation, even in the brain.

In the current model, we don't even "look" to see. In Descartes's version, the judgment presided over seeing, making it seem a very deliberate process. In fact, there are three kinds of unconscious eye movement involved in seeing. Convergence movements direct both eyes to the same place. Pursuit movements track moving objects. Within these overall ways of seeing, research has shown that the building block of close seeing is the *saccade*, a spontaneous scanning by the eyes that move from one point to the next. Saccades are very rapid. Saccades can be set off voluntarily, as when we direct our eyes at a painting. Or they can be involuntary, in response to a moving object, a noise, or any other unexpected event. So, the old idea of a single gaze or look, fixing people or objects under its stare, has to be modified. Our eyes are always busy, boogieing back and forth. The resulting mental "image" that we "see" remains stable because the brain computes it in that way.

The saccades produce data points that allow us to make calculations, such as how to pick up a coffee cup. We should think of looking as a whole as a form of doing, or a performance, as we called it in Chapter 1. We make a world in which the way we look makes sense and enables the actions we want to perform. And it is also a form of computing because we use that model to calculate how to be active in that world. So, if we're trying to count basketball passes, we can easily ignore the gorilla in the famous video described above because it has nothing to do with our counting action. If we are used to observing passes, though, we have more mental space to notice the gorilla. When the actions that we are trying to perform fail, we may simply have failed to throw an apple core into the trash. Or we may be totally disoriented by a hitherto unknown experience, such as an accident or disaster. We continuously rework these systems, absorb information, and change the way we perceive in order to account for it. In games that require hand-eye coordination, players are familiar with the experience of "getting their eye in," meaning that after the first few efforts, it seems to become easier to hit the ball. Our minds and bodies are continuously interacting, forming one system.

For all the discoveries of recent decades, the place of visual perception has nonetheless become less precise. Vision now seems more akin to the puzzling one does in front of a complex painting like *Las Meninas* (Chapter 1)—the moving back and forth to gain different effects, the homing in on certain details, the changing affect that often returns to the beginning point—than it does to the instant affirmation of photography. Eighteenth-century art theorists once proposed a theory of *papillotage*, meaning "flickering" or "blinking" vision. Painters like François Boucher aimed to create this sense of moving surface that would later come to influence the Impressionists. It accepted the constructed nature of vision. A blinking vision is aware of its effort to see, of the difference between it and what it is trying to see, and even the eyelid that comes between the two.

In his small painting *Diana Leaving Her Bath* (1742), Boucher sets off the sensuousness of the goddess of the hunt's naked body with a dramatically vibrant background of golds and greens, representing the natural world behind her. The painting creates a sense of tactility to heighten its allure. By contrast, neoclassical painting, which came into fashion in the late eighteenth century at the time of the French Revolution, stressed the primacy of drawing, using hard edges for figures and objects. If we glance at Jacques-Louis David's much-larger portrait of *Antoine-Laurent Lavoisier and His Wife* (1788), the difference is at once apparent.

There is certainly sexual intrigue here. But the crisp precision of David's lines and almost hallucinatory clarity

Figure 26. Boucher, *Diana Leaving Her Bath*

Figure 27. David, *Antoine-Laurent Lavoisier and His Wife*

of the figures and their outlines could not be more distinct from those of his predecessor Boucher. It was this approach that would be imported into photography as "correct" focus. This realism wanted to efface all flickering and insist that what we see is what is there and nothing else.

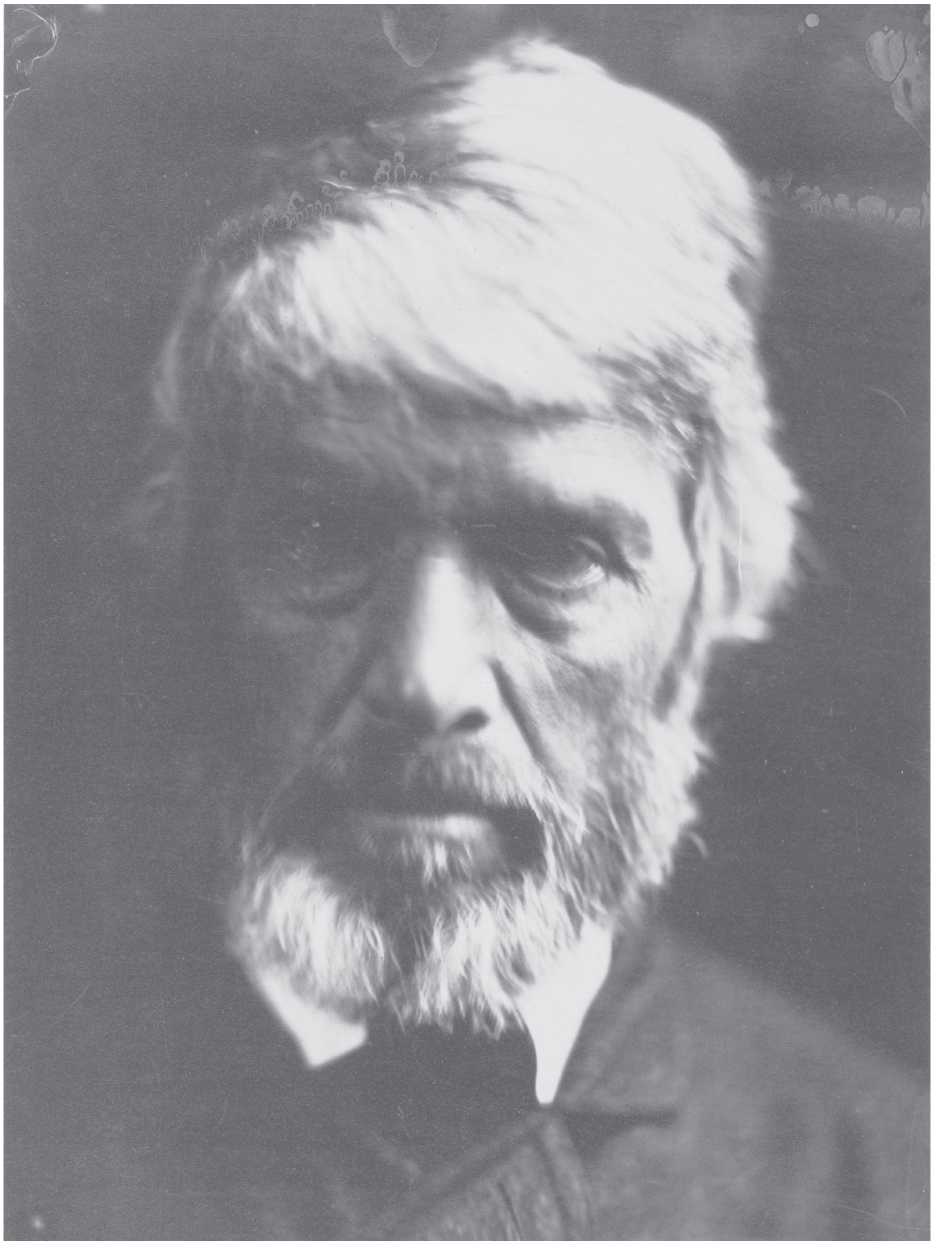

Figure 28. Cameron, *Thomas Carlyle*

Early photographers, like the British artist Julia Margaret Cameron, disputed this, demanding to know in a letter from 1864, “What is focus?” (Mavor 1999). Do we concentrate on producing sharp lines, or insight into the content of the photograph? In her portrait of Thomas Carlyle, for example, Cameron produced an image of the historian as mystic that is arguably closer to Carlyle's sense of himself as a "seer" (meaning both a prophet and a person who sees) than a sharp-edged image would have been.

The point is not that earlier periods correctly anticipated neuroscience or that the current ideas of the neuroscientists must be accepted as unquestioned truth. Rather, what these works of art reveal is that the current research fits into a well-established line of thought about vision, even if it is based on very different evidence and in a very different context. Indeed, it is noticeable that people today often put more trust in a less-than-perfect photograph or video that takes an effort to decipher than they do into a professionally finished work, because they suspect that the latter will have been manipulated.

W. J. T. Mitchell, often considered to be the founder of visual culture studies, was prompted by the rise of digital visual culture to propose "there are no visual media" (2005). By this apparently paradoxical statement, Mitchell meant that all media involve every sense and so it is inaccurate to describe a painting, which is made by an artist touching a canvas with brushes, as being nothing but visual. Now we can reinforce that interpretation with the understanding that perception is not a single action but a process carefully assembled within the brain. That work centers on what neuroscientists call body maps, our sense of where and who we are. Thus we always know what posture our body is in, even with our eyes closed. Think of how you brush off an insect that lands on you. You have to be able to coordinate a sense of yourself, which is now called proprioception—a sense of the relative positioning of all the different parts of your body—with the perception of an insect on your skin. Then you use a hand to swipe at the bug at sufficient speed to prevent it biting you but without inflicting excessive pain on yourself.

These maps do not always accord with the objective state of our bodies. To take extreme examples, a person withdrawing from addictive drugs may feel insects crawling all over their skin and swipe at them constantly to no effect. A person who has suffered amputation often feels a “phantom” limb where the actual member used to be. They know perfectly well it’s not there, but it hurts, itches, and otherwise calls attention to itself. By the same token, anorexic people may look at their image and see themselves as obese, rather than perceiving the very thin body that others would see. In this case, such anorexics have long been assumed to be distorting their own image, which is part of the clinical diagnosis. However, some “pro-ana” (proanorexia) websites and online communities share images of very thin people, including celebrities and models, as “thinspiration.” So, some people with anorexia (as the medical profession sees it) are as aware of their body map as everyone else. They just draw radically different conclusions from it (and, to be clear, not ones that I would personally support).

Some striking experiments have shown that these maps can be relearned. The neurologist Ramachandran performed a famous therapy with phantom-limb patients by rigging up a mirrored box, so that the surviving hand (for example) was reflected by mirrors in the place of the amputated one. Using this simple visual bypass, people have been able to redraw their body maps so that phantom limbs can be "moved" out of awkward positions, itches scratched and so on. It seems that all the redundancy in the visualizing system allows for the possibility of such relearning. Stroke victims and even people suffering from chronic pain have been able to benefit from this therapy. The visual information seems to "overwrite" the other information available. There is more processing space in the brain devoted to vision than to all the other senses combined, which might account for why this illusion is so irresistible.

In less dramatic cases, many optical illusions continue to "work" even if you know that they are not real. Think of the philosopher Wittgenstein's duck-rabbit, a drawing that can look like either a rabbit or a duck, or both, depending on how you see it. We might even see it as a collection of lines and shading. By extension, you might take an active choice to see only in one way, like the person with anorexia who sees their thinness as affirmative. Or you can succumb to the illusion and "move" the amputated limb that no longer exists.

It turns out, according to recent research, that this is because there are two "streams" of brain activity: one for perception and one for action (Nassi and Callaway 2009). Vision is a plural noun, it appears. One stream (perception) recognizes a friend. The other (action) reaches for the friend's hand and shakes it. There are now thought to be over eighty physical locations in the brain that process vision, connected by at least twelve parallel processing pathways. Despite all this input, key "attributes such as motion, shape and color must be computed from these sensory cues." That is to say, we don't "see" color, shape, or speed so much as work out what they must be. The brain is not a camera. It's a sketch pad. Intriguingly, the medieval theologian St. Thomas Aquinas also held that vision was not relevant to determining where an object was or what color it was, albeit for very different reasons. I take it from this that people have always known in some way that vision is constantly being learned and relearned.

Figure 29. Wittgenstein, "Rabbit and Duck"

THE MIRROR AND THE COMMUNITY

One of the most intriguing of all the insights from the new research is that we do indeed learn mostly from each other, rather than by ourselves, and that our brains are specifically designed for that purpose. Sense experience is not individual but common. This was the surprising result of a set of experiments on monkeys by Italian scientists. The Italian group was working on a project to analyze what nerve impulses were produced by a monkey reaching for a peanut. By chance, they discovered that the monkeys watching had exactly the same neurological response as the one actually grabbing the treat. The same part of the brain "lit up" in the watching monkeys as in those of the monkey doing the action. The result has been reproduced many times in many different contexts and it has led to some radical conclusions.

Neuroscientists have now discovered that we have mirror neurons whose function is to respond to others. Ramachandran calls them "'Gandhi neurons' because they blur the boundary between the self and others—not just metaphorically, but quite literally, since the neuron can't tell the difference."[5] The quality of empathy is, in the current metaphor, hard-wired. Vittorio Gallese, an Italian neuroscientist, has explored the implications of his experiment with the monkeys here described (2003). For Gallese, the brain is in effect a shared space with a "we" that is not a crowd of individuals but a collective formation. It is from this formation that individuals emerge, meaning that we move from the social to the individual. By this he means that we don't generalize what we learn on our own, but apply the general to the specific, because we all form a "theory of mind." Such a theory, well known to philosophers, is vital to human interaction because without it, no one could begin to imagine how others might act. Like all visual culture, this insight is rooted in the everyday. If, in a café, I see a person reach for a coffee cup, I believe she is going to drink because that conforms to my theory. Of course, she may throw the coffee over her spouse, but that would be an exception to which my theory would then have to adapt. Rather than being a distraction from reality, the imagination is key to our very understanding of how we exist in the world.

In short, mirror neurons do not only allow me to see the world from my point of view but to visualize it from the point of view of others. As Ramachandran explains, humans are exceptional because we develop so slowly, learning everything from basic motor skills to language *after* birth, unlike most animals: "Obviously we must gain some very large advantage from this costly, not to mention risky up-front investment and we do: It's called culture."[6] So, it would not be unreasonable to say that the collective theory of seeing we develop would be something that we can call visual culture. Certainly this culture varies from place to place and time to time. As part of the interfaced sensory apparatus of the brain, visual culture is not only visual (in the commonsense meaning of the term), but is also concerned with the entire body map. The final implication is that humans have developed in the dramatically rapid way that we have, not only by Darwin's natural selection but also through cultural development.

All of the experience we call vision is multiply processed, multiply analyzed, and subject to constant feedback from other areas of the body, rather than being a single, independent sense. Researchers now want to add many new senses, like the proprioception described earlier, to the classical five senses. For example, we feel hunger and thirst as strong needs without having to think about them and without external stimulation. But that is not to say, as some skeptical philosophers have done, that we all see differently or that we can make no use of the sensing that we do. Because seeing is a specific instance of our collective theory of mind, vision is a commons, meaning a shared resource that we can nonetheless make use of in ways that also suit our individual needs. And this is what visual culture seeks to debate, explore, and explain.