UNCORRECTED PAGE PROOFS

MATERIAL FROM THIS COPY SHOULD NOT BE QUOTED OR USED WITHOUT FIRST CHECKING WITH THE PUBLISHER, AS SOME OF THIS MATERIAL MAY NOT APPEAR IN THE FINISHED BOOK.

WE KNOW IT
WHEN
WE SEE IT

WE KNOW IT WHEN WE SEE IT

What the Neurobiology of Vision Tells Us About How We Think

Richard Masland

BASIC BOOKS
New York

Copyright © 2020 by Richard Masland
Cover design by XXX
Cover image [Credit here]
Cover copyright © 2020 Hachette Book Group, Inc.

Hachette Book Group supports the right to free expression and the value of copyright. The purpose of copyright is to encourage writers and artists to produce the creative works that enrich our culture.

The scanning, uploading, and distribution of this book without permission is a theft of the author's intellectual property. If you would like permission to use material from the book (other than for review purposes), please contact permissions@hbgusa.com. Thank you for your support of the author's rights.

Basic Books
Hachette Book Group
1290 Avenue of the Americas, New York, NY 10104
www.basicbooks.com

Printed in the United States of America
First Edition: March 2020
Published by Basic Books, an imprint of Perseus Books, LLC, a subsidiary of Hachette Book Group, Inc. The Basic Books name and logo is a trademark of the Hachette Book Group.

The Hachette Speakers Bureau provides a wide range of authors for speaking events. To find out more, go to www.hachettespeakersbureau.com or call (866) 376–6591.

The publisher is not responsible for websites (or their content) that are not owned by the publisher.

Print book interior design by Linda Mark.

Library of Congress Cataloging-in-Publication Data has been applied for.
ISBNs: 978-1-5416-1850-3 (hardcover), 978-1-5416-1849-7 (ebook)

LSC-C

10 9 8 7 6 5 4 3 2 1

Contents

Introduction 000

PART I. THE FIRST STEPS TOWARD VISION 000

1 | The Wonder of Perception 000
2 | Neurons That Sing to the Brain 000
3 | A Microprocessor in the Eye 000
4 | Ghost Neurons 000
5 | What the Eye Tells the Brain 000

PART II. INTO THE WILD 000

6 | Sensory Messages Enter the Brain 000
7 | What Happens Next: Not One Cortex but Many 000
8 | The Malleable Senses 000
9 | Inventing the Nerve Net: Neurons that Fire Together Wire Together 000
10 | Machine Learning, Brains, and Seeing Computers 000
11 | A Vision of Vision 000

PART III. TO THE HORIZON 000

12 | Why Evolution Loved Nerve Nets 000
13 | Some Mysteries, Some Progress 000
14 | In the Distance 000

Glossary 000
Acknowledgments 000
Figure Credits 000
Bibliography 000
Notes 000
Index 000

Introduction

THIS IS A BOOK ABOUT HOW WE SEE. THINKERS HAVE PONDERED vision for a long time, but most of their ideas were, by modern standards, naive: the eye *is*, in fact, something like a camera, but there is a whole lot more to vision than that. It may seem natural and simple that we can recognize the face of a friend—so much so that the ancients didn't even identify it as a problem—but there is actually nothing simple about it. To truly understand vision, you have to understand more than just how our eyes work. You also must understand how our brains make sense of the outside world.

Paradoxically, brains are pretty slow; neurons and their synapses work millions of times more slowly than modern computers. Yet they beat computers at many perceptual tasks. You are able to recognize your child among the crowd on the playground in milliseconds. How does your brain do it? How does it take a blunt stimulus—a patch of light, a vibration in the air, a change of pressure on the skin—and give it meaning? We have only glimpses of the ways, but what we have learned is fascinating.

We'll go through vision step by step. You'll hear that the world you see is not the world that actually exists: it has been broken into fragments by your retina and sent to your brain in separate channels, each telling the brain its specific little thing about the image. You'll learn how this recoding is accomplished by neurons in your retina, and why. We'll follow these signals into the brain, where they build our perceptions.

The brain holds many mysteries, but an important insight is that much of the brain works not by fixed point-to-point connections, like the telephone system, but by means of swarms of neurons interconnected, like a spiderweb, into nerve nets. These days, nerve nets are often associated with computers, but in fact they were thought up a half century ago by a far-seeing Canadian neuroscientist, Donald Hebb. A few years later the idea was co-opted by computer scientists. During the next decades nerve nets moved in and out of fashion, but better computers eventually allowed computer scientists to create the field of machine learning, better known as artificial intelligence. They showed that computer nerve nets can learn to perform dramatic feats, leading neuroscientists to look again at nerve nets in the brain. So today we have a remarkable alliance between neurobiology and computer science, each field informing the other.

Do brains use nerve nets to interpret the world? Does the brain work by "machine learning"? The answer seems to be yes—and brains do it a whole lot better than computers. Outside some specialized tasks, it turns out that computers are very bad brains, and a search is on to make them more brain-like.

The key to machine learning, as imagined long ago by Hebb, is that a nerve net connected by fixed wiring cannot do very much. As it happens, the synapses that connect the neurons of a nerve net (or the simulated "neurons" of a computer) are modifiable by experience. This plasticity is a general rule in the brain—not just in sensory systems. It helps the brain recover from injury, and allows it to allocate extra brain resources to tasks that are particularly important. In vision, the nerve nets of the brain can learn to anticipate the identity of an object in the world—to supplement the raw information coming from the retina with its knowledge of images it has seen before. Boiled down, this means that much of perception is not just a fixed response to the visual scene but is learned. The brain's nerve nets recognize certain combinations of features when they see them.

Where does this lead in our search for understanding the actual experience of perception, thinking, emotion? We don't have a detailed answer, but we can see, far in the distance, how the final answer may look. Known, verifiable science can take us to an entry point. I will take us part of the way, to the seam where sensory experience turns into perception and thought.

Finally, where are "you" in all this? It's easy enough to talk about the brain as we see it from the outside, but where is the inner person that we imagine to be looking out through our eyes? There we can barely begin—and we run inexorably into the nature of consciousness, the self. We'll go there at the very end, with no answer but an attempt to see the problem more clearly.

* * *

I HAVE BEEN a neuroscientist since I was twenty-five—before the discipline of neuroscience officially existed—and I care as much about it now as I did then. I've watched our understanding evolve, and I've participated in the work myself. The basic narrative of this book is "how vision works"—from the retina to the highest visual centers deep in the temporal lobe. But I also want to let you follow the scientific journey, to see how basic neurobiology—not the talk-show kind—looks from beside the laboratory bench. So I'll mix in some scenes from the lab, and sketch some of the players.

PART I

THE FIRST STEPS TOWARD VISION

During the 1960s, a good teacher named Jacob Beck gave a college course titled simply “Perception.” The course met in a small auditorium tucked into a corner of Memorial Hall, a nineteenth-century brownstone colossus erected as a memorial to Harvard’s Civil War dead. The lecture hall’s gradual slope accommodated perhaps a hundred brown wooden desks, covered with a century’s coats of yellowing varnish. A black chalkboard stretched the width of the front wall. High on the left wall were sparse windows. The room was otherwise lit by a few incandescent bulbs, turning the auditorium a soft yellow. Thirty or forty students were thinly scattered around the room.

Beck was as straightforward a teacher as the name of his course would suggest. His manner was pleasant enough, but he was not particularly interested in charming students—his main mission was to present his material in a clear and organized way. He used careful notes and stuck to them. He spent the first few minutes of each lecture reviewing the main points covered in the previous one.

Beck did not need showmanship. The material was fascinating in itself. To be sure, he taught us the basics: pressure on the skin deforms a nerve ending, which sends a signal up the spinal cord to the brain. Some of our skin sensors signal light touch, some signal heat, and some are for things moving across our skin—say, a venomous bug dropping on your arm from the forest canopy. Facts like these were interesting in their own right. But the most wondrous thing of all—the great challenge Beck posed to his roomful of nineteen-year-olds—was object recognition.

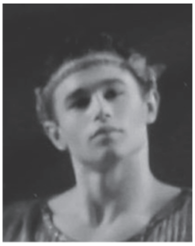
On one hand, this is a problem of sensation—how the eye works, how it signals to the brain. But it connects with the great issues of perception: thinking, memory, the nature of consciousness. We can get our hands on the pathways of sensation. We can record the electrical signals in sensory pathways. We can tease the neurons to tell us what they see. We now know a lot about how the sensory signals are handled—how they are passed from station to station in the brain. This gives us a handle on the larger questions; it is a place where we have certain kinds of secure knowledge. We are only starting to understand where the brain takes things from there. But taking vision step by step gives us a platform from which to peer toward the great mysteries.

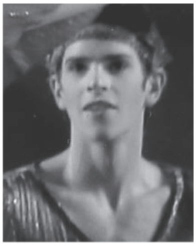
1 | The Wonder of Perception

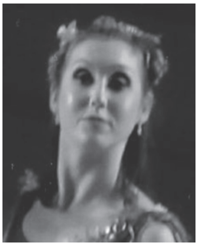
The pears are not viols,
Nudes or bottles.
They resemble nothing else.

They are yellow forms
Composed of curves
Bulging toward the base.
They are touched red.
—Wallace Stevens

Consider these three faces. Although the images are slightly blurry and the contrast is poor, you can tell them apart. The woman on the right has a slightly rounder face; the boy on the left has a strong chin. If they were your son or your daughter, your friend or your mother, you would recognize them across







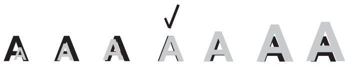
an amazing variety of situations. You would recognize them in plain clothes, without their makeup. You would recognize them in front of you or from an angle. You would recognize them in bright light or dim, nearby or at a distance, glad or sad, laughing or silent.

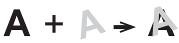
Yet *how* do you recognize them in all those different instances? The actual image that falls upon your retina is physically different in each case. Your brain adjusts to each version: larger or smaller, brighter or dimmer, smiling or glum. The permutations of faces, received as physical stimuli falling on your retina, are almost infinite. Yet you recognize familiar faces instantly, without effort. And you can tell apart not just these three but hundreds or thousands of faces. How can the brain—which is only a physical machine, like any other—perform this task so well?

It may help to think about a simpler example. Imagine that you must design a computer program that can recognize the letter A. Modern computers do this with ease, right? But in comparison with brains, they cheat. More on that in a minute.

The solution seems obvious: somewhere in the computer (or your brain) there must be a map or template of the letter A. Then the computer (or the brain) can just compare an A with the template and match them. But what if the size of the A to be recognized is different from the size of the template? The computer (or the brain) would have to conclude that they are not the same letter.

Well, why not just have the computer test a bunch of different-sized templates? That would fix the problem:



No doubt about it, that would work. Suppose, however, that the test A is now tipped a bit: 

They won't match, no matter how perfectly the computer has guessed the size.

OK, then, let's have the computer compare against all possible sizes and all possible angles. If the computer is fairly fast, that might work. But in the end, we'd have too many variants—line thickness, color, font, and so on. And then we'd have to multiply all those variants by each other. The computer ends up having to test all possible sizes *times* all possible angles *times* all possible fonts *times* all possible colors, and so on. The number of combinations that has to be tested becomes very, very large, impractically large. All this hassle for a simple letter.

With faces, there is almost no limit to the variants. A face can be smiling or frowning, dim or bright, viewed from the front or at an angle. And the components of brains—neurons and synapses—are, compared with computers, very slow. It takes about a thousandth of a second for a neuron in a human brain to transmit its most basic signal across a synapse to one of its fellow neurons. During that time, a pretty fast modern computer performs something like 1 million operations. This superhuman speed is why I said that computers cheat—they do something that ordinary wet biology could never do. Say it takes one hundred operations for a computer to make one comparison. A computer could therefore make a hundred thousand comparisons in the time it takes a brain to transmit a single nerve impulse across a synapse. And that's not counting the time it takes the signal to travel down the nerve fibers that connect neurons. If it were making comparisons in the same way computers do, your poor old brain would take minutes to recognize even the most familiar face. In other words, making lots of guesses is not an option for brains.

Here's another example, drawn from a different sense—hearing.[1] It's the problem called segmentation. If I say to you "The dog is blue," you'll generally hear the words as they are written on this page. But normal spoken speech does not have breaks between words. In the actual acoustics of that sentence (unless you speak with artificial breaks), there are no empty spaces between the sounds "the," "dog," "is," and blue." In physical reality, the sentence is a single long sound. To make sense of it, our brains break that long sound into words we know from a lifetime of speaking English (or whatever language we are using).

Once again, it is virtually impossible to see how the brain could use a template and match words against it. How many sounds would the template include? Certainly far more than the words in a dictionary. And this is to say nothing of different accents, rates of speaking, background noise, and more. So the brain isn't using a template to understand this string of sounds.

This whole big mystery—an act that we perform many times daily with such ease—is termed the problem of object recognition. We think of it as being about sensory experience, but it is just as much a problem of memory: object recognition is matching a present stimulus to the memory of an object seen in the past. To figure out how it works is a spectacular technical challenge—the Mount Everest of sensory neurobiology.

2 | Neurons That Sing to the Brain

> You learn something general by studying
> something specific.
>
> —Stephen Kuffler

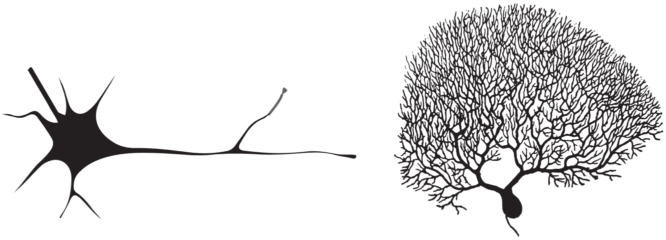
I HAVE TOLD YOU THAT THE WORLD YOU THINK YOU SEE IS NOT THE world that actually exists. It has been altered by your retina, fragmented into dozens of different signals for transmission to the brain. The retina parses the visual image into its most telling components and sends a separate stream of signals about each of them to the brain. The rest is ignored, treated as background noise. This kind of stripped-down signaling, which is a search for economy you'll hear more about, is not just a result of evolution amusing itself; it is one of the most fundamental principles of all perception.

To see how it happens, we have to get down to basics.

A SINGLE NEURON

A neuron is not a complicated thing. It is a physical object, albeit a very small one, made of materials we understand. It has the normal parts that make up any animal cell, with only a few unique features.

When you concatenate a few hundred million neurons, though, big things happen: recognizing a friend, hearing Beethoven, a one-handed catch of a thirty-yard forward pass.



A neuron, like all vertebrate cells, is a bag of water separated from the surrounding water by a thin, fluid membrane. The membrane serves to divide the space inside the cell (black in these drawings) from everything outside it. A few neurons are more or less round, like a kid's balloon. Others take more complex, amoeboid shapes. Still others can have bizarre and complicated arrangements. Many neurons look like skeletons—a tree in winter. The twigs and branches reflect the neuron's connections with its near or far neighbors. No matter how baroque the shape, however, the cell always consists of a single space enclosed by a membrane. Its thin twigs enclose long thin spaces, like branching, convoluted soda straws.

What is this cell membrane? It is made of lipid, a variety of fat. Since fat and water do not mix, the cell membrane stays separate, a bit like a soap bubble. By itself, the cell membrane cannot accomplish much of anything. In the lab you can make an artificial cell that has only a cell membrane. Such a cell just sits there. An actual cell membrane is studded with a myriad of fancy little machines that do specific tasks—for example, embedded protein molecules that sense other molecules impinging from outside and then open a gate between the outside and the inside of the cell, which allows an electrical charge to pass. This is the basis of the nerve impulse.

While nerves have an impressive functional repertoire, their main function—the function that distinguishes them from almost all other cells—is to communicate with other neurons. In most cases they do so by transmitting brief impulses of electrical activity, known as spikes. Spikes can travel short distances or long. Some neurons talk to others (we say they "conduct nerve impulses") only within their own restricted neighborhood. These so-called interneurons (local circuit neurons) signal over distances as small as 10 micrometers, which is just a hundredth of a millimeter. Alternatively, some spikes travel all the way from your brain to the bottom of your spinal cord, as when you seek to wiggle your big toe, or in the reverse direction, as when you stub it on a brick.

Spikes are not electrical currents, like the currents that a copper wire conducts. They are a more complicated biological event, in which the cell membrane actively participates: they are the electrical reflection of the movements of charged ions into and out of the cell, guided by specialized proteins that sit in the cell membrane. For that reason, they travel very slowly compared to electrical conduction down a wire. Nerve impulses travel along an axon at a speed ranging from roughly 10 to 100 meters per second, depending on the axon. Electricity flows down a wire at about 300 million meters per second. From the point of view of our brain's ability to compute things, this slowness of conduction is a big deal. It is the main reason brains cannot use brute-force, dumb strategies to solve problems.

At the end of the axon is usually a synapse. A synapse allows one neuron to talk to other neurons across the gap that separates them. At the synapse, an electrical signal in one neuron is changed into a chemical signal; specialized synaptic machinery allows the spike to trigger the release of chemicals that are sensed by the second neuron. These are neurotransmitters, about which we hear so much in the news. Because there are lots of different kinds of neurotransmitters, used for different purposes at various places around the brain, and lots of steps involved in their release, this is a point where we can manipulate brain function—for therapeutic goals or for recreation.[1]

Nicotine acts on synapses. So do antipsychotic drugs, and those that control epileptic seizures. So does Valium, to make you calm, or Prozac, to make you happy.

A neurotransmitter released by one neuron can make another neuron more excited or less excited. (In reality, a neuron is rarely receiving just a single signal, but for the present purposes let's just assume it does.) The second neuron integrates all the inputs it receives. When enough impulses reach that neuron within a short time, what we call an "action potential" is triggered in the neuron. That action potential can propagate autonomously within the second neuron and excite or inhibit a third neuron, and so on.

At this point, we see the second big thing neurons do: they decide which inputs to pass on to further neurons and which inputs not to pass on. They make this decision solely by adding up all the inputs the cell receives. This is a bit of a simplification, as the ways in which inputs can be received are wonderfully varied. But to take a simple example, they add excitatory inputs and subtract inhibitory ones. The study of this process constitutes a field of its own within neurobiology; some of my smartest colleagues have spent their lives unraveling the many and elegant ways in which synaptic communication can occur.

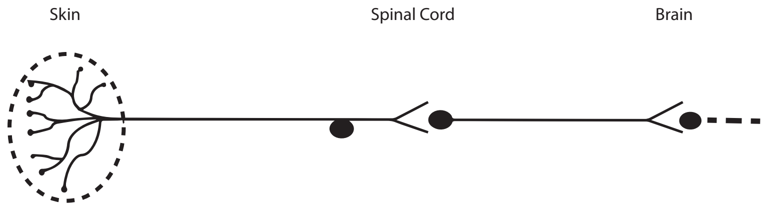
Now, though, we'll think of neurons in their simplest mode, waiting for inputs to come along and firing an action potential when those inputs reach a certain magnitude. But just sending messages from neuron to neuron does not make a brain a brain. It is the combination of neuronal signaling and neuronal decision-making that makes a brain a brain. I'm simplifying because my task here is to tell you about perception. For that, we only need to understand a few things. The most important is that an action potential causes an electrical change wherever it goes. Critically for our story, that electrical change—the spike—can be eavesdropped upon by mortals armed with long thin probes called microelectrodes.

HOW SENSORY NEURONS SIGNAL

As I've said, neurons carry messages from place to place over a distance that can be short or long. In a giraffe, the neurons that control walking can span two and a half meters, reaching from the brain to the lower spinal cord. In all but a few cases, however, the means of signaling is the same: somewhere on the surface of the cell, there is a stimulus that initiates an action potential that spreads throughout the neuron.

All neurons that sense the outside world—whether through touch, hearing, vision, or smell—do the same fundamental thing: they detect an event in the world and transmit a signal about it, sometimes with a relay or two, to the brain. But they do this in quite different ways, mirroring the events in the world they are sensing, which are also physically different.

Consider the sense of touch. A perception of touch originates when the skin is deformed through pressure. This could occur by a finger stroking your wrist, a mosquito walking gingerly in search of a soft spot to stab, or your brusque collision with some solid object. These deforming pressures, hard or soft, are detected by nerve endings located just below the surface of the skin. Each ending is part of a neuron.



This image shows two neurons on the touch pathway; the patch of skin, shown by the dashed circle, is known as the receptive field. Information travels from left to right in this diagram. The first neuron has a long fiber (axon) that runs from a place on the skin—where the nerve ending forms many small branchlets—to the spinal cord. When, say, a mosquito lands on your arm, the mosquito's foot ever so slightly depresses the skin over the nerve ending. That pressure is transmitted to the neuron and a nerve impulse is initiated. That impulse travels along the axon, through the cell body and ends at a synapse (indicated by a forked line) upon another neuron, located in the spinal cord, which then projects to the brain. (Other pathways to the brain exist. This is just one of the simplest.)

The tactile neuron's branches can detect the indentation of the surface of the skin through what is known as a mechanosensitive ion channel. This channel is a protein in the cell membrane. Deforming the mechanosensitive channel allows positive ions to flow from outside the cell into the nerve ending. The flow of positive ions tends to excite the ending. When the excitation reaches a certain threshold, the ending begins to fire action potentials. These travel up the skin sensory nerve (axon) and past the cell body to a collecting site in the spinal cord, where the axon encounters a second neuron that will transmit the information toward the brain for interpretation. Note that the information from this skin sensory nerve has told the rest of the nervous system three things: that there is something touching your skin, that it is located just above your right wrist, and that the thing is fairly light.

First, the "where," which is easy. The endings of an individual tactile neuron in the skin cover a limited space on the skin. This space may be tiny, such as on the hand or lip, or broader, such as on the skin of the back. The brain knows what region each nerve surveys, and from that it knows where on the skin the stimulus fell—where the receptive field of that neuron is located.[2] Obviously, if the stimulus falls on an area of the body like your fingertip, which is covered by many tiny nerve endings, the brain will know more precisely where a small stimulus fell than for places with only a few huge endings, like on your back.

I introduced an important piece of nomenclature when I referred to the dashed circle in the diagram. I called this area under the terminal branchlets of a sensory axon a cell's receptive field. The receptive field is the specific part of the skin from which a particular sensory axon can be excited. As you will see, we use same term to talk about vision, where "receptive field" refers to the area of the retina that excites a particular visual neuron—in the retina or later in the visual system.

Now for the "how much" question: how light or heavy the stimulus. How does the skin sensory nerve convey that? All sensory axons—for touch, hearing, vision, smell—communicate with the brain by a coded frequency of action potentials. A light touch produces only a few action potentials; a stronger one produces a more rapid string of them. That's how the brain, or an experimenter who can monitor the rate of firing, can tell how strong the stimulus was.

Many scientists (including me) have speculated in print that additional information may be contained in the detailed *pattern* of the action potentials, just as a pattern of key taps conveys information in Morse code.[3] The pattern could tell the brain, for example, what type of receptor a particular axon is carrying signals from (see the next paragraph). Certainly the pattern of spikes influences how the brain responds; we know that closely spaced action potentials (spikes) excite the postsynaptic cell more powerfully than widely spaced ones. But nobody has proposed and tested a specific code that has turned out to be convincing.

Even more interesting is the "what" part of our question. The brain wants to know: "What kind of a thing is touching my wrist?" All touches are not created equal. There are several different kinds of touch neurons, responding to different aspects of touch. One type of touch receptor is moderately sensitive to light touch on the surface of the skin and keeps sending a signal to the brain as long as the lightly touching thing is still touching. Another type of receptor