Second Nature

BRAIN SCIENCE AND HUMAN KNOWLEDGE

Gerald M. Edelman

YALE UNIVERSITY PRESS
New Haven and London

Contents

PREFACE ix
ACKNOWLEDGMENTS xi
INTRODUCTION 1

one
The Galilean Arc and Darwin's Program 5

two
Consciousness, Body, and Brain 12

three
Selectionism
A Prerequisite for Consciousness 23

four
From Brain Activity to Consciousness 35

five
Epistemology and Its Discontents 43

six
A Brain-Based Approach 53

seven
Forms of Knowledge
The Divorce between Science and the Humanities 68

eight
Repairing the Rift 76

nine
Causation, Illusions, and Values 88

ten
Creativity
The Play between Specificity and Range 98

eleven
Abnormal States 106

twelve
Brain-Based Devices
Toward a Conscious Artifact 125

thirteen
Second Nature
The Transformation of Knowledge 142

NOTES 159

BIBLIOGRAPHY 173

INDEX 183

Introduction

FROM TIME TO TIME I HAVE a dream. In it, the historian Henry Adams appears, moaning about complexity, muttering about the Virgin and the Dynamo. There usually is no more to the dream than that. When I remember enough detail in my waking state, I connect it to the famous chapter in *The Education of Henry Adams.*[1] In that chapter, Adams recounts his sense of inadequacy before the forty-foot dynamos that his friend Langley the engineer showed him at the Paris Exhibition of 1900. Adams contrasts the complexity of these engines to the simplicity of the religious turn to the Virgin Mary. That theme, its variations, and Adams's sense of not being comfortable in his time run through the *Education.*

Adams, a scion of the great family descended from John Adams, was an accomplished historian. His sense of alienation prompts speculation. Could it have simply been the symptom of clinical depression? Could it have been entwined with the circumstances leading to his wife's suicide? Or could it have reflected a genuine rift between the way a person sees the world from the standpoint of science and the way that person sees it reflected in the humanities?

We do not know. But one thing is sure. There is a divorce between science and the humanities, and between the so-called hard sciences such as physics and the human sciences such as sociology. Perhaps my recurrent dreams of Henry Adams come from my persistent interest in the origin of this estrangement.

I have long puzzled over the gap between scientific explanations and everyday experience, whether by individuals or in historical settings. Is the divorce between science and the humanities inevitable? Can the human sciences be reconciled with the hard sciences?

Views on these questions have ranged widely, and some might say that they aren't worth bothering with. As this book attests, I believe the opposite—that understanding how we arrive at knowledge, whether by scientific inquiry, by reason, or by happenstance, is of major importance. Wrongheadedness, severe reductionism, or insouciance can each have unfortunate long-range consequences for human welfare.

This book is the result of a line of thought leading to what I have called brain-based epistemology. This term refers to efforts to ground the theory of knowledge in an understanding of how the brain works. It is an extension of the notion of naturalized epistemology, a proposal made by the philosopher Willard Van Orman Quine.[2]

My line of argument differs from his, which stopped, as it were, at the skin and other sensory receptors. I deal with the issue by considering a wider-ranging interaction—that between brain, body, and environment. I believe that above all, it is particularly important to understand the basis of consciousness. Quine with his usual ironic candor said,

> I have been accused of denying consciousness, but I am not conscious of having done so. Consciousness is to me a mystery, and not one to be dismissed. We know what it is like to be conscious, but not how to put it into satisfactory scientific terms. Whatever it precisely may be, consciousness is a state of the body, a state of nerves.
>
> The line I am urging as today's conventional wisdom is not a denial of consciousness. It is often called, with more reason, a repudiation of mind. It is called a repudiation of mind as a second substance, over and above body. It can be described less harshly as an identification of mind with some of the faculties, states, and activities of the body. Mental states and events are a special subclass of the states and events of the human or animal body.[3]

I believe we are now in a position to reduce the mystery. In this book, I lay out thoughts that reveal this position and bear directly on how we know, on how we discover and create, and on our search for truth. I follow in the footsteps of William James, who pointed out that consciousness is a process whose function is knowing.[4]

There is nature and human nature. How do they intersect? The title I have chosen reflects this question and is to some extent a play on words. The term "second nature" usually refers to an act done spontaneously, easily and without the need for exertion or learning. I use the term here to include this meaning but also to call attention to the fact that our thoughts often float free of our realistic descriptions of nature. They are a "second nature." I aim to explore here how nature and second nature interact.

one

The Galilean Arc and Darwin's Program

Almost everything that distinguishes the modern world from earlier centuries is attributable to science, which achieved its most spectacular triumphs in the seventeenth century.

—BERTRAND RUSSELL

The Origin of Species *introduced a mode of thinking that in the end was bound to transform the logic of knowledge, and hence the treatment of morality, politics, and religion.*

—JOHN DEWEY

Something definite happens when to a certain brain state a certain 'sciousness corresponds.

—WILLIAM JAMES

Henry Adams didn't know the half of what was coming. But he did sense a transformation of our existence by scientific technology. We are in the midst of a revolution: communication, computers, the Internet, the explosion of travel by land and air, atomic power, biological manipulation of our genetic makeup. One could go on and on about the technological substrate and the globalization that has changed the pace of our lives, the modes of our thought, our place in nature, and our threat to it.

What has happened to our conception of nature and of our second nature? To answer this question we have to take a longer view of Western science, particularly of physics and biology. I pick two figures, Galileo Galilei and Charles Darwin, to highlight the developments that have so changed our lives.

First, Galileo, who can be taken to represent the birth in the seventeenth century of modern physics, the broadest of modern sciences. The philosopher Alfred North Whitehead, in his book *Science and the Modern World,* called Galileo's achievement "the quiet commencement of the most intimate change in outlook which the human race had yet encountered."[1] Surely, we must be impressed by the arc of modern physics ranging from Galileo's ideas on the heavens and his experiments on inertia to our present cosmology and theories of matter. We must confront the weird domain of the very small described by quantum mechanics as well as the grand elegance of general relativity as it opens up vistas of the very large, the universe itself. So now the Galilean arc ranges from nuclear power to solid-state physics, to the exploration of space, and to the origin of the universe itself in the Big Bang.

Even before these advances, a vision of the basis of life, the evolution of living things, was laid down by Charles Darwin in the second half of the nineteenth century.[2] Darwin's development of the idea of natural selection provided the theoretical basis for understanding life itself, particularly when it was coupled with Mendelian genetics in the twentieth century.[3] The further development of molecular biology in the latter part of that century has made it possible to change the very basis of biological reproduction.

In looking over the domains of nature it may appear that, if we include Darwin, the Galilean arc has provided an enlightened understanding of all major subject categories: galaxies, stars, and planets, the structure of matter, the nature of genes and biological evolution. Henry Adams's plate today would be more than full of scientific matters covering much of our existence. But there is a gap or an incompleteness in the Galilean arc. We have not yet scientifically founded the bases of consciousness in the brain, an issue that, until recently, has been left to the philosophers.

There are reasons for this. Until recently, noninvasive methods of examining events in the brain were lacking. More than that, consciousness is a first-person affair, whereas the objective methodology of science is a third-person affair. Opinions, subjectivity, and the like cannot be admitted in scientific experiments. An equally large factor affecting the scientific approach to consciousness can be attributed to the influential thoughts of René Descartes.[4] Sometime after Galileo, Descartes essentially removed the mind from nature. He did this by thought alone, concluding that there were two substances: *res extensa,* extended things that were susceptible to physics, and *res cogitans,* thinking things that were not extended in space and that were unavailable to physics. Descartes's dualism and its various subsequent derivations have had a profound effect on the approach to consciousness as a valid scientific target.

This state of affairs is most curious. In principle, no subject is a priori immune from scientific inquiry. Yet the very ground of our awareness has been left outside the pale! Science is imagination in the service of the verifiable truth. And as such, imagination is actually dependent on consciousness. Science itself is so dependent. As the great physicist Erwin Schrödinger observed, no scientific theory in physics includes sensations and perceptions and to get ahead it must therefore assume these phenomena as being outside of science's grasp.[5]

Must we accept this state of affairs? Or can science complete the Galilean arc? If it cannot, must it leave the ground of consciousness to the philosophers, to the humanities, and thereby acquiesce to the divorce that so concerned Henry Adams?

Thanks to discoveries about the brain and advances in brain theory in the past twenty years, it seems we do not have to remain in this predicament. We can study consciousness even in the face of subjectivity. My aim here is to show how. But first, let us turn to the significance of a scientific understanding of consciousness.

I have found that some people do not believe that a scientific account of consciousness offers much in the way of consequences. My remarks here are not specifically aimed at these doubters, but I hope they will persuade some at least to consider the contrary position. I start with a big assumption: that we have a satisfactory scientific theory of consciousness based on brain activity. What would its significance be?

First, it would clarify the relation between mental and physical events and clear up some outstanding philosophical puzzles. We would no longer have to consider dualism, panpsychism, mysterianism, and spooky forces as worth pursuing.[6] Time would be saved, at the least. And in clarifying these issues, we would have a better view of our place in the natural order. We would be able to corroborate Darwin's view that the human mind is the outcome of natural selection and thereby complete his program.[7]

We would also have a better picture of the bases of human illusions, useful and otherwise. One illusion I hope to dispel is the notion that our brains are computers and that consciousness could emerge from computation. Furthermore, a successful theory of consciousness might clarify the place of values in a world of facts. In connection with both of these issues, a brain-based theory would be of great use in understanding psychiatric and neuropsychological syndromes and diseases.

Tangent to these matters, a brain-based theory might contribute to our notions of creativity. It might even provide a clearer view of the connection of objective descriptions derived from hard science to normative issues that arise in aesthetics and ethics. To that degree, it may help undo the divorce between science and the humanities.

Above all, achieving these ends may contribute to and affect the formulation of a biologically based epistemology—an account of knowledge that relates truth to opinion and belief, and thought to emotion by including aspects of brain-based subjectivity in an analysis of human knowledge.

The most remarkable outcome of a satisfactory brain theory would be the construction of a conscious artifact.[8] Although that goal is presently in the realm of fantasy, scientists at The Neurosciences Institute in La Jolla, California, have already built brain-based devices that have perceptual and memorial capabilities. Of course, a minimum requirement for us to believe that we had constructed a conscious device would rest in its ability to report, through a language, its internal phenomenal states while we measure its neural and bodily performance. This requirement is presently far from being met. But if it were, we would have an unparalleled opportunity to explore brain, body, and environment as they interact in such a device. Would it "see" or "sense" the world in ways we cannot imagine? Only a receipt of messages from outer space would exceed this enterprise in excitement. We shall have to wait.

I propose now to provide some support for the assumption I made that we have a satisfactory theory of consciousness. I shall do so by presenting a brief account of consciousness and of the brain dynamics from which it emerges. Following that, we can return to analyze their consequences in greater detail.

two

Consciousness, Body, and Brain

Like the entomologist in search of brightly colored butterflies, my attention hunted, in the garden of gray matter, cells with delicate and elegant forms, the mysterious butterflies of the soul.

— Santiago Ramón y Cajal

I HAVE WRITTEN EXTENSIVELY on the details of brain structure and dynamics as they relate to perception, memory, and consciousness. I have no intention of repeating these details here. Instead, I shall describe some of the main features of consciousness. Then I will give a brief account of brain activity in terms of a theory called Neural Darwinism.[1] This will allow me to show how consciousness emerges from brain dynamics. I shall not hesitate to make large statements without detailed proof; such proof can be found elsewhere.[2]

We all know implicitly what consciousness is. It is what you lose on entering a dreamless deep sleep and, less commonly, deep anesthesia or coma. And it is what you regain after emerging from these states. In the awake conscious state, you experience a unitary scene composed variably of sensory responses—sight, sound, smell, and so on—as well as images, memories, feeling tones and emotions, a sense of willing or agency, a feeling of situatedness, and other aspects of awareness. Being conscious is a unitary experience in the sense that you cannot at any time become totally aware of just one thing to the complete exclusion of others. But you can direct your attention to various aspects of a less inclusive but still unitary scene. Within a short time, that scene will vary in one degree or another and, though still integrated, will become differentiated, yielding a new scene. The extraordinary fact is that the number of such privately experienced scenes is apparently limitless. The transitions seem to be continuous, and in their complete detail they are private, first-person subjective experiences.

Conscious states are often, but not always, about things or events, a property called intentionality. But they do not necessarily always show this property; they can, for example, be about a mood. There is often a just-aware "fringe," as William James called certain barely perceived states. Conscious states can also involve the awareness of agency or the willing of an action.

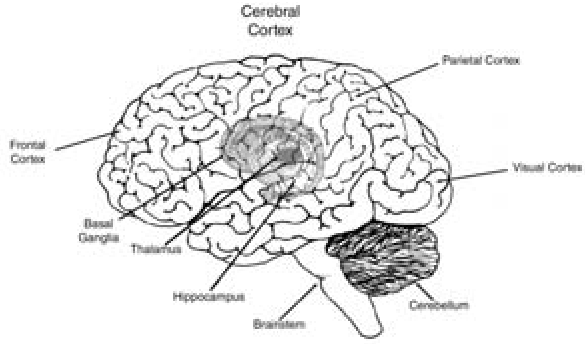
The property most often described as particularly mysterious is the phenomenal aspect of consciousness, the experience of qualia. Qualia are, for example, the greenness of green and the warmness of warmth. But several students of the subject, myself included, go beyond these simple qualities and consider the whole ensemble of conscious scenes or experiences to be qualia.

Many consider explaining qualia to be the acid test of a consciousness theory. How can we explain not only qualia but all the other features of consciousness? The answer I propose is to look into how the brain works, formulating a global brain theory that can be extended to explain consciousness. Before I do so, however, one more distinction will prove useful. As human beings, we know what it is like to be conscious. Moreover, we are conscious of being conscious and can report on our experience. Although we cannot experience the consciousness of members of another species, we surmise that animals like dogs are conscious. We do this on the basis of their behavior and the close similarity of their brains to ours. But we do not usually attribute consciousness of consciousness to them.

This is the basis for a useful distinction. Dogs and other mammals, if they are aware, have primary consciousness. This is the experience of a unitary scene in a time period of at the most seconds that I call the remembered present—a bit like the illumination by a flashlight beam in a dark room. Although they are aware of ongoing events, animals with primary consciousness are not conscious of being conscious and do not have a concept of the past, the future, or a nameable self.

Such notions require the ability to experience higher-order consciousness, and this depends on having semantic or symbolic capabilities. Chimpanzees appear to have the rudiments of these capabilities. In our case, they exist in full flower because we have syntax and true language. With the ability to speak, we can free ourselves temporarily from the limitations of the remembered present. Nonetheless, at all times when higher-order consciousness is present we also possess primary consciousness.

Against this background, let us turn to the organ responsible for all these extraordinary traits: the brain. The human brain weighs about three pounds. It is one of the most complicated material objects in the known universe. Its connectivity is awe-inspiring: the wrinkled cortical mantle of the brain


Cerebral Cortex
Parietal Cortex
Frontal Cortex
Visual Cortex
Basal Ganglia
Thalamus
Hippocampus
Cerebellum
Brainstem

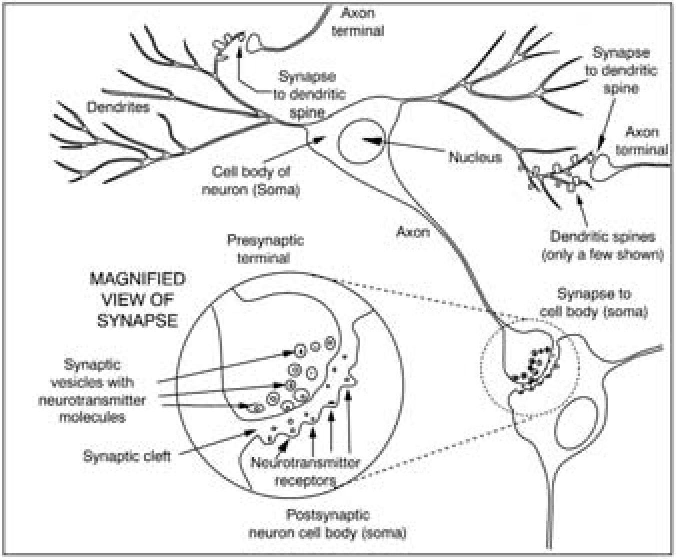

Axon terminal
Synapse to dendritic spine
Dendrites
Synapse to dendritic spine
Axon terminal
Nucleus
Cell body of neuron (Soma)
Axon
Dendritic spines (only a few shown)
Presynaptic terminal
MAGNIFIED VIEW OF SYNAPSE
Synapse to cell body (soma)
Synaptic vesicles with neurotransmitter molecules
Synaptic cleft
Neurotransmitter receptors
Postsynaptic neuron cell body (soma)

Figure 1.
Top: Relative locations of major parts of the human brain. The cerebral cortical mantle, which has about thirty billion neurons, receives projections from the thalamus and sends reciprocal projections back; this constitutes the thalamocortical system. Beneath the mantle are three major cortical appendages: the basal ganglia and cerebellum (both of which regulate movement), and the hippocampus, which is necessary for memory. Below them is the oldest part of the brain in evolutionary terms, the brain stem, which contains several diffusely projecting value systems.

Bottom: Synaptic connections between two neurons. An action potential traveling down the axon of the presynaptic neuron causes the release of a neurotransmitter into the synaptic cleft. The transmitter molecules bind to receptors in the postsynaptic membrane, changing the probability that the postsynaptic cell will, in turn, fire its own action potential. Particular sequences of activity can either strengthen or weaken the synapse, changing its efficacy. (Because of the number of different shapes and kinds of neurons, this drawing is a greatly simplified cartoon.)

(figure 1, *top)* has about thirty billion nerve cells or neurons and one million billion connections. The number of possible active pathways of such a structure far exceeds the number of elementary particles in the known universe.

This is not the place to go into detail about how the brain gives rise to consciousness. I have done that in several books, which may be consulted. But I do want to provide a working picture of brain structure and activity. I propose to use a mixture of down-to-earth description, analogy, and metaphor—just enough to give an idea of how consciousness arises.

To start with, let's consider the fundamental cells that carry signals in the brain. These are the neurons, which have a treelike set of branches (dendrites) and usually a single extended process (the axon) that serves to connect one neuron to another. This connection, called the synapse (figure 1, *bottom)*, is a critical element in ensuring the function of brain circuits. This is so because electricity traveling down the axon releases little packets of chemicals called neurotransmitters at the synapse. These chemicals cross the small distance inside the synapse and bind to certain receptors present often at the dendrites of the receiving cell. If the release happens often enough, the receiving or postsynaptic cell fires and can repeat the process and signal yet another cell. Imagine such a process summing up across a myriad of synapses, and you will get an idea of why with modern methods we can actually record the otherwise minute currents and potentials over the scalp. Neurophysiologists can in fact record more precisely from single cells by invading the brain and inserting microscopic electrodes within individual neurons.

A key property of synapses is that they are plastic: various activities and biochemical events can change their strength. These changes can in turn determine which neuronal pathways are selected to transmit signals. Patterns of such changes in synaptic strength provide a basis for memory. At this point, it may be useful to mention that synapses come in two flavors: excitatory and inhibitory. Both can exhibit plasticity; together they help select the functioning signal pathways of the brain.

Now an important next step in this bowdlerized account is to point out that the overall anatomical connections and pathways in the brain of a given animal species are selected during evolution and development. The result is a stunning set of different brain areas and cell collections called nuclei. Each of these has both short-range and long-range inputs and outputs.

Let us look at the visual pathway in monkeys as an example. Light, striking cells in the retina, excites the optic nerve, whose signals ultimately reach a structure called the thalamus, a central player in our story. The thalamus is a small structure that is of great importance in any account of consciousness. Thalamic neurons mediating vision send axons to an area of the cerebral cortex called V_1. From there, all kinds of pathways within the cortex are elaborated to areas called V_2, V_3, and V_4, among others. Indeed, at least thirty-three different cortical areas are involved in one way or another in the process of vision.

Two important facts about this and several other sensory systems have emerged. The first is that, in general, each brain area is functionally segregated: in vision, V_1 for orientation of objects, V_4 for color, V_5 for object motion. The second fact is that there is no one area controlling and coordinating the responses of all the rest when a complex visual signal comes from, say, a colored moving object of particular shape. As we shall see, the brain nevertheless has means to coordinate the segregated perceptual events that occur when such a stimulus strikes the retina. The net result of such coordination is perceptual categorization—the carving up of the world of inputs into objects significant for a given animal species' recognition. The brain carries out pattern recognition. We could go on about sensory systems other than vision, but the principles are similar even if their receptors and inputs differ.

What about outputs? Well, different sensory areas connect to "higher" areas in the cortex so that the brain speaks mainly to itself. Of course, one set of cortical areas sends motor output signals to the spinal cord and thence to our muscles to elicit various actions and movements. Furthermore, the cortex receives additional inputs from, and yields outputs to, a number of subcortical structures besides the thalamus. These (see figure 1) include the basal ganglia and cerebellum, which help to regulate movement, and the hippocampus, which helps establish long-term memory of events and episodes by interacting with the cortex.

So far, what I have said could superficially be thought to describe a system analogous to an electronic device such as a computer. Indeed, in many scientific circles, there remains a widespread belief that the brain is a computer. This belief is mistaken for a number of reasons.[3] First, the computer works by using logic and arithmetic in very short intervals regulated by a clock. As we shall see, the brain does not operate by logical rules. To function, a computer must receive unambiguous input signals. But signals to various sensory receptors of the brain are not so organized; the world (which is not carved beforehand into prescribed categories) is not a piece of coded tape. Second, the brain order that I have briefly described is enormously variable at its finest levels. As neural currents develop, variant individual experiences leave imprints such that no two brains are identical, even those of identical twins. This is so in large measure because, during the development and establishment of neuroanatomy, neurons that fire together wire together. Furthermore, there is no evidence for a computer program consisting of effective procedures that would control a brain's input, output, and behavior. Artificial intelligence doesn't work in real brains. There is no logic and no precise clock governing the outputs of our brains no matter how regular they may appear.

Last, it should be stressed that we are not born with enough genes to specify the synaptic complexity of higher brains like ours. Of course, the fact that we have human brains and not chimpanzee brains does depend on our gene networks. But these gene networks, like those in the brain themselves, are enormously variable since their various expression patterns depend on environmental context and individual experience.

If the mammalian brain is not a computer, what is it? How does it work? We must answer these questions before we can explain the brain bases of consciousness.