

*Self
Comes to Mind*

CONSTRUCTING THE
CONSCIOUS BRAIN

Antonio Damasio



PANTHEON BOOKS, NEW YORK

CONTENTS

Cover
Other Books by This Author
Title Page
Copyright
Dedication
Epigraph

PART I / STARTING OVER

Chapter 1. Awakening

Goals and Reasons

Approaching the Problem

The Self as Witness

Overcoming a Misleading Intuition

An Integrated Perspective

The Framework

A Preview of Main Ideas

Life and the Conscious Mind

Chapter 2. From Life Regulation to Biological Value

The Implausibility of Reality

Natural Will

Staying Alive

The Origins of Homeostasis

Cells, Multicellular Organisms, and Engineered Machines

Biological Value

Biological Value in Whole Organisms

The Success of Our Early Forerunners

Developing Incentives

Connecting Homeostasis, Value, and Consciousness

PART II / WHAT'S IN A BRAIN THAT A MIND CAN BE?

Chapter 3. Making Maps and Making Images

Maps and Images

Cutting Below the Surface

Maps and Minds

The Neurology of Mind

The Beginnings of Mind

Closer to the Making of Mind?

Chapter 4. The Body in Mind

The Topic of the Mind

Body Mapping

From Body to Brain

Representing Quantities and Constructing Qualities

Primordial Feelings

Mapping Body States and Simulating Body States

The Source of an Idea

The Body-Minded Brain

Chapter 5. Emotions and Feelings

Situating Emotion and Feeling

Defining Emotion and Feeling

Triggering and Executing Emotions

The Strange Case of William James

Feelings of Emotion

How Do We Feel an Emotion?

The Timing of Emotions and Feelings

The Varieties of Emotion

Up and Down the Emotional Range

An Aside on Admiration and Compassion

Chapter 6. An Architecture for Memory

Somehow, Somewhere

The Nature of Memory Records

Dispositions Came First, Maps Followed

Memory at Work

A Brief Aside on Kinds of Memory

A Possible Solution to the Problem

More on Convergence-Divergence Zones

The Model at Work

The How and Where of Perception and Recall

PART III / BEING CONSCIOUS

Chapter 7. Consciousness Observed

Defining Consciousness

Breaking Consciousness Apart

Removing the Self and Keeping a Mind

Completing a Working Definition

Kinds of Consciousness

Human and Nonhuman Consciousness

What Consciousness Is Not

The Freudian Unconscious

Chapter 8. Building a Conscious Mind

A Working Hypothesis

Approaching the Conscious Brain

Previewing the Conscious Mind

The Ingredients of a Conscious Mind

The Protoself

Constructing the Core Self

The Core Self State

Touring the Brain as It Constructs a Conscious Mind

Chapter 9. The Autobiographical Self

Memory Made Conscious

Constructing the Autobiographical Self

The Issue of Coordination

The Coordinators

A Possible Role for the Posteromedial Cortices

The PMCs at Work

Other Considerations on the Posteromedial Cortices

A Closing Note on the Pathologies of Consciousness

Chapter 10. Putting It Together

By Way of Summary

The Neurology of Consciousness

The Anatomical Bottleneck Behind the Conscious Mind

From the Ensemble Work of Large Anatomical Divisions to the Work of Neurons

When We Feel Our Perceptions

Qualia I

Qualia II

Qualia and Self

Unfinished Business

PART IV / LONG AFTER CONSCIOUSNESS

Chapter 11. Living with Consciousness

Why Consciousness Prevailed

Self and the Issue of Control

An Aside on the Unconscious

A Note on the Genomic Unconscious

The Feeling of Conscious Will

Educating the Cognitive Unconscious

Brain and Justice

Nature and Culture

Self Comes to Mind

The Consequences of a Reflective Self

Appendix

Notes

Acknowledgments

About the Author

PART I

Starting Over

1

Awakening

When I woke up, we were descending. I had been asleep long enough to miss the announcements about the landing and the weather. I had not been aware of myself or my surroundings. I had been unconscious.

Few things about our biology are as seemingly trivial as this commodity known as consciousness, the phenomenal ability that consists of having a mind equipped with an owner, a protagonist for one's existence, a self inspecting the world inside and around, an agent seemingly ready for action.

Consciousness is not merely wakefulness. When I woke up, two brief paragraphs ago, I did not look around vacantly, taking in the sights and the sounds as if my awake mind belonged to no one. On the contrary, I knew, almost instantly, with little hesitation if any, without effort, that this was me, sitting on an airplane, my flying identity coming home to Los Angeles with a long to-do list before the day would be over, aware of an odd combination of travel fatigue and enthusiasm for what was ahead, curious about the runway we would be landing on, and attentive to the adjustments of engine power that were bringing us to earth. No doubt, being awake was indispensable to this state, but wakefulness was hardly its main feature. What was that main feature? The fact that the myriad contents displayed in my mind, regardless of how vivid or well ordered, *connected* with me, the proprietor of my mind, through invisible strings that brought those contents together in the forward-moving feast we call self; and, no less important, the fact that the connection was *felt*. There was a *feelingness* to the experience of the connected me.

Awakening meant having my temporarily absent mind returned, but with *me* in it, both property (the mind) and proprietor (me) accounted for. Awakening allowed me to reemerge and survey my mental domains, the sky-wide projection of a magic movie, part documentary and part fiction, otherwise known as the conscious human mind.

We all have free access to consciousness, bubbling so easily and abundantly in

our minds that without hesitation or apprehension we let it be turned off every night when we go to sleep and allow it to return every morning when the alarm clock rings, at least 365 times a year, not counting naps. And yet few things about our beings are as remarkable, foundational, and seemingly mysterious as consciousness. Without consciousness—that is, a mind endowed with subjectivity—you would have no way of knowing that you exist, let alone know who you are and what you think. Had subjectivity not begun, even if very modestly at first, in living creatures far simpler than we are, memory and reasoning are not likely to have expanded in the prodigious way they did, and the evolutionary road for language and the elaborate human version of consciousness we now possess would not have been paved. Creativity would not have flourished. There would have been no song, no painting, and no literature. Love would never have been love, just sex. Friendship would have been mere cooperative convenience. Pain would never have become suffering—not a bad thing, come to think of it—but an equivocal advantage given that pleasure would not have become bliss either. Had subjectivity not made its radical appearance, there would have been no knowing and no one to take notice, and consequently there would have been no history of what creatures did through the ages, no culture at all.

Although I have not yet provided a working definition of consciousness, I hope I am leaving no doubt as to what it means *not* to have consciousness: in the absence of consciousness, the personal view is suspended; we do not know of our existence; and we do not know that anything else exists. If consciousness had not developed in the course of evolution and expanded to its human version, the humanity we are now familiar with, in all its frailty and strength, would never have developed either. One shudders to think that a simple turn not taken might have meant the loss of the biological alternatives that make us truly human. But then, how would we ever have found out that something was missing?

We take consciousness for granted because it is so available, so easy to use, so elegant in its daily disappearing and reappearing acts, and yet, when we think of it, scientists and nonscientists alike, we do puzzle. What is consciousness made of? Mind with a twist, it seems to me, since we cannot be conscious without having a mind to be conscious of. But what is mind made of? Does mind come from the air or from the body? Smart people say it comes from the brain, that it *is* in the brain, but that is not a satisfactory reply. How does the brain *do* mind?

The fact that no one sees the minds of others, conscious or not, is especially mysterious. We can observe their bodies and their actions, what they do or say or write, and we can make informed guesses about what they think. But we cannot observe their minds, and only we ourselves can observe ours, from the inside, and through a rather narrow window. The properties of minds, let alone conscious minds, appear to be so radically different from those of visible living

matter that thoughtful folk wonder how one process (conscious minds working) meshes with the other process (physical cells living together in aggregates called tissues).

But to say that conscious minds are mysterious—and on the face of it they are—is different from saying that the mystery is insoluble. It is different from saying that we shall never be able to understand how a living organism endowed with a brain develops a conscious mind.¹

Goals and Reasons

This book is dedicated to addressing two questions. First: how does the brain construct a mind? Second: how does the brain make that mind conscious? I am well aware that addressing questions is not the same as answering them, and that on the matter of the conscious mind, it would be foolish to presume definitive answers. Moreover, I realize that the study of consciousness has expanded so much that it is no longer possible to do justice to all contributions being made to it. That, along with issues of terminology and perspective, make current work on consciousness resemble a walk through a minefield. Nonetheless, at one's own peril, it is reasonable to think through the questions and use the current evidence, incomplete and provisional as it is, to build testable conjectures and dream about the future. The goal of this book is to reflect on the conjectures and discuss a framework of hypotheses. The focus is on how the human brain needs to be structured and how it needs to operate in order for conscious minds to emerge.

Books should be written for a reason, and this one was written to start over. I have been studying the human mind and brain for more than thirty years, and I have previously written about consciousness in scientific articles and books.² But I have grown dissatisfied with my account of the problem, and reflection on relevant research findings, new and old, has changed my views, on two issues in particular: the origin and nature of feelings and the mechanisms behind the construction of the self. This book is an attempt to discuss the current views. In no small measure, the book is also about what we still do not know but wish we did.

The remainder of Chapter 1 situates the problem, explains the framework chosen to address it, and previews the main ideas that will emerge in the chapters ahead. Some readers may find that the long presentation in Chapter 1 slows down the reading, but I promise it will also make the rest of the book all the more accessible.

Approaching the Problem

Before we attempt to make some headway on the matter of how the human brain constructs a conscious mind, we need to acknowledge two important legacies. One of them consists of prior attempts to discover the neural basis of consciousness, in efforts that date back to the middle of the twentieth century. In a series of pioneering studies conducted in North America and Italy, a small band of investigators pointed with astonishing certainty to a brain sector that is now unequivocally related to the making of consciousness—the brain stem—and identified it as a critical contributor to consciousness. Not surprisingly, in light of what we know today, the account provided by these pioneers—Wilder Penfield, Herbert Jasper, Giuseppe Moruzzi, and Horace Magoun—was incomplete, and parts of it were less than correct. But one should have nothing but praise and admiration for the scientists who intuited the right target and aimed at it with such precision. This was the brave beginning of the enterprise to which several of us wish to contribute today.³

Also part of this legacy are studies performed more recently in neurological patients whose consciousness was compromised by focal brain damage. The work of Fred Plum and Jerome Posner launched the effort.⁴ Over the years these studies, complementing those of the consciousness-research pioneers, have yielded a powerful collection of facts regarding the brain structures that are or are not involved in making human minds conscious. We can build on that foundation.

The other legacy to be acknowledged consists of a long tradition of formulating conceptions of mind and consciousness. It has a rich history, as long and varied as the history of philosophy. From the wealth of its offerings, I have come to favor the writings of William James as an anchor for my own thinking, although this does not imply a full endorsement of his positions on consciousness and especially on feeling.⁵

The title of this book, as well as its first pages, leave no doubt that in approaching the conscious mind, I privilege the self. I believe conscious minds arise when a self process is added onto a basic mind process. When selves do not occur within minds, those minds are not conscious in the proper sense. This is a predicament faced by humans whose self process is suspended by dreamless sleep, anesthesia, or brain disease.

Defining the self process that I regard as so indispensable for consciousness, however, is easier said than done. That is why William James is so helpful to this preamble. James wrote eloquently about the importance of the self, and yet he also noted that, on many occasions, the presence of the self is so subtle that the contents of the mind dominate consciousness as they stream along. We need to confront this elusiveness and decide on its consequences before we go any

further. Is there a self, or is there not? If there is a self, is it present whenever we are conscious, or is it not?

The answers are unequivocal. There is indeed a self, but it is a process, not a thing, and the process is present at all times when we are presumed to be conscious. We can consider the self process from two vantage points. One is the vantage point of an observer appreciating a dynamic *object*—the dynamic object constituted by certain workings of minds, certain traits of behavior, and a certain history of life. The other vantage point is that of the self as *knower*, the process that gives a focus to our experiences and eventually lets us reflect on those experiences. Combining the two vantage points produces the dual notion of self used throughout the book. As we shall see, the two notions correspond to two stages of evolutionary development of the self, the self-as-knower having had its origin in the self-as-object. In everyday life each notion corresponds to a level of operation of the conscious mind, the self-as-object being simpler in scope than the self-as-knower.

From either vantage point, the process has varied scopes and intensities and its manifestations vary with the occasions. The self can operate on a subtle register, as “a hint half hinted” of the presence of a living organism,⁶ or on a salient register that includes personhood and identity for the owner of the mind. Now you sense it, now you don’t, but you always *feel* it, is my way of summing up the situation.

James thought that the self-as-object, the material me, was the sum total of all that a man could call his—“not only his body and his psychic powers, but his clothes and his wife and children, his ancestors and friends, his reputation and works, his lands and horses, and yacht and bank account.”⁷ Leaving aside the political incorrectness, I agree. But James also thought something else with which I am in even greater agreement: what allows the mind to know that such dominions exist and belong to their mental owners—body, mind, past and present, and all the rest—is that the perception of any of these items generates emotions and feelings, and, in turn, the feelings accomplish the separation between the contents that belong to the self and those that do not. From my perspective, such feelings operate as *markers*. They are the emotion-based signals I designate as somatic markers.⁸ When contents that pertain to the self occur in the mind stream, they provoke the appearance of a marker, which joins the mind stream as an image, juxtaposed to the image that prompted it. These feelings accomplish a distinction between self and nonself. They are, in a nutshell, *feelings of knowing*. We shall see that the construction of a conscious mind depends, at several stages, on the generation of such feelings. As for my working definition of the material me, the self-as-object, it is as follows: *a dynamic collection of integrated neural processes, centered on the representation of the living body, that finds expression in a dynamic collection of integrated mental processes.*

The self-as-subject, as knower, as the “I,” is a more elusive presence, far less collected in mental or biological terms than the *me*, more dispersed, often dissolved in the stream of consciousness, at times so annoyingly subtle that it is there but almost not there. The self-as-knower is more difficult to capture than the plain *me*, unquestionably. But that does not diminish its significance for consciousness. The self-as-subject-and-knower is not only a very real presence but a turning point in biological evolution. We can imagine that the self-as-subject-and-knower is stacked, so to speak, on top of the self-as-object, as a new layer of neural processes giving rise to yet another layer of mental processing. There is no dichotomy between self-as-object and self-as-knower; there is, rather, a continuity and progression. The self-as-knower is grounded on the self-as-object.

Consciousness is not merely about images in the mind. It is, in the very least, about an *organization of mind contents centered on the organism that produces and motivates those contents*. But consciousness, in the sense that reader and author can experience anytime they wish, is more than a mind organized under the influence of a living, acting organism. It is also a mind capable of knowing that such a living, acting organism exists. To be sure, the fact that the brain succeeds in creating neural patterns that map things experienced as images is an important part of the process of being conscious. Orienting the images in the perspective of the organism is also a part of the process. But that is not the same as automatically and explicitly *knowing* that images exist within me and are mine and, in current lingo, actionable. The mere presence of organized images flowing in a mental stream produces a mind, but unless some supplementary process is added on, the mind remains *unconscious*. What is missing from that unconscious mind is a *self*. What the brain needs in order to become conscious is to acquire a new property —*subjectivity*—and a defining trait of subjectivity is the feeling that pervades the images we experience subjectively. For a contemporary treatment of the importance of subjectivity from the perspective of philosophy, read John Searle’s *The Mystery of Consciousness*.⁹

In keeping with this idea, the decisive step in the making of consciousness is not the making of images and creating the basics of a mind. The decisive step is *making the images ours*, making them belong to their rightful owners, the singular, perfectly bounded organisms in which they emerge. In the perspective of evolution and in the perspective of one’s life history, the knower came in steps: the protoself and its primordial feelings; the action-driven core self; and finally the autobiographical self, which incorporates social and spiritual dimensions. But these are dynamic processes, not rigid things, and on any day their level fluctuates (simple, complex, somewhere in between) and can be readily adjusted as the circumstances dictate. A knower, by whatever name one may want to call it—self, experiencer, protagonist—needs to be generated in the brain if the mind is to become conscious. When the brain manages to introduce a

knower in the mind, subjectivity follows.

Should the reader wonder if this defense of the self is necessary, let me say that it is quite justified. At this very moment, those of us in neuroscience whose work aims at elucidating consciousness subscribe to very different attitudes toward the self. The attitudes range from considering the self as an indispensable topic of the research agenda to thinking that the time has not come to deal with the subject (literally!).¹⁰ Given that the work associated with either attitude continues to produce useful ideas, there is no need, as yet, to decide which approach will turn out to be more satisfactory. But we need to acknowledge that the resulting accounts are different.

In the meantime, it is noteworthy that these two attitudes perpetuate a difference of interpretation that separated William James from David Hume, one that is generally overlooked in such discussions. James wanted to make certain that his conceptions of self had a firm biological grounding: his “self” would not be mistaken for a metaphysical knowing agency. But that did not prevent him from recognizing a knowing function for the self, even when the function was subtle rather than exuberant. David Hume, on the other hand, pulverized the self to the point of doing away with it. The following passages illustrate Hume’s views: “I never can catch *myself* at any time without a perception and never can observe anything but the perception.” And further on: “I may venture to affirm of the rest of mankind, that they are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perceptual flux and movement.”

Commenting on Hume’s dismissal of the self, James was moved to issue a memorable rebuke and affirm the existence of the self, emphasizing the odd mixture of “unity and diversity” within it and calling attention to the “core of sameness” running through the ingredients of the self.¹¹

The foundation discussed here has been modified and expanded upon by philosophers and neuroscientists to include different aspects of self.¹² But the significance of the self for the construction of the conscious mind has not been diminished. I doubt that the neural basis for the conscious mind can be comprehensively elucidated without first accounting for the self-as-object—the material me—and for the self-as-knower.

Contemporary work on philosophy of mind and psychology has extended the conceptual legacy, while the extraordinary development of general biology, evolutionary biology, and neuroscience has capitalized on the neural legacy, produced a wide array of techniques to investigate the brain, and amassed a colossal amount of facts. The evidence, conjectures, and hypotheses presented in this book are grounded on all these developments.

The Self as Witness

Countless creatures for millions of years have had active minds, but only in those who developed a self capable of operating as a witness to the mind was its existence acknowledged, and only after minds developed language and lived to tell did it become widely known that minds did exist. The self as witness is the something extra that reveals the presence, in each of us, of events we call mental. We need to understand how that something extra is created.

The notions of witness and protagonist are not meant as mere literary metaphors. I hope they help illustrate the range of roles that the self assumes in the mind. For one thing, the metaphors can help us see the situation we face when we attempt to understand mental processes. A mind unwitnessed by a self protagonist is still a mind. However, given that the self is our only natural means to know the mind, we are entirely dependent on the self's presence, capabilities, and limits. And given this systematic dependence, it is extremely difficult to imagine the nature of the mind process independently of the self, although from an evolutionary perspective, it is apparent that plain mind processes preceded self processes. The self permits a view of the mind, but the view is clouded. The aspects of the self that permit us to formulate interpretations about our existence and about the world are still evolving, certainly at the cultural level and, in all likelihood, at the biological level as well. For instance, the upper reaches of self are still being modified by all manner of social and cultural interactions and by the accrual of scientific knowledge about the very workings of mind and brain. One entire century of movie viewing has certainly had an impact on the human self, as has the spectacle of globalized societies now instantly broadcast by electronic media. As for the impact of the digital revolution, it is just beginning to be appreciated. In brief, our only direct view of the mind depends on a part of that very mind, a self process that we have good reason to believe cannot provide a comprehensive and reliable account of what is going on.

At first glance, after acknowledging the self as our entry into knowledge, it may appear paradoxical, not to mention ungrateful, to question its reliability. And yet that is the situation. Except for the direct window that the self opens into our pains and pleasures, the information it provides must be questioned, most certainly when the information pertains to its very nature. The good news, however, is that the self also has made reason and scientific observation possible, and reason and science, in turn, have been gradually correcting the misleading intuitions prompted by the unaided self.

Overcoming a Misleading Intuition

It is arguable that cultures and civilizations would not have come to pass in the absence of consciousness, thus making consciousness a notable development in biological evolution. And yet the very nature of consciousness poses serious problems for those attempting to elucidate its biology. Viewing consciousness from where we stand today, mindful and armed with a self, can be blamed for an understandable but troubling distortion of the history of mind and consciousness studies. Viewed from the top, the mind acquires a special status, discontinuous with the remainder of the organism to which it belongs. Viewed from the bottom, the mind appears to be not just very complex, which it certainly is, but also different in kind from the biological tissues and functions of the organism that begets it. In practice, we adopt two sorts of optic when we observe our beings: we see the mind with eyes that are turned inward; and we see biological tissues with eyes that are turned outward. (To boot, we use microscopes to extend our vision.) Under the circumstances, it is not surprising that the mind appears to have a nonphysical nature and that its phenomena appear to belong to another category.

Viewing the mind as a nonphysical phenomenon, discontinuous with the biology that creates and sustains it, is responsible for placing the mind outside the laws of physics, a discrimination to which other brain phenomena are not usually subject. The most striking manifestation of this oddity is the attempt to connect the conscious mind to heretofore undescribed properties of matter and, for example, explain consciousness in terms of quantic phenomena. The rationale for this idea appears to be as follows: the conscious mind seems mysterious; because quantum physics remains mysterious, perhaps the two mysteries are connected.¹³

Given our incomplete knowledge of both biology and physics, one should be cautious before dismissing alternative accounts. After all, in spite of neurobiology's remarkable success, our understanding of the human brain is quite incomplete. Nonetheless, the possibility of explaining mind and consciousness parsimoniously, within the confines of neurobiology as currently conceived, remains open; it should not be abandoned unless the technical and theoretical resources of neurobiology are exhausted, an unlikely prospect at the moment.

Our intuition tells us that the mercurial, fleeting business of the mind lacks physical extension. I believe this intuition is false and attributable to the limitations of the unaided self. I see no reason to give to it more credence than to previously evident and powerful intuitions such as the pre-Copernican view of what the sun does to the earth or, for that matter, the view that the mind resides in the heart. Things are not always what they seem. White light is a composite of

the colors of the rainbow, although that is not apparent to the naked eye.¹⁴

An Integrated Perspective

Most of the progress made to date on the neurobiology of conscious minds has been based on combining three perspectives: (1) the direct-witness perspective on the individual conscious mind, which is personal, private, and unique to each one of us; (2) the behavioral perspective, which allows us to observe the telltale actions of others whom we have reason to believe also have conscious minds; and (3) the brain perspective, which allows us to study certain aspects of brain function in individuals whose conscious mind states are presumed to be either present or absent. Evidence from these three perspectives, even when intelligently aligned, is usually not enough to generate a smooth transition across the three kinds of phenomena—introspective, first-person inspection; external behaviors; and brain events. In particular, there appears to be a major gap between the evidence from first-person introspection and the evidence from brain events. How can we bridge such gaps?

A fourth perspective is needed, one that requires a radical change in the way the history of conscious minds is viewed and told. In earlier work I advanced the idea of turning life regulation into the support and justification of self and consciousness, and that idea suggested a path into this new perspective: a search for antecedents of self and consciousness in the evolutionary past.¹⁵ Accordingly, the fourth perspective is grounded on facts from evolutionary biology and neurobiology. It requires us to consider early living organisms first, then gradually move across evolutionary history toward current organisms. It requires us to note incremental modifications of nervous systems and link them to the incremental emergence of, respectively, behavior, mind, and self. It also requires an internal working hypothesis: that mental events are equivalent to certain kinds of brain events. Of course, mental activity is caused by the brain events that antecede it, but at the end of the day, the mental events correspond to certain states of brain circuits. In other words, some neural patterns *are* simultaneously mental images. When some other neural patterns generate a rich enough self process subject, the images can become *known*. But if no self is generated, the images still *are*, although no one, inside or outside the organism, knows of their existence. Subjectivity is not required for mental states to exist, only for them to be privately known.

In brief, the fourth perspective asks us to construct, simultaneously, with the help of available facts, a view from the past, and from within, literally an imagined view of a brain caught in the state of containing a conscious mind. To be sure, this is a conjectural, hypothetical view. There are facts to support parts of this imaginarium, but it is in the nature of the “mind-self-body-brain

problem” that we must live for quite a while with theoretical approximations rather than complete explanations.

It might be tempting to regard the hypothesized equivalence of mind events to certain brain events as a crude reduction of the complex to the simple. This would be a false impression, however, given that neurobiological phenomena are immensely complex to begin with, anything but simple. The explanatory reductions involved here are not from the complex to the simple but rather from the extremely complex to the slightly less so. Although this book is not about the biology of simple organisms, the facts to which I allude in Chapter 2 make it clear that the lives of cells occur in extraordinary complex universes that formally resemble, in many ways, our elaborate human universe. The world and behavior of a single-cell organism such as the paramecium are a wonder to behold, far closer to who we are than meets the eye.

It is also tempting to interpret the proposed brain-mind equivalence as a neglect of the role of culture in the generation of the mind or as a downgrading of the role of individual effort in the shaping of the mind. Nothing could be farther from my formulation, as will become clear.

Using the fourth perspective, I can now rephrase some of the statements made earlier in a way that takes into account facts from evolutionary biology and includes the brain: countless creatures for millions of years have had active minds happening in their *brains*, but only after those *brains* developed a protagonist capable of bearing witness did consciousness begin, in the strict sense, and only after those *brains* developed language did it become widely known that minds did exist. The witness is the something extra that reveals the presence of implicit *brain* events we call mental. Understanding how the brain produces that something extra, the protagonist we carry around and call self, or me, or I, is an important goal of the neurobiology of consciousness.

The Framework

Before I sketch the framework guiding this book, I need to introduce some basic facts. Organisms make minds out of the activity of special cells known as neurons. Neurons share most of the characteristics of other cells in our body, and yet their operation is distinctive. They are sensitive to changes around them; they are excitable (an interesting property they share with muscle cells). Thanks to a fibrous prolongation known as the axon, and to the end region of the axon known as the synapse, neurons can send signals to other cells—other neurons, muscle cells—often quite far away. Neurons are largely concentrated in a central nervous system (the brain, for short), but they send signals to the organism’s body, as well as to the outside world, and they receive signals from both.

The number of neurons in each human brain is on the order of billions, and the synaptic contacts that the neurons make among themselves number in the trillions. Neurons are organized in small microscopic circuits, whose combination constitutes progressively larger circuits, which in turn form networks or systems. For more on neurons and brain organization, see Chapter 2 and the Appendix.

Minds emerge when the activity of small circuits is organized across large networks so as to compose momentary patterns. The patterns represent things and events located outside the brain, either in the body or in the external world, but some patterns also represent the brain's own processing of other patterns. The term *map* applies to all those representational patterns, some of which are coarse, while others are very refined, some concrete, others abstract. In brief, the brain maps the world around it and maps its own doings. Those maps are experienced as *images* in our minds, and the term *image* refers not just to the visual kind but to images of any sense origin such as auditory, visceral, tactile, and so forth.

Let us now turn to the framework proper. Using the term *theory* to describe proposals for how the brain produces this or that phenomenon is somewhat out of place. Unless the scale is large enough, most theories are just hypotheses. What is being proposed in this book, however, is more than that, since it articulates several hypothetical components for one aspect or another of the phenomena I am addressing. What we hope to explain is too complex to be addressed by a single hypothesis and be accounted for by one mechanism. So I have settled for the term *framework* to designate the effort.

In order to qualify for the lofty title, the ideas presented in the chapters ahead need to accomplish certain goals. Given that we wish to understand how the brain makes the mind conscious, and given that it is manifestly impossible to deal with all levels of brain function in assembling an explanation, the framework must specify the level at which the explanation applies. This is the large-scale systems level, the level at which macroscopic brain regions constituted by neuron circuits interact with other such regions to form systems. Of necessity, those systems are macroscopic, but the underlying microscopic anatomy is known in part, as are the general operating rules of the neurons that constitute them. The large-scale systems level is amenable to research via numerous techniques, old and new. They include the modern version of the lesion method (which relies on the study of neurological patients with focal brain damage investigated with structural neuroimaging and experimental cognitive and neuropsychological techniques); functional neuroimaging (based on magnetic resonance scanning, positron-emission tomography, magnetoencephalography, and assorted electrophysiological techniques); direct neurophysiological recording of neuron activity in the setting of neurosurgical treatments; and transcranial magnetic stimulation.

The framework must interconnect behavior, mind, and brain events. On this second goal, the framework aligns behavior, mind, and brain closely; and because it relies on evolutionary biology, it places consciousness in a historical setting, a placement suitable to organisms undergoing evolutionary transformation by natural selection. Moreover, the maturation of neuron circuitries in each brain is also seen as subject to selection pressures resulting from the very activity of organisms and the processes of learning. The repertoires of neuron circuitries initially provided by the genome are changed accordingly.¹⁶

The framework indicates the placement of regions involved in mind-making, at whole-brain scale, and proposes how some brain regions might operate in concert to produce the self. It suggests how a brain architecture that features convergence and divergence of neuron circuitries plays a role in the high-order coordination of images and is essential for the construction of the self and of other aspects of mental function, namely memory, imagination, language, and creativity.

The framework needs to break down the phenomenon of consciousness in components amenable to neuroscience research. The result is two researchable domains, namely, mind processes and self processes. Furthermore, it decomposes the self process into subtypes. The latter separation offers two advantages: presuming and investigating consciousness in species that are likely to have self processes albeit less elaborate; and creating a bridge between the high levels of self and the sociocultural space in which humans operate.

Another goal: the framework must address the issue of how system macroevents are built from microevents. Here the framework hypothesizes the equivalence of mental states to certain states of regional brain activity. The framework assumes that when certain ranges of intensity and frequency of neuron firing occur in small neuron circuits, when some of these circuits are synchronously activated, and when certain conditions of network connectivity are met, the result is a “mind with feelings.” In other words, as a result of the growing size and complexity of neural networks, there is a scaling up of “cognition” and “feeling,” from the microlevel to the macrolevel, across hierarchies. The model for this scaling up to mind with feeling can be found in the physiology of movement. The contraction of a single microscopic muscle cell is a negligible phenomenon, whereas the simultaneous contraction of large numbers of muscle cells can produce visible movement.

A Preview of Main Ideas

I

Of the ideas advanced in the book, none is more central than the notion that the body is a foundation of the conscious mind. We know that the most stable aspects of body function are represented in the brain, in the form of maps, thereby contributing images to the mind. This is the basis of the hypothesis that the special kind of mental images of the body produced in body-mapping structures, constitutes the *protoself*, which foreshadows the self to be. Of note, the critical body-mapping and image-making structures are located below the level of the cerebral cortex, in a region known as the upper brain stem. This is an old part of the brain shared with many other species.

II

Another central idea is based on the consistently overlooked fact that the brain's protoself structures are not merely *about* the body. They are literally and inextricably *attached* to the body. Specifically, they are attached to the parts of the body that bombard the brain with their signals, at all times, only to be bombarded back by the brain and, by so doing, creating a resonant loop. This resonant loop is perpetual, broken only by brain disease or death. Body and brain *bond*. As a result of this arrangement, the protoself structures have a privileged and direct relationship to the body. The images they engender regarding the body are conceived in circumstances different from those of other brain images, say, visual or auditory. In light of these facts, the body is best conceived as the rock on which the protoself is built, while the protoself is the pivot around which the conscious mind turns.

III

I hypothesize that the first and most elementary product of the protoself is *primordial feelings*, which occur spontaneously and continuously whenever one is awake. They provide a direct experience of one's own living body, wordless, unadorned, and connected to nothing but sheer existence. These *primordial feelings* reflect the current state of the body along varied dimensions, for example, along the scale that ranges from pleasure to pain, and they originate at the level of the brain stem rather than the cerebral cortex. All feelings of

emotion are complex musical variations on primordial feelings.¹⁷

In the functional arrangement outlined here, pain and pleasure are body events. The events are *also* mapped in a brain that at no instant is separated from its body. Thus primordial feelings are a special kind of image generated thanks to the obligate body-brain interaction, to the characteristics of the circuitry accomplishing the connection, and possibly to certain properties of neurons. It is not enough to say that feelings are felt because they map the body. I hypothesize that in addition to holding a unique relationship to the body, the brain-stem machinery responsible for making the kinds of images we call feelings is capable of richly mixing signals from the body and thus creating complex states with the special and novel properties of feeling rather than mere slavish maps of the body. The reason why nonfeeling images are also felt is that they are normally *accompanied* by feelings.

The foregoing implies that the notion of a sharp border separating body and brain is problematic. It also suggests a potentially fruitful approach to the vexing problem of why and how normal mental states are invariably imbued with some form of feeling.

IV

Brains begin building conscious minds not at the level of the cerebral cortex but rather at the level of the brain stem. Primordial feelings are not only the first images generated by the brain but also immediate manifestations of sentience. They are the protoself foundation for more complex levels of self. These ideas run counter to widely accepted views, although Jaak Panksepp (cited earlier) has defended a comparable position and so has Rodolfo Llinás. But the conscious mind as we know it is a far different affair from the conscious mind that emerges in the brain stem, and on this point there probably is universal agreement. The cerebral cortices endow the mind-making process with a profusion of images that, as Hamlet might put it, go far beyond anything that poor Horatio could ever dream of, in heaven or earth.

Conscious minds begin when self comes to mind, when brains add a self process to the mind mix, modestly at first but quite robustly later. The self is built in distinct steps grounded on the *protoself*. The first step is the generation of primordial feelings, the elementary feelings of existence that spring spontaneously from the protoself. Next is the *core self*. The core self is about action—specifically, about a relationship between the organism and the object. The core self unfolds in a sequence of images that describe an object engaging the protoself and modifying that protoself, including its primordial feelings. Finally, there is the *autobiographical self*. This self is defined in terms of biographical knowledge pertaining to the past as well as the anticipated future.

The multiple images whose ensemble defines a biography generate pulses of core self whose aggregate constitutes an autobiographical self.

The protoself with its primordial feelings, and the core self, constitute a “material me.” The autobiographical self, whose higher reaches embrace all aspects of one’s social persona, constitute a “social me” and a “spiritual me.” We can observe these aspects of self within our own minds or study their effects in the behavior of others. In addition, however, the core and autobiographical selves within our minds construct a knower; in other words, they endow our minds with another variety of subjectivity. For practical purposes, normal human consciousness corresponds to a mind process in which all of these self levels operate, offering to a limited number of mind contents a momentary link to a pulse of core self.

V

At neither modest nor robust levels do self and consciousness *happen* in one area or region or center of the brain. Conscious minds result from the smoothly articulated operation of several, often many, brain sites. The key brain structures in charge of implementing the requisite functional steps include specific sectors of the upper brain stem, a set of nuclei in a region known as the thalamus, and specific but widespread regions of the cerebral cortex.

The ultimate consciousness product occurs *from* those numerous brain sites at the same time and not in one site in particular, much as the performance of a symphonic piece does not come from the work of a single musician or even from a whole section of an orchestra. The oddest thing about the upper reaches of a consciousness performance is the conspicuous absence of a conductor *before* the performance begins, although, as the performance unfolds, a conductor comes into being. For all intents and purposes, a conductor is now leading the orchestra, although the performance has created the conductor—the self—not the other way around. The conductor is cobbled together by feelings and by a narrative brain device, although this fact does not make the conductor any less real. The conductor undeniably exists in our minds, and nothing is gained by dismissing it as an illusion.

The coordination on which conscious minds depend is achieved by a variety of means. At the modest core level, it begins quietly, as a spontaneous assembly of images that emerge one after the other in close time proximity, the image of an object, on the one hand, and the image of the protoself changed by the object, on the other. No additional brain structures are needed for a core self to emerge, at this simple level. The coordination is natural, sometimes resembling a mere musical duo, played by organism and object, sometimes resembling a chamber music ensemble, and in both cases managing quite well without a conductor. But

when the contents being processed in the mind are more numerous, other devices are required to accomplish coordination. In that case a variety of brain regions below the level of the cerebral cortices and within them play a key role.

Building a mind capable of encompassing one's lived past and anticipated future, along with the lives of others added to the fabric and a capacity for reflection to boot, resembles the execution of a symphony of Mahlerian proportions. But the marvel, as hinted at earlier, is that the score and the conductor become reality only as life unfolds. The coordinators are not mythical, sapient homunculi in charge of interpreting anything. And yet the coordinators do help with the assembly of an extraordinary media universe and with the placement of a protagonist in its midst.

The grand symphonic piece that is consciousness encompasses the foundational contributions of the brain stem, forever hitched to the body, and the wider-than-the-sky imagery created in the cooperation of cerebral cortex and subcortical structures, all harmoniously stitched together, in ceaseless forward motion, interruptible only by sleep, anesthesia, brain dysfunction, or death.

No single mechanism explains consciousness in the brain, no single device, no single region, or feature, or trick, any more than a symphony can be played by one musician or even a few. Many are needed. What each of them contributes does count. But only the ensemble produces the result we seek to explain.

VI

Managing and safekeeping life efficiently are two of the recognizable achievements of consciousness: neurological patients whose consciousness is compromised are unable to manage their lives independently even when their basic life functions operate normally. And yet mechanisms for managing and maintaining life are not a novelty in biological evolution and are not necessarily dependent on consciousness. Such mechanisms already exist in single cells and are coded in their genome. They are also widely replicated within ancient, humble, *un-minded* and *un-conscious* neuron circuits, and they are very much present deep in human brains. We shall see that managing and safekeeping life is the fundamental premise of biological value. Biological value has influenced the evolution of brain structures, and in any brain it influences almost every step of brain operations. It is expressed as simply as in the release of chemical molecules related to reward and punishment, or as elaborately as in our social emotions and in sophisticated reasoning. Biological value naturally guides and colors, so to speak, almost everything that happens inside our very minded, very conscious brains. Biological value has the status of a principle.

In brief, the conscious mind emerges within the history of life regulation. Life

regulation, a dynamic process known as *homeostasis* for short, begins in unicellular living creatures, such as a bacterial cell or a simple amoeba, which do not have a brain but are capable of adaptive behavior. It progresses in individuals whose behavior is managed by simple brains, as is the case with worms, and it continues its march in individuals whose brains generate both behavior and mind (insects and fish being examples). I am ready to believe that whenever brains begin to generate primordial feelings—and that could be quite early in evolutionary history—organisms acquire an early form of sentience. From there on, an organized self process could develop and be added to the mind, thereby providing the beginning of elaborate conscious minds. Reptiles are contenders for this distinction, for example; birds make even stronger contenders; and mammals get the award and then some.

Most species whose brains generate a self do so at core level. Humans have both core self and autobiographical self. A number of mammals are likely to have both as well, namely wolves, our ape cousins, marine mammals and elephants, cats, and, of course, that off-the-scale species called the domestic dog.

VII

The march of mind progress does not end with the arrival of the modest levels of self. Throughout the evolution of mammals, especially primates, minds become ever more complex, memory and reasoning expanding notably, and the self processes enlarge their scope. The core self remains, but it is gradually surrounded by an autobiographical self, whose neural and mental natures are very different from those of the core self. We become able to use a part of our mind's operation to monitor the operation of other parts. The conscious minds of humans, armed with such complex selves and supported by even greater capabilities of memory, reasoning, and language, engender the instruments of culture and open the way into new means of homeostasis at the level of societies and culture. In an extraordinary leap, homeostasis acquires an extension into the sociocultural space. Justice systems, economic and political organizations, the arts, medicine, and technology are examples of the new devices of regulation.

The dramatic reduction of violence along with the increase in tolerance that has become so apparent in recent centuries would not have occurred without sociocultural homeostasis. Neither would the gradual transition from coercive power to the power of persuasion that hallmarks advanced social and political systems, their failures notwithstanding. The investigation of sociocultural homeostasis can be informed by psychology and neuroscience, but the native space of its phenomena is cultural. It is reasonable to describe those who study the rulings of the U.S. Supreme Court, the deliberations of the U.S. Congress, or the workings of financial institutions as engaging, indirectly, in studying the

vagaries of sociocultural homeostasis.

Both basic homeostasis (which is nonconsciously guided) and sociocultural homeostasis (which is created and guided by reflective conscious minds) operate as curators of biological value. Basic and sociocultural varieties of homeostasis are separated by billions of years of evolution, and yet they promote the same goal—the survival of living organisms—albeit in different ecological niches. That goal is broadened, in the case of sociocultural homeostasis, to encompass the *deliberate* seeking of well-being. It goes without saying that the way in which human brains manage life requires both varieties of homeostasis in continuous interaction. But while the basic variety of homeostasis is an established inheritance, provided by everyone's genome, the sociocultural variety is a somewhat fragile work in progress, responsible for much of human drama, folly, and hope. The interaction between these two kinds of homeostasis is not confined to each individual. There is growing evidence that, over multiple generations, cultural developments lead to changes in the genome.

VIII

Viewing the conscious mind in the optic of evolution from simple life-forms toward complex and hypercomplex organisms such as ours helps naturalize the mind and shows it to be the result of stepwise progressions of complexity within the biological idiom.

We can look at human consciousness and at the functions it made possible (language, expanded memory, reasoning, creativity, the whole edifice of culture) as the curators of value inside our modern, very minded, very social beings. And we can imagine a long umbilical cord that links the barely weaned, perennially dependent conscious mind to the depths of very elementary and very *un*-conscious regulators of the value principle.

The history of consciousness cannot be told in the conventional way. Consciousness came into being because of biological value, as a contributor to more effective value management. But consciousness did not *invent* biological value or the process of valuation. Eventually, in human minds, consciousness revealed biological value and allowed the development of new ways and means of managing it.

Life and the Conscious Mind

Is it reasonable to devote a book to the question of how brains make conscious minds? It is sensible to ask if understanding the brain work behind mind and self

has any practical significance besides satisfying our curiosity about human nature. Does it make any difference in daily life? For many reasons, large and small, I think it does. Brain science and its explanations are not about to provide for all people the satisfaction that so many obtain from experiencing the arts or cultivating spiritual beliefs. But there certainly are other compensations.

Understanding the circumstances in which conscious minds emerged in the history of life, and specifically how they developed in human history, allows us to judge perhaps more wisely than before the quality of the knowledge and advice those conscious minds provide. Is the knowledge reliable? Is the advice sound? Do we gain from understanding the mechanisms behind the minds that give us counsel?

Elucidating the neural mechanisms behind conscious minds reveals that our selves are not always sound and that they are not in control of every decision. But the facts also authorize us to reject the false impression that our ability to deliberate consciously is a myth. Elucidating conscious as well as nonconscious mind processes increases the possibility of fortifying our deliberative powers. The self opens the way for deliberation and for the adventure of science, two specific tools with which all the misleading guidance of the unaided self can be countered.

The time will come when the issue of human responsibility, in general moral terms as well as on matters of justice and its application, will take into account the evolving science of consciousness. Perhaps the time is now. Armed with reflexive deliberation and scientific tools, an understanding of the neural construction of conscious minds also adds a welcome dimension to the task of investigating the development and shaping of cultures, the ultimate product of collectives of conscious minds. As humans debate the benefits or perils of cultural trends, and of developments such as the digital revolution, it may help to be informed about how our flexible brains create consciousness. For example, will the progressive globalization of human consciousness brought on by the digital revolution retain the goals and principles of basic homeostasis, as current sociocultural homeostasis does? Or will it break away from its evolutionary umbilical cord, for better or worse?¹⁸

Naturalizing the conscious mind and planting it firmly in the brain does not diminish the role of culture in the construction of human beings, does not reduce human dignity, and does not mark the end of mystery and puzzlement. Cultures arise and evolve from collective efforts of human brains, over many generations, and some cultures even die in the process. They require brains that have already been shaped by prior cultural effects. The significance of cultures to the making of the modern human mind is not in question. Nor is the dignity of that human mind diminished by connecting it to the astonishing complexity and beauty to be found inside living cells and tissues. On the contrary, connecting personhood to biology is a ceaseless source of awe and respect for anything human. Last,

naturalizing the mind may solve one mystery but only to raise the curtain on other mysteries quietly awaiting their turn.

Placing the construction of conscious minds in the history of biology and culture opens the way to reconciling traditional humanism and modern science, so that when neuroscience explores human experience into the strange worlds of brain physiology and genetics, human dignity is not only retained but reaffirmed.

F. Scott Fitzgerald wrote memorably, “His was a great sin who first invented consciousness.” I can understand why he said so, but his condemnation is only half the story, appropriate for moments of discouragement with the imperfections of nature that conscious minds expose so nakedly. The other half of the story should be occupied with praise for such an invention as the enabler of all the creations and discoveries that trade loss and grief for joy and celebration. The emergence of consciousness opened the way to a life worth living. Understanding how it comes about can only strengthen that worth.¹⁹

Does knowing about how the brain works matter at all for how we live our lives? I believe it matters very much, all the more so if, besides knowing who we presently are, we care at all for what we may become.