

CHAPTER ONE

Introduction to the Study of Consciousness

Consciousness is what makes the mind-body problem really intractable...
Without consciousness the mind-body problem would be much less
interesting. With consciousness it seems hopeless.

From *What Is It Like to Be a Bat?* by Thomas Nagel

In Thomas Mann's unfinished novel, *Confessions of Felix Krull, Confidence Man*, Professor Kuckuck comments to the Marquis de Venosta on the three fundamental and mysterious stages of creation. Foremost is the creation of something—namely, the universe—out of nothing. The second act of genesis is the one that begat life from inorganic, dead matter. The third mysterious act is the birth of consciousness¹ and conscious beings out of organic matter, beings that can reflect upon themselves. Humans and at least some animals not only detect light, move their eyes, and perform other actions, but also have “feelings” associated with these events. This remarkable feature of the world cries out for an explanation. Consciousness remains one of the key puzzles confronting the scientific worldview.

1.1 | WHAT NEEDS TO BE EXPLAINED?

Throughout recorded history, men and women have wondered how we can see, smell, reflect upon ourselves, and remember. How do these sensations arise? The fundamental question at the heart of the mind-body problem is, *what*

¹The word consciousness derives from the Latin *conscientia*, composed of *cum* (with or together) and *scire* (to know). Until the early 17th century, consciousness was used in the sense of moral knowledge of right or wrong, what is referred to today as *conscience*.

*is the relation between the conscious mind and its physical basis in the electro-chemical interactions in the body?*² How do the salty taste and crunchy texture of potato chips, the unmistakable smell of dogs after they have been in the rain, or the feeling of hanging on tiny fingerholds on a wall a couple of meters above the last secure foothold, emerge from networks of neurons? These sensory qualities, the building blocks of conscious experience, have traditionally been called *qualia*. The puzzle is, how can a physical system have qualia?

Furthermore, why is a particular quale the way it is and not different? How come red feels the way it does, quite distinct from the sensation of seeing blue? These are not abstract, arbitrary symbols; they represent something *meaningful* to the organism. Philosophers talk about the mind's capacity to represent or to be about things. How meaning arises from electrical activity in the vast neural networks making up the brain remains a deep mystery. The structure of these networks, their connectivity, surely plays a role, but how so?³

How is it that humans and many animals have experiences? Why can't people live, beget, and raise children without consciousness? From a subjective vantage point, this would resemble not being alive at all, like sleepwalking through life. Why, then, from the point of view of evolution, does consciousness exist? What survival value is attached to subjective, mental life?

In Haitian lore, a zombie is a dead person who, by the magical power of a sorcerer, must act out the wishes of the person controlling him. In philosophy, a *zombie* is an imaginary being who behaves and acts just like a normal person, but has absolutely no conscious life, no sensations, and no feelings. A particularly insidious zombie will even lie, claiming that she is experiencing something when she is not.

The fact that it is so difficult to imagine such a scenario is living proof of the fundamental importance of consciousness to daily life. Following René Descartes's famous remark—made in the context of establishing his existence—I can ascertain with certainty that “I am conscious.” Not always,

²No consensual usage of objective and subjective terms has emerged across disciplines. I adopt the following convention throughout the book: *detection* and *behavior* are objective terms that can be operationalized (see, Dennett, 1991), as in “the retina detects the red flash, and the observer presses her finger in response.” Detection and behavior can occur in the absence of consciousness. I use *sensation*, *perception*, *seeing*, *experience*, *mind*, and *feeling* in their subjective senses, as in “conscious sensation” and so on. While I'm on the topic of convention, one last one. Throughout the book, I use *awareness* and *consciousness* (or *aware* and *conscious*) as synonyms. Some scholars distinguish between these two on ontological (Chalmers, 1996), conceptual (Block, 1995), or psychological (Tulving, 1993) grounds. At this point, little empirical evidence justifies such a distinction (see, however, Lamme, 2003). I might have to revise this standpoint in the future. Curiously, the contemporary scientific literature discourages the usage of the word consciousness, while awareness is acceptable. This is more a reflection of sociological trends than deep insight.

³The exact relationship between qualia and meaning intentionality is unclear (see the anthology by Chalmers, 2002).

not in a dreamless sleep or while under anesthesia, but often: when I read, talk, climb, think, discuss, or just sit and admire the beauty of the world.⁴

The mystery deepens with the realization that much of what goes on in the brain bypasses consciousness. Electrophysiological experiments prove that furious activity in legions of neurons can fail to generate a conscious percept or memory. In a reflex action, you will instantly and vigorously shake your foot if you detect an insect crawling over it, even though the realization of what is happening only comes later on. Or your body reacts to a fearful sight, a spider or gun, before it's been consciously registered: The palms become sweaty, the heartbeat and blood pressure increases, and adrenaline is released. And all this before you know that you are afraid and why. Many relatively complex sensory-motor behaviors are similarly rapid and nonconscious. Indeed, the point of training is to teach our bodies to quickly execute a complex series of movements—returning a serve, evading a punch, or tying shoelaces—without thinking about it. Nonconscious processing extends to the highest echelons of the mind. Sigmund Freud argued that childhood experiences—often of a traumatic nature—can profoundly determine adult behavior without any conscious insight into this influence. Much high-level decision making and creativity occurs without conscious thought, a topic treated in more depth in Chapter 18.

So much of what constitutes the ebb and flow of daily life takes place beyond the pale of consciousness. Some of the best evidence for this comes from the clinic. Consider the strange case of the neurological patient D.F. She is unable to see shapes or recognize pictures of everyday objects, yet can catch a ball. Even though she can't tell the orientation of a thin mail box-like slit (is it horizontal?) she can deftly post a letter into the slit. By studying such patients, neuropsychologists have inferred the existence of *zombie agents* in the brain that bypass awareness; that is, they don't involve consciousness (recall that in the second footnote to this chapter, I equate awareness with consciousness). These agents are dedicated to stereotypical tasks, such as shifting the eyes or positioning the hand. They usually operate fairly rapidly and don't have access to explicit memory. I'll return to these themes in Chapters 12 and 13.

Why, then, isn't the brain just a large collection of specialized, zombie agents? Life might be boring if it were, but since such agents work effort-

⁴Strictly speaking, I don't know whether you are conscious or not. You might even be a zombie! However, because you act and speak just as I do, because your brain is similar to mine, and because you and I share the same evolutionary heritage, I sensibly assume that you are also conscious. At present, the proper scientific understanding of consciousness is insufficient to prove this, but everything about the natural world appears compatible with this assumption. *Solipsism* denies this and argues that only the subject himself is truly conscious while everybody else is a zombie. This seems implausible and also rather arbitrary. After all, why should I, out of all the people in the world, be singled out for consciousness?

lessly and rapidly, why is consciousness needed at all? What is its function? In Chapter 14, I argue that consciousness gives access to a general-purpose and deliberate processing mode for planning and contemplating a future course of action.

Consciousness is an intensely private matter. A sensation cannot be directly conveyed to somebody else but is usually circumscribed in terms of other experiences. Try to explain your experience of seeing red. You'll end up relating it to other percepts, such as "red as a sunset" or "as a Chinese flag" (this task becomes next to impossible when communicating to a person blind from birth). You can talk meaningfully about the relationships among different experiences but not about any single one. This too needs to be explained.

This, then, is the charter for our quest: To understand how and why the neural basis of a specific conscious sensation is the basis of that sensation rather than another, and rather than a completely nonconscious state; why sensations are structured the way they are, how they acquire meaning, and why they are private; and, finally, how and why so many behaviors occur in the absence of consciousness.

1.2 | A SPECTRUM OF ANSWERS

Philosophers and scientists have pondered the mind-body problem in its present form since the publication of Descartes's *Traité de l'homme* in the mid-17th century. Until the 1980s, however, the vast majority of work in the brain sciences made no references to consciousness. In the last two decades, philosophers, psychologists, cognitive scientists, clinicians, neuroscientists, and even engineers have published dozens of monographs and books aimed at "discovering," "explaining," or "reconsidering" consciousness. Much of this literature is either purely speculative or lacks any detailed scientific program for systematically discovering the neuronal basis of consciousness and, therefore, does not contribute to the ideas discussed in this book.

Before introducing the approach Francis Crick and I have taken to address these problems, I will survey the philosophical landscape to familiarize readers with some of the possible categories of answers that people have considered. Keep in mind that only cartoon-like pocket sketches of these positions are provided here.⁵

⁵I can't possibly do justice to the sophisticated nature of these arguments. Anyone interested in all the subtle twists and turns is urged to consult the philosophical anthologies by Block, Flanagan, and Güzeldere (1997) and by Metzinger (1995). The textbook by the philosopher Patricia Churchland (2002) overviews contemporary mind-body problems with an emphasis on the relevant neuroscience. I also recommend the compact and readable monograph by Searle (1997). For the reverberations of these discussions among theologians, see Brown, Murphy, and Malony (1998) and the thoughtful McMullin (2000).

Consciousness Depends on an Immaterial Soul

Plato, the patriarch of Western philosophy, is widely credited with the concept of a person as an immortal soul imprisoned in a mortal body. He also proposed that ideas have a real existence and are eternal. These Platonic views were subsequently absorbed into the New Testament and form the basis of the classical Roman Catholic doctrine of a *soul*. The belief that at the heart of consciousness lies a transcendent and immortal soul is widely shared by many religions and faiths throughout the world.⁶

In modern times, Descartes distinguished between *res extensa*—such as the animal spirits that run through nerves and fill the muscles, substances with spatial extent—and *res cogitans*, thinking substances. He argued that *res cogitans* are unique to humans and give rise to consciousness. Descartes's ontological division constitutes the very definition of *dualism*. Weaker forms of dualism had been proposed earlier by Aristotle and by Thomas Aquinas. The most famous modern defenders of dualism are the philosopher Karl Popper and the neurophysiologist and Nobel laureate John Eccles.

While logically consistent, strong dualist positions are dissatisfying from a scientific viewpoint. Particularly troublesome is the mode of interaction between the soul and the brain. How and where is this supposed to take place? Presumably, this interaction would have to be compatible with the laws of physics. This, however, would require an exchange of energy that needs to be accounted for. And what happens to this spooky substance, the soul, once its carrier, the brain, dies? Does it float around in some hyperspace, like a ghost?⁷

The concept of an immaterial essence can be saved by postulating that the soul is immortal and completely independent of the brain. This leaves it as something ineffable, undetectable, a “ghost in the machine,” to use a phrase coined by Gilbert Ryle, outside of science.

⁶Being raised in a devout Roman Catholic family, I have much sympathy for this point of view. Flanagan's book (2002) explores the clash between the notion of soul (and free will) and the modern scientific view that tends to deny both (see also Murphy, 1998).

⁷Popper and Eccles (1977) argued that brain-soul interactions are camouflaged by Heisenberg's uncertainty principle, according to which it is impossible to know precisely the position and the momentum of a microscopic system, such as an electron, at the same time. In 1986, Eccles postulated that the conscious mind interferes with the release probability of vesicles at synapses in a way that does not violate conservation of energy yet is sufficient to influence the brain's behavior. These ideas have not been received with enthusiasm by the scientific community. Yet what is refreshing about the Popper and Eccles (1977) monograph is that they take consciousness seriously. They assume that sensations are a product of evolution that cries out for some function (see, in particular, Eccles, 1988). This is a remarkable sentiment after so many decades of behaviorism that disregarded consciousness entirely.

Consciousness Cannot Be Understood by Scientific Means

Quite a different philosophical tradition is the *mysterian*⁸ position, which claims that human beings are unable to comprehend consciousness because it is just too complex. This limitation is either a principled, formal one (how can any system completely understand itself?) or a practical one, expressed as a pessimism about the human mind's inability to perform the necessary massive conceptual revisions (what chance does an ape have of understanding general relativity?).

Other philosophers assert that they don't see how the physical brain can give rise to consciousness. Therefore, any scientific program to explore the physical basis of consciousness is doomed to failure. This is an argument from ignorance: The current absence of a compelling argument for a link between the brain and the conscious mind cannot be taken as evidence that such a link does not exist. Of course, to answer these critics, science will have to come up with the relevant concepts and evidence to support this link.

Although scientists may never fully comprehend—even in principle, let alone in practice—the workings of brains and the genesis of consciousness, it is premature to conclude so now. Neuroscience is a young discipline, accumulating new knowledge with ever-more-refined methods at a breathtaking pace. Before much of this development has run its course, there is no reason to come to this defeatist conclusion. Just because one particular scholar is unable to understand how consciousness might arise does not mean that it must be beyond all human comprehension!

Consciousness Is Illusory

Another type of philosophical reaction to the mind-body dilemma is to deny that there is any real problem at all. The most lively contemporary exponent of this rather counterintuitive notion—originating in the behaviorist tradition—is the philosopher Daniel Dennett from Tufts University. In *Consciousness Explained*, he argues that consciousness as most people conceive of it is an elaborate illusion, mediated by the senses in collusion with motor output, and supported by social constructions and learning. While acknowledging that people claim that they are conscious and that this persistent, but erroneous, belief needs to be explained, he denies the inner reality of the ungraspable aspects of qualia. He thinks that the usual way of thinking about consciousness is wildly wrong. Dennett seeks to explain the *third-person account* of con-

⁸The term *mysterian* originates with Flanagan (1992), who used it to characterize the approaches of Lucas (1961), Nagel (1974), and McGinn (1991).

consciousness while rejecting those aspects of the *first-person account* that render it resistant to reduction.⁹

Having dental pain is about expressing, or wanting to express, certain behaviors: To want to stop chewing on that side of the mouth, to run away and hide until the pain has subsided, to grimace, and so on. These “reactive dispositions,” as he calls them, are real. But not the badness of the pain. That elusive feeling doesn’t exist.¹⁰

Given the centrality of subjective feelings to everyday life, it would require extraordinary factual evidence before concluding that qualia and feelings are illusory. Philosophical arguments, based on logical analysis coupled to introspection, are not powerful enough to deal with the real world with all of its subtleties in a decisive manner. The philosophical method is at its best when formulating questions, but does not have much of a track record at answering them. The provisional approach I take in this book is to consider first-person accounts as brute facts of life and seek to explain them.¹¹

Consciousness Requires Fundamentally New Laws

Some have called for new scientific laws to explain the puzzle of consciousness, rather than just more facts and principles about the brain. Roger Penrose, at Oxford University, argues in the wonderful *The Emperor’s New Mind* that present-day physics is incapable of dealing with the intuition of mathematicians—and, by extension, people at large. Penrose believes that a yet-to-be-formulated theory of quantum gravity will explain how human consciousness can carry out processes that no possible digital (Turing) computer could implement. In conjunction with the anesthesiologist Stuart Hameroff,

⁹A third-person account recognizes only objective events, such as light of a certain wavelength impinging upon the retina, causing the person to exclaim “I see red,” while the first-person account is concerned with subjective events, such as the sensation of red. The late Francisco Varela labeled the program of mapping first-person experiences onto the brain *neuropsychophenomenology* (Varela, 1996).

¹⁰I refer the reader to Dennett’s book (1991), and to Dennett and Kinsbourne (1992). See Ryle (1949) for an antecedent in the behavioral tradition. For an update on his views, consult Dennett (2001). In his 1991 book, Dennett rightly takes aim at the notion of a *Cartesian theater*, a single place in the brain where conscious perception must occur (note that this does not exclude the possibility of a distributed set of neuronal processes that express consciousness at any one point in time). He proposes a *multiple drafts* model to account for various puzzling aspects of consciousness, such as the nonintuitive role of time in the organization of experience. Dennett’s writing is characterized by his skillful use of colorful metaphors and analogies, of which he is overly fond. It is difficult to relate these to specific neuronal mechanisms.

¹¹These are deep waters. Dennett retorts that innocently accepting feelings as facts to be explained is giving a hostage to fortune; that to talk about real qualia is a highly ideological move akin to presupposing the existence of “real magic,” full of epistemological implications (Dennett, 2003).

at the University of Arizona at Tucson, Penrose has proposed that microtubules, self-assembling cytoskeletal proteins found throughout all cells in the body, are critically involved in mediating coherent quantum states across large populations of neurons.¹²

While Penrose has generated a vigorous debate regarding the extent to which mathematicians can be said to have access to certain noncomputable truths and whether these can be instantiated by computers, it remains utterly mysterious how quantum gravity could explain how consciousness occurs in certain classes of highly organized matter. Both consciousness and quantum gravity have enigmatic features, but to conclude that one is therefore the cause of the other seems rather arbitrary. Given the lack of any positive evidence for macroscopic quantum-mechanical effects occurring in the brain, I will not pursue this idea further.

The philosopher David Chalmers, at the University of Arizona at Tucson, has sketched an alternative proposal in which information has two aspects, one physically realizable that is used in computers, and a phenomenal or experiential aspect that is inaccessible from the outside. In his view, any information-processing system, from a thermostat to a human brain, can be conscious in at least some rudimentary sense (although Chalmers admits that it probably doesn't feel like much "to be a thermostat"). While the audacity of endowing all systems that represent information with experience has a certain appeal and elegance, it is not clear to me how Chalmers's hypothesis could be tested scientifically. For now, this modern-day *panpsychism* can only be accepted as a belief. Over time, though, a theory couched in the language of probabilities and information theory might well prove necessary to understand consciousness. Even if Chalmers's framework is accepted, a more quantitative structure must be worked out. Do certain types of processing architectures, such as massive parallel versus serial, facilitate the development of consciousness? Does the richness of experience relate to the amount or organization of memory (shared or not, hierarchical or not, static or dynamic memory, and so on)?¹³

¹²Penrose's books (Penrose, 1989, 1994) are among the most lucid and best-written accounts of Turing machines, Gödel's theorems, computing, and modern physics I have encountered. However, given that both monographs nominally deal with the human mind and brain, they are equally remarkable for the almost complete absence of any serious discussion of psychology and neuroscience. Hameroff and Penrose (1996) outline their proposal that microtubules, a major component of cellular scaffolding, are critical to the processes underlying consciousness. The Achilles' heel of this idea is the lack of any biophysical mechanism that would permit neurons, and not just any cells, to rapidly form highly specific coalitions across large regions of the brain on the basis of quantum-coherency effects. All of this is supposed to take place, of course, at body temperature, a rather hostile environment for sustaining such events over macroscopic scales.

¹³I definitely recommend at least browsing through Chalmers's (1996) book, in particular his Chapter 8. For a theoretical approach toward consciousness based on measures of complexity and information theory, see Tononi and Edelman (1998) and Edelman and Tononi (2000). Nagel (1988) examines panpsychism.

While I cannot rule out that explaining consciousness may require fundamentally new laws, I currently see no pressing need for such a step.

Consciousness Requires Behavior

The *enactive* or *sensorimotor* account of consciousness stresses the fact that a nervous system can't be considered in isolation. It is part of a body living in a habitat that has acquired, through myriad of sensorimotor interactions over its lifetime, knowledge about the way that the world (including its own body) acts. This knowledge is put to skillful use in the body's ongoing encounters with the world. Proponents of this view acknowledge that the brain supports perception but claim that neural activity is not sufficient for consciousness, and that it is futile to look for physical causes or correlates of consciousness. The behaving organism embedded in a particular environment is what generates feelings.¹⁴

While proponents of the enactive point of view rightly emphasize that perception usually takes place within the context of action, I have little sympathy for their neglect of the neural basis of perception. If there is one thing that scientists are reasonably sure of, it is that brain activity is both necessary and sufficient for biological sentience. Empirical support for this fact derives from many sources. For instance, in dreaming, a highly conscious state, almost all voluntary muscles are inhibited. That is, each night, most of us have episodes of phenomenal feelings yet fail to move.¹⁵ Another example is that direct brain stimulation with electrical or magnetic pulses triggers simple percepts, such as flashes of colored light, the basis for ongoing research in neuroprosthetic devices for the blind. Also, many patients are unfortunate enough to lose the use of their motor system, either during short-lived episodes¹⁶ or permanently,¹⁷ yet continue to experience the world.

¹⁴The manifesto of this movement is (O'Regan and Noë, 2001; Noë, 2004). For a related approach, see (Järvilehto, 2000). Historical antecedents of the enactive movement in philosophy and psychology are (Merleau-Ponty, 1962) and (Gibson, 1966) respectively.

¹⁵The eyes move, of course, during periods of heightened dream activity. Revonsuo (2000) overviews the form and putative functions of dream content.

¹⁶A transient form of paralysis is one of the characteristic features of *narcolepsy*, a neurological disorder. Triggered by a strong emotion—laughter, embarrassment, anger, excitement—the afflicted subject suddenly loses skeletal muscle tone without becoming unconscious. Severe *cataplectic* attacks can last for minutes and leave the patient collapsed on the floor, utterly unable to move or to signal, but fully aware of her surroundings (Guilleminault, 1976; Siegel, 2000).

¹⁷The most dramatic of these have *locked-in syndrome* (Feldman, 1971; see also Celesia, 1997). Take the case of Jean-Dominique Bauby, editor of the French fashion magazine *Elle*, who retained nothing but the ability to move his eyes up and down following a massive stroke. He composed an entire book on his inner experiences using eye movements as a form of Morse code. Bauby's 1997 *Le Scaphandre et le Papillon* (translated as *The Diving-Bell and the Butterfly*) is a strangely uplifting and inspirational volume written under appalling circumstances. If his last link with the world, his vertical eye movements, had been severed, Bauby would have

I conclude that action is not necessary for consciousness. Of course, this is not to argue that motion of the body, eyes, limbs, and so on, isn't important in shaping awareness. It is! Yet behavior is not strictly necessary for qualia to occur.

Consciousness Is an Emergent Property of Certain Biological Systems

The working hypothesis of this book is that consciousness emerges from neuronal features of the brain.¹⁸ Understanding the material basis of consciousness is unlikely to require any exotic new physics, but rather a much deeper appreciation of how highly interconnected networks of a large number of heterogeneous neurons work. The abilities of coalitions of neurons to learn from interactions with the environment and from their own internal activities are routinely underestimated. Individual neurons themselves are complex entities with unique morphologies and thousands of inputs and outputs. Their interconnections, the *synapses*, are molecular machines that come equipped with learning algorithms that modify the strength and dynamics of synapses across many timescales. Humans have no real experience with such a vast organization. Hence, even biologists struggle to appreciate the properties and power of the nervous system.

A reasonable analogy can be made with the debate raging at the turn of the 20th century concerning vitalism and the mechanisms underlying heredity. How can mere chemistry store all the information needed to specify a unique individual? How can chemistry explain how splitting a single frog embryo at the two-cell stage gives rise to two tadpoles? Doesn't this require some *vitalistic* force, or new law of physics, as Erwin Schrödinger was postulating?

The central difficulty faced by researchers at the time was that they could not imagine the great specificity inherent in individual molecules. This is perhaps best expressed by William Bateson, one of England's leading geneticists in the early part of the 20th century. His 1916 review of *The Mechanism of Mendelian Heredity*, a book by the Nobel laureate Thomas Hunt Morgan and his collaborators, states:

been condemned to living a fully conscious life while appearing all but dead! He and other such patients appear to see much like normals, although this has never been systematically studied. *Frozen addicts*, discussed in Chapter 7, are yet another living proof that complete lack of mobility and consciousness can coexist.

¹⁸I use *emergence* in the sense that the initiation and propagation of the action potential in axonal fibers, a highly nonlinear phenomenon, is the result of—and can be predicted from—the attributes of voltage-dependent ionic channels inserted into the neuronal membrane. In this sense, the laws of heredity have their basis in the properties of DNA and other macromolecules in a lawful, reducible manner. For a general introduction to the problem of emergence, see Beckermann, Flohr, and Kim (1992).

1.3 | Our Approach Is a Pragmatic, Empirical One

11

The properties of living things are in some way attached to a material basis, perhaps in some special degree to nuclear chromatin; and yet it is inconceivable that particles of chromatin or of any other substance, however complex, can possess those powers which must be assigned to our factors or gens. The supposition that particles of chromatin, indistinguishable from each other and indeed almost homogeneous under any known test, can by their material nature confer all the properties of life surpasses the range of even the most convinced materialism.

What Bateson and others did not know at the time, given the technology available, was that chromatin (that is, the chromosomes) is only homogeneous statistically, being composed of roughly equal amounts of the four nucleic bases, and that the exact linear sequence of the nucleotides encodes the secrets of heredity. Geneticists underestimated the ability of these nucleotides to store prodigious amounts of information. They also underestimated the amazing specificity of protein molecules, which has resulted from the action of natural selection over a few billion years of evolution. These mistakes must not be repeated in the quest to understand the basis of consciousness.

Once again, I assume that the physical basis of consciousness is an emergent property of specific interactions among neurons and their elements. Although consciousness is fully compatible with the laws of physics, it is not easy to predict or understand consciousness from these.

1.3 | OUR APPROACH IS A PRAGMATIC, EMPIRICAL ONE

In order to make progress on these difficult questions without getting bogged down in diversionary skirmishes, I will have to make some assumptions without justifying them in too much detail. These provisional working hypotheses might well need to be revised or even rejected later on. The physicist turned molecular biologist Max Delbrück advocated “The Principle of Limited Sloppiness” when it comes to experiments. He recommended trying things in a rough and ready manner to see whether they might work out. I apply this principle to the realm of ideas about the brain.

A Working Definition

Most everyone has a general idea of what it means to be conscious. According to the philosopher John Searle, “Consciousness consists of those states of sentience, or feeling, or awareness, which begin in the morning when we awake from a dreamless sleep and continue throughout the day until we fall into a coma or die or fall asleep again or otherwise become unconscious.”¹⁹ If I ask

¹⁹The definition, taken from Searle (1997), leaves out an entire domain of conscious experiences that are usually not remembered: vivid dreams that can’t be distinguished from real life. More

you to describe what you see and you respond in an appropriate manner, I will assume for now that you are conscious. Some form of attention is required, but is not sufficient. Operationally, consciousness is needed for nonroutine tasks that require retention of information over seconds.

Although fairly vague, this provisional definition is good enough to get started. As the science of consciousness advances, it will need to be refined and expressed in more fundamental neuronal terms. Until the problem is better understood, a more formal definition of consciousness is likely to be either misleading or overly restrictive, or both. If this seems evasive, try defining a *gene*. Is it a stable unit of hereditary transmission? Does a gene have to code for a single enzyme? What about structural and regulatory genes? Does a gene correspond to one continuous segment of nucleic acid? What about introns? And wouldn't it make more sense to define a gene as the mature mRNA transcript after all the editing and splicing have taken place? So much is now known about genes that any simple definition is likely to be inadequate. Why should it be any easier to define something as elusive as consciousness?²⁰

Historically, significant scientific progress has commonly been achieved in the absence of formal definitions. For instance, the phenomenological laws of electrical current flow were formulated by Ohm, Ampère, and Volta well before the discovery of the electron in 1892 by Thompson. For the time being, therefore, I adopt the above working definition of consciousness and will see how far I can get with it.

Consciousness Is Not Unique to Humans

It seems plausible that some species of animals—mammals, in particular—possess some, but not necessarily all, of the features of consciousness; that they see, hear, smell, and otherwise experience the world. Of course, each species has its own unique sensorium, matched to its ecological niche. But I assume that these animals have feelings, subjective states. To believe otherwise seems presumptuous and flies in the face of all experimental evidence for the continuity of behaviors between animals and humans. We are all Nature's children.

This is particularly true for monkeys and apes, whose behavior, development, and brain structure are remarkably similar to those of humans (it takes

elaborate definitions of consciousness are no more helpful. For instance, Schiff and Plum (2000), two neurologists who treat severely neurologically impaired patients, state: "At its least, normal human consciousness consists of a serially time-ordered, organized, restricted, and reflective awareness of self and the environment. Moreover, it is an experience of graded complexity and quantity." While useful clinically, this definition presupposes notions of awareness, the self, and so on. The *Oxford English Dictionary* is no better, having eight entries under 'consciousness' and 12 under 'conscious.'

²⁰See Keller (2000) and Ridley (2003) for the checkered history of the term "genes" and Churchland (1986, 2002) for the role of definitions in science.

an expert to distinguish a cubic millimeter of monkey brain tissue from the corresponding chunk of human brain tissue). In fact, the best way to study stimulus awareness at the neuronal level today is based on behaving monkeys. Given this likeness, appropriate experiments on nonhuman primates—carried out in a humane and ethical manner—are necessary to discover the mechanisms underlying consciousness.²¹

Of course, humans do differ fundamentally from all other organisms in their ability to talk. True language enables *homo sapiens* to represent and disseminate arbitrarily complex concepts. Language leads to writing, representative democracy, general relativity, and the Macintosh computer, activities and inventions that are beyond the capabilities of our animal friends. The primacy of language for most aspects of civilized life has given rise to a belief among philosophers, linguists, and others that consciousness is impossible without language and that, therefore, only humans can feel and introspect. While this might be true, to a limited extent, about self-consciousness (as in, “I know that I am seeing red”), all of the evidence from split-brain patients, autistic children, evolutionary studies, and animal behavior is fully compatible with the position that at least mammals experience the sights and sounds of life.²²

At present, it is unknown to what extent conscious perception is common to *all* animals. It is probable that consciousness correlates to some extent with the complexity of the organism’s nervous system. Squids, bees, fruit flies, and even roundworms are all capable of fairly sophisticated behaviors. Perhaps they too possess some level of awareness; perhaps they too can feel pain, experience pleasure, and see.

How Can Consciousness Be Approached in a Scientific Manner?

Consciousness takes many forms, but it seems best to begin with the form that is easiest to investigate. Studying vision has several advantages over studying other senses, at least when it comes to understanding consciousness.

²¹A few words on some of the approximately 200 primate species, of which humans are but one member. The order of *primates* is divided into two suborders, *prosimians* (literally, “before monkeys”) and *anthropods*, encompassing monkeys, apes and humans. There are two superfamilies of monkeys, which have distinct geographical distributions, *New World* and *Old World monkeys*. Old World monkeys, which include *baboons* and *macaques*, have larger and more convoluted brains than New World monkeys, are easily bred in captivity, and are not endangered. They are popular as a model system for human brain organization. *Gorillas*, *orangutans* and the two species of *chimpanzees* constitute the *great apes*. Given their highly developed cognitive abilities and kinship to humans, little invasive research is carried out on apes. Most of what is known about their brains derives from postmortem studies.

²²The position that only humans are conscious and that animals are mere automatons, advocated most famously by Descartes, used to be widespread. After Darwin and the rise of evolutionary explanations, it became less so. However, even today some argue that language is a *sine qua non* for consciousness (Macphail, 1998). Griffin (2001) is the classical reference surviving consciousness throughout the animal kingdom.

First, humans are visual creatures. This is reflected in the large amount of brain tissue dedicated to the analysis of images and in the importance of seeing in daily life. If you have a cold, for instance, your nose becomes stuffy and you may lose your sense of smell, but this impairs you only mildly. A transient loss of vision, as occurs during snow blindness on the other hand, devastates you.

Second, visual percepts are vivid and rich in information. Pictures and movies are highly structured, yet easy to manipulate using computer-generated graphics.

Third, a sheer endless number of illusions directly manipulate visual experience. Take *motion-induced blindness*: a bunch of randomly moving blue lights are superimposed onto three highly salient but stationary yellow spots. Fixate anywhere on the display, and after a while one, two, or even all three disks simply disappear.²³ Gone! It is an amazing sight: The swirling blue cloud wipes the yellow spots from sight, even though the spots continue to stimulate the retina. Following a brief eye movement the spots reappear. While such sensory phenomena are far removed from “intentionality,” the “aboutness of consciousness,” “free will,” and other concepts dear to philosophers, understanding the neuronal basis of visual illusions can teach much about the physical basis of consciousness in the brain. In the early days of molecular biology, Max Delbrück focused on the genetics of phages, simple viruses that prey on bacteria. You might have thought that the way phages pass information on to their descendants is irrelevant to human heredity. Yet this is not the case. Likewise, Eric Kandel’s belief that the lowly marine snail *Aplysia* has much to teach us about the molecular and cellular strategies underlying memory has proven to be prophetic.²⁴

Last, and most important, the neuronal basis of many visual phenomena and illusions has been investigated throughout the animal kingdom. Perceptual neuroscience has advanced to such a point that reasonably sophisticated computational models have been constructed and have proven their worth in guiding experimental agendas and summarizing the data.

I therefore concentrate on visual sensation or awareness. Antonio Damasio, the eminent neurologist at the University of Iowa, refers to such sensory forms of awareness as *core consciousness*, and differentiates these from *extended consciousness*.²⁵ Core consciousness is all about the here and now, while extended

²³The motion-induced blindness phenomenon was discovered by Bonne, Cooperman, and Sagi (2001).

²⁴Kandel (2001).

²⁵See Damasio’s book (1999). A pithy formulation of his ideas can be found in Damasio (2000). The cognitive psychologist Endel Tulving at the University of Toronto refers to percepts as involving *noetic* (knowing) consciousness, in contrast to *autonoetic* (self-knowing) consciousness that is characteristic of episodic memory (Tulving, 1985). Edelman and Tononi (2000) refer to *primary* and *higher-order* consciousness, and Block refers to phenomenal consciousness on the one hand and reflective and self-consciousness on the other (Block, 1995).

consciousness requires a sense of self—the self-referential aspect that for many people epitomizes consciousness—and of the past and the anticipated future.

My research program neglects, for now, these and other aspects such as language and emotions. This is not to say that these are not critically important to humans. They are. Aphasics, children with severe autism, or patients who have lost their sense of self are severely impaired, confined to hospitals or nursing homes. For the most part, however, they can still see and feel pain. Extended consciousness shares with sensory consciousness the same mysterious stance, but it is much less amenable to experimental investigations since these capabilities can't easily be studied in laboratory animals, making access to the underlying neurons difficult.

Underlining my choice is the tentative assumption that all the different aspects of consciousness (smell, pain, vision, self-consciousness, the feeling of willing an action, and so on) employ one or perhaps a few common mechanisms. Figuring out the neuronal basis for one modality, therefore, will simplify understanding them all. From an introspective point of view, this hypothesis is quite radical. What is the communality between a sound, a sight, and a smell? Their content feels quite different, yet all three have that magical buzz about them. Is it possible that the subjective sensations associated with each are caused by similar neuronal events and circuits?

I allude to other promising lines of work, too, such as olfaction and Pavlovian conditioning, particularly if they have features that make them easy to study in the laboratory. Given the desirability of relating consciousness to the firing activity of individual neurons and their arrangements, it will be imperative to carry out relevant experiments in behaving mice. The amazing development of ever-more-powerful molecular biology tools permits scientists to manipulate rodent brains in a deliberate, delicate, and reversible manner, something currently not possible in primates.

Altered states of consciousness—hypnosis, out-of-body experience, lucid dreaming, hallucination, meditation, and so on—are not covered in this book. While all are fascinating case studies of the human condition, it is difficult to access their underlying neuronal representations (can a monkey be hypnotized?). A comprehensive theory of consciousness will ultimately have to account for these unusual phenomena.²⁶

²⁶Blackmore (1982), Grüsser and Landis (1991), and Blanke et al. (2002) describe the psychology and neurology of out-of-body experiences, a fascinating phenomenon that has, until recently, been almost entirely co-opted by new age mystics. Manford and Andermann (1998) review complex visual hallucinations from a clinical vantage point.

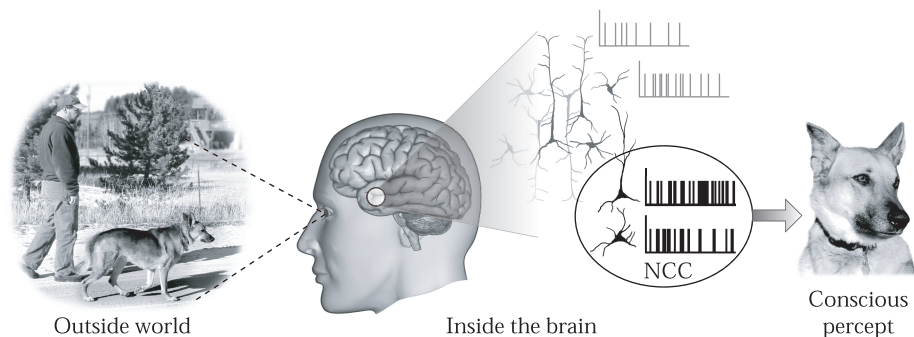


FIGURE 1.1 *The Neuronal Correlates of Consciousness* The NCC are the minimal set of neural events—here synchronized action potentials in neocortical pyramidal neurons—sufficient for a specific conscious percept.

1.4 | THE NEURONAL CORRELATES OF CONSCIOUSNESS

Francis and I focus on discovering the *neuronal correlates of consciousness* (NCC). Whenever information is represented in the NCC you become conscious of it. The goal is to discover *the minimal set of neuronal events and mechanisms sufficient for a specific conscious percept* (Figure 1.1). The NCC involve the firing activity of neurons in the forebrain.²⁷ As detailed in the next chapter, by firing activity I mean the sequences of pulses, about a tenth of a volt in amplitude and 0.5–1 msec in duration. These binary *spikes* or *action potentials* can be treated as the principal output of forebrain neurons. Stimulating the relevant cells with some yet-to-be-invented technology that replicates their exact spiking pattern should trigger the same percept as using natural images, sounds or smells. As I emphasized a few pages earlier, our approach is preconditioned on observations that consciousness depends on what is inside the head, not necessarily on the behavior of the organism.

The notion of the NCC is significantly more subtle than illustrated by the figure and must also specify over what range of circumstances and data the correlation between neuronal events and conscious percept holds. Is the relationship true only when the subject is awake? What about dreams or various pathologies? Is the relationship the same for all animals? These complications are taken up in Chapter 5.

Using the NCC in this way implies that if I am aware of an event, the NCC in my head must directly express this. *There must be an explicit correspondence*

²⁷A three-sentence primer on gross brain architecture: I follow a tripartite division of the vertebrate brain into the *forebrain*, the *midbrain*, and the *hindbrain*. The forebrain consists, by and large, of the neocortex, the basal ganglia, hippocampus, amygdala, olfactory bulb, and the thalamus and its associated structures. The hindbrain includes the pons, medulla, and the cerebellum.

between any mental event and its neuronal correlates. Another way of stating this is that any change in a subjective state must be associated with a change in a neuronal state.²⁸ Note that the converse need not necessarily be true; two different neuronal states of the brain may be mentally indistinguishable.

It is possible that the NCC are not expressed in the spiking activity of some neurons but, perhaps, in the concentration of free, intracellular calcium ions in the postsynaptic dendrites of their target cells.²⁹ Or the invisible partners of neurons, *glia* cells that support, nurture and maintain nerve cells and their environment in the brain, might be directly involved (although this is unlikely³⁰). But whatever the correlates are, they must map directly, rather than indirectly, onto conscious perception because the NCC are all that are needed for that particular experience.

The NCC may be associated with some special type of activity in one or more sets of neurons with some special pharmacological, anatomical, and biophysical properties that must exceed some threshold and last for some minimal amount of time.

As I shall argue in Chapter 14, it is quite unlikely that consciousness is a mere epiphenomenon. Rather consciousness enhances the survival of its carrier. This means that the NCC activity must affect other neurons in some manner. This post-NCC activity influences other neurons that ultimately cause some behavior. This activity can also feed back to the NCC neurons and to previous stages in the hierarchy, significantly complicating matters.

Discovering the NCC would constitute a major step forward on the road to a final understanding of consciousness. Identification of the NCC would enable

²⁸This stance implies that in the absence of a physical carrier, consciousness can't exist. Put succinctly: No matter, never mind.

²⁹The proposition that the NCC are closely related to subcellular processes is not as outlandish as it may sound. Cellular biophysicists have realized over the past few years that the distribution of calcium ions within dendrites represents a crucial variable for processing and storing information (Koch, 1999). Calcium ions enter spines and dendrites through voltage-gated channels. This, along with their diffusion, buffering, and release from intracellular stores, leads to rapid local modulations of the calcium concentration within the dendritic tree. The concentration of calcium can, in turn, influence the membrane potential (via calcium-dependent membrane conductances) and—by binding to buffers and enzymes—turn on or off intracellular signaling pathways that initiate plasticity and form the basis of learning. The dynamics of calcium in thick dendrites and cell bodies spans the right time scale (on the order of hundreds of milliseconds) for perception. Indeed, it has been established experimentally in the cricket that the concentration of free, intracellular calcium in the omega interneuron correlates well with the degree of auditory masking, a time-dependent modulation of auditory sensitivity in these animals (Sobel and Tank, 1994).

³⁰*Glia* cells are as numerous as neurons but are rarely talked about. Their behavior is sluggish and they show little of the elaborate sensitivity associated with forebrain neurons (Laming et al., 1998). This behavior is why they are unlikely to play a direct role in perception. Some glia cells exhibit all-or-none propagating calcium events, akin to action potentials, except that they occur over seconds (Cornell-Bell et al., 1990; Sanderson, 1996).

neuroscientists to manipulate its cellular substrate on the basis of pharmacological intervention and genetic manipulation. It may be possible to fashion transgenic mice whose NCC can be rapidly and safely switched on and off. What behaviors might such zombie rodents be capable of? Clinical benefits will flow from this discovery as well, such as a better understanding of mental diseases and the design of new and powerful anesthetics, with few side effects.

Eventually, a theory that bridges the explanatory gap, that explains why activity in a subset of neurons is the basis of (or, perhaps, is identical to) some particular feeling, is required. This theory needs to make comprehensible why that activity means something for the organism (e.g., why does it hurt?) and why qualia feel the way they do (e.g., why does red look one way, quite different from blue?).³¹

Along the way, the great debate that swirls around the question of the exact relationship between neuronal and mental events needs to be resolved. *Physicalism* asserts that the two are identical; that is, that the NCC for the percept of purple is the percept. Nothing else is needed. While the former is measured by microelectrodes, the latter is experienced by brains. A favorite analogy is with temperature of a gas and the average kinetic energy of the gas molecules. The former is a macroscopic variable that is recorded by a thermometer, while the latter is a microscopic variable that requires quite a different set of tools to study. Yet the two are identical. Even though, superficially, they appear quite distinct, temperature is equivalent to the average kinetic energy of molecules. The faster the molecules move, the higher the temperature. It does not make sense to talk of the rapid molecular motion causing temperature as if one is the cause and the other the effect. One is sufficient and necessary for the other.³²

At this point, I am not sure whether this sort of strong identity holds for the NCC and the associated percept. Are they really one and the same thing, viewed from different perspectives? The character of brain and phenomenal states appear too different to be completely reducible to each other. I suspect that their relationship is more complex than traditionally envisioned. For now, it is best to keep an open mind on this matter and to concentrate on identifying the correlates of consciousness in the brain.

³¹There is no guarantee that science will discover a final, objective theory of consciousness. As Chalmers (1996) and others have argued, one might have to settle for some sort of nonreductionist physicalist account of consciousness or for an ontological dualism with rigorous, quantitative bridging principles linking the domain of subjective experiences with objective reality. Only time will tell.

³²There is an extensive philosophical literature on this topic, with many, many variants. I refer the curious reader to Patricia Churchland's books, which deal extensively with this topic (1986, 2002).

1.5 | RECAPITULATION

Consciousness resides at the nexus of the mind-body problem. It appears as mysterious to 21st-century scholars as when humans first started to wonder about their minds several millennia ago. Nevertheless, scientists today are better positioned than ever to investigate the physical basis of consciousness.

My approach is a direct one that many of my colleagues consider naive or ill-advised. I take subjective experience as given and assume that brain activity is both necessary and sufficient for biological creatures to experience something. Nothing else is needed. I seek the physical basis of phenomenal states within brain cells, their arrangements and activities. My goal is to identify the specific nature of this activity, the neuronal correlates of consciousness, and to determine to what extent the NCC differ from activity that influences behavior without engaging consciousness.

The focus of this book is on sensory forms of consciousness—vision in particular. More than other aspects of sensation, visual awareness is amenable to empirical investigations. Emotions, language, and a sense of the self and of others are critical to daily life, but these facets of consciousness are left for later, when their neural bases will be better understood. Similar to the quest to understand life, discovering and characterizing the molecular, biophysical, and neurophysiological operations that constitute the NCC will likely help solve the central enigma, how events in certain privileged systems can be the physical basis of, or even be, feelings.

It would be contrary to evolutionary continuity to believe that consciousness is unique to humans. I assume that the human mind shares some basic properties with animal minds—in particular, with mammals such as monkeys and mice. I ignore niggling debates about the exact definition of consciousness and whether or not my spinal cord is conscious but is not telling me. These questions must be answered eventually, but today they only impede progress. You don't win a war by fighting the most arduous battle first.

Blunders will be committed and oversimplifications will be made in the course of this colossal, long-term undertaking, but these will only become apparent as time passes. For now, science should rise to the challenge and explore the basis of consciousness in the brain. Like the partially occluded view of a snow-covered mountain summit during a first ascent, the lure of understanding this puzzle is irresistible. As Lao Tsu remarked many years ago, "A journey of a thousand miles begins with a single step."

Now that we have started, let me acquaint you with some key concepts that will guide our quest. In particular, I need to flesh out the notions of explicit and implicit neuronal representations, essential nodes, and the various forms of nervous activity.

