world, and making decisions. We will review this content in later sections. In Sections 3 and 4, we address fundamental issues about how neuroscientists gain knowledge: how they study neural processes, and whose neural processes they study. A major aim of neuroscientists is to offer explanations for behavior and cognition, and Section 6 will offer accounts of what is required of an explanation. Sections 7 and 8 focus on more specialized issues of neuroscience explanations: the levels at which explanations are offered and whether explanations should attribute representations to neural processes.

Both in neuroscience and in philosophy it has been common to adopt a cortico-centric view of the brain, but in fact there is extensive research in neuroscience on subcortical areas. Subcortical processing is extremely important in determining how we behave. This is significant since cortical areas constitute a different type of neural processing system than subcortical areas, and in Section 9, we focus on what is distinctive about the neocortex in particular. We then turn to the question of how the whole brain is organized. It is often viewed as organized as a hierarchy with the neocortex at the top, and indeed, one part of the neocortex, the prefrontal region, at the very top, operating as a central executive. In Section 10, we contrast this with a heterarchical perspective that views neural processes as organized in an interactive network, with different regions exercising control over different aspects of behavior and cognition. Finally, in Section 11, we pull from various topics addressed in earlier sections to address the neurophilosophical question of what neuroscience has to teach us about ourselves as agents in the world.

2 What Are Neurons and Neural Processes?

Most people have seen multiple (typically idealized) pictures of the human brain as it would appear if one opened up the skull. The first thing one notices is a highly convoluted gray structure (at the top of Figure 1) in which the projecting areas are known as gyri and the indented areas as sulci. This structure, known as the *neocortex*, is often divided into four lobes: frontal, occipital, parietal, and temporal. As the part of the brain that has most expanded in the lineage of primates, including us, it has assumed a central focus in much philosophical theorizing. However, as the characterization of it as *neo* suggests, there is more to the cortex (often termed the *cerebral cortex*), including very important structures such as the hippocampus. The term *cortex* is derived from the Latin term for the bark of a tree and, as that suggests, it refers just to the outer structure. There is much of the brain beneath the cortex.

In this Element, we seek to avoid the all too frequent cortico-centric take on the brain by focusing as much on what is beneath the cortex and the

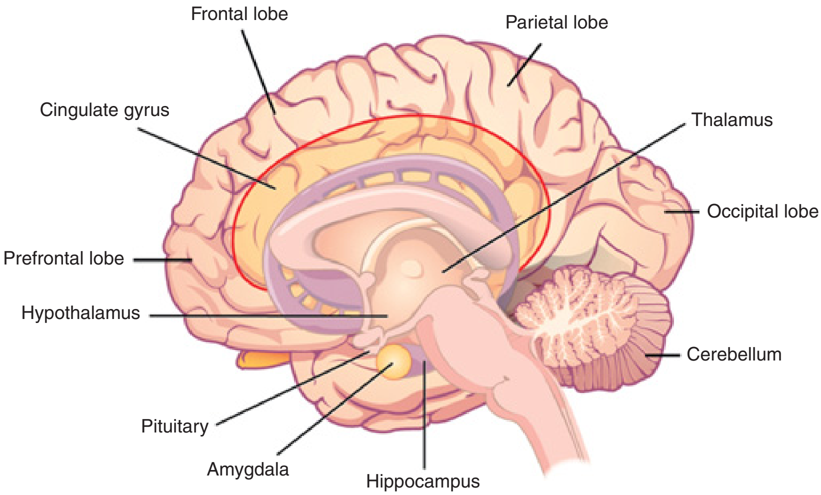

Figure 1 Major areas of the human brain. Adapted from OpenStax College – Anatomy & Physiology, Connexions website: http://cnx.org/content/col11496/1.6/, June 19, 2013., distributed under CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=30148029.

philosophical questions that those areas engender. By taking into account subcortical brain regions and their role in behavior and cognition, we will be in position, by Section 9, to address what is different about the neocortex and how it provides humans with distinctive cognitive abilities. Even as it enables these distinctive abilities, the neocortex does so through interacting with subcortical regions. For now, we start from the building block of all nervous tissue, the neuron, and then consider ways in which neurons are organized.

2.1 The Neuron

The neuron is a specialized type of cell. Although neurons are too small to be seen with the naked eye, ancient anatomists did observe nerves (bundles of neurons), recognized their importance in transmitting signals through the body, and speculated on their constitution. Most hypotheses viewed them as functioning much like blood vessels, with very fine matter (animal spirits, where *spirit* refers to fine matter, as in *spirits of alcohol*) flowing through them. Only in recent centuries did researchers ascertain that neurons transmit electrical current.

The research that would reveal electrical transmission of neurons began around 1600, when investigators (and the lay public!) began experimenting with electrical shocks, including those generated by friction machines. Many were fascinated by how these could cause muscle contractions. Based on extensive experiments with frog legs, involving the elicitation of muscle contraction by spark-generating machines or by lightning, Galvani (1791) argued that muscles possessed their own source of what he termed *animal electricity*. Extensive research through the nineteenth and first half of the twentieth centuries revealed that what he had identified was an electrical potential due to different concentrations of potassium and sodium ions across the membranes of both muscles and neurons. Changes in these concentrations propagate along the neuron, often in the form of action potentials, also referred to as *spikes*. Action potentials are large changes in the membrane potential at one location on the membrane that cause similar changes at adjacent locations, creating a wave of electrical current that passes along the neuron until it reaches the synaptic terminal at the end of the neuron. Figure 2 shows the now canonical representation of the action potential, which begins with the neuron negatively polarized (approximately −70 mV; referred to as the *resting potential*). When a stimulus is sufficient to push the potential above threshold, it rapidly and temporarily depolarizes to approximately +40 mV before repolarizing. When neurons propagate action potentials, they are often said to *fire*, capturing the fact that action potentials represent relatively discrete signals propelled along neurons.[1]

During the same period (the nineteenth century), other researchers were examining biological tissues with the light microscope. They identified what are termed *cells* and advanced the theoretical framework in which cells are the basic living units. Adding stains enabled researchers to see the projections – axons and dendrites – that differentiate neurons from other cells. One stain, a silver nitrate stain introduced by Comillo Golgi, was particularly informative since, for reasons still not understood, it only stains some neurons in a preparation. This makes it possible to visualize individual neurons. Golgi, however, did not interpret what he saw as individual cells but rather as a continuous reticular network of nerve tissue. Adopting Golgi's stain and visualizing such things as developing or degenerating nerve fibers, Santiago Ramón y Cajal concluded that the network was not continuous; rather, there were gaps between projections from different nerve cells. Drawing upon Cajal's work, Waldeyer invented the term *neurone*, now *neuron*, and articulated the *neuron doctrine* according to which discrete neurons are the units of nerve tissue. At the end of the nineteenth century, the dispute between Golgi and Cajal was very contentious, and even as both were awarded the Nobel Prize in 1906, Golgi

[1] Not all neurons generate action potentials. Some transmit graded potentials. Instead of discrete, digital signals, they generate responses of varying magnitude. An important advantage of signaling with action potentials is that they can be maintained over long distances without loss of content.

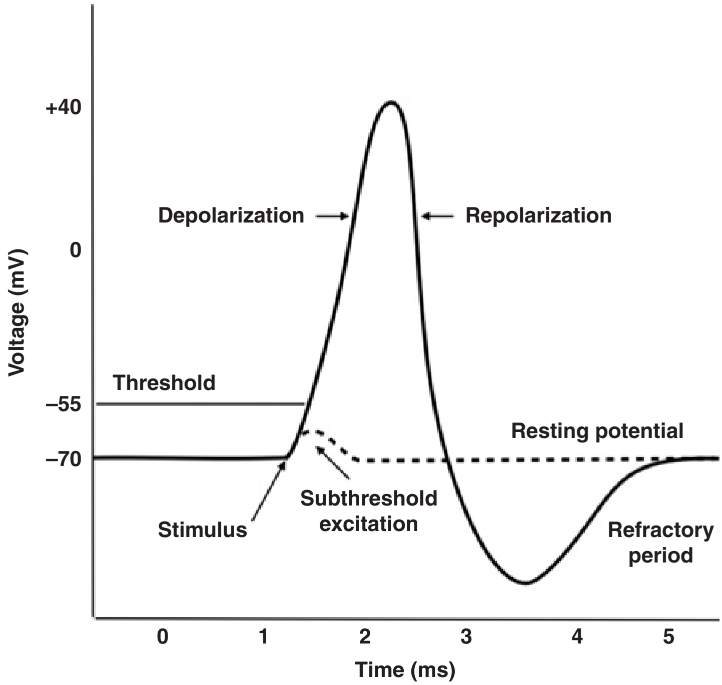

Figure 2 Characteristic graph of voltage changes during an action potential. Adapted from en:User:Chris 73, updated by en:User:Diberri, converted to SVG by tiZom – distributed under CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=2241513.

continued to defend the reticular view, arguing that only if nerves consisted of an interconnected network would they be able to communicate messages through the body. The conflict between Golgi and Cajal is an illuminating example of how skilled observers can reach conflicting conclusions and how such conflicts are resolved (for discussion and details, see Mundale, 2001; Shepherd, 2016).[2]

Cajal argued that the two types of processes extending from the neuron cell body play different roles. He interpreted the typically short and highly branching structures, known as *dendrites*, as receiving inputs from other neurons, and the longer, less branched structures, known as *axons*, as carrying output to other cells. He supported this by the observation that sensory neurons have their dendrites oriented toward the sense organ (e.g., the eye) and axons oriented

[2] One might think such a schism could be resolved simply by looking carefully through the microscope, but the gap between neurons is too small to be seen with the light microscope. When the electron microscope was applied to nerve tissue in the 1950s, it did reveal the gap, but ironically it also revealed the presence in some cases of direct contacts between nerve cells, known as gap junctions.

toward the brain. In making this distinction, he proposed that there was one-way transmission through the nervous system. In 1897, Charles Scott Sherrington characterized the gap between neurons as *synapses* (derived from the Greek for "to clasp"). Figure 3 shows a prototypical neuron. Although we will not develop the point, one should note that neurons exhibit enormous variety both in appearance and function.

The discovery of synapses presented a new challenge: How do signals cross the gaps between neurons? The initial assumption of many researchers was that electrical charges could jump synapses, much as sparks from a spark generator can jump to a grounded surface. A long lineage of research, especially in the first half of the twentieth century, ultimately revealed this was incorrect and transmission between neurons is chemical.

Most of the initial work that led to this conclusion focused on the junction between nerve and muscle. Around the turn of the century, a few pharmacologists and chemists began investigating substances (such as an extract from the adrenal gland initially referred to as noradrenaline and later as norepinephrine) that elicited or inhibited responses of muscles. A notable finding was the accumulation of another chemical, acetylcholine, in heart tissue when stimulated by the vagus nerve (which projects from the central brain to the heart, lung, and intestines). Many investigators, however, initially resisted the idea that acetylcholine was released by the nerve and caused contraction of heart muscles

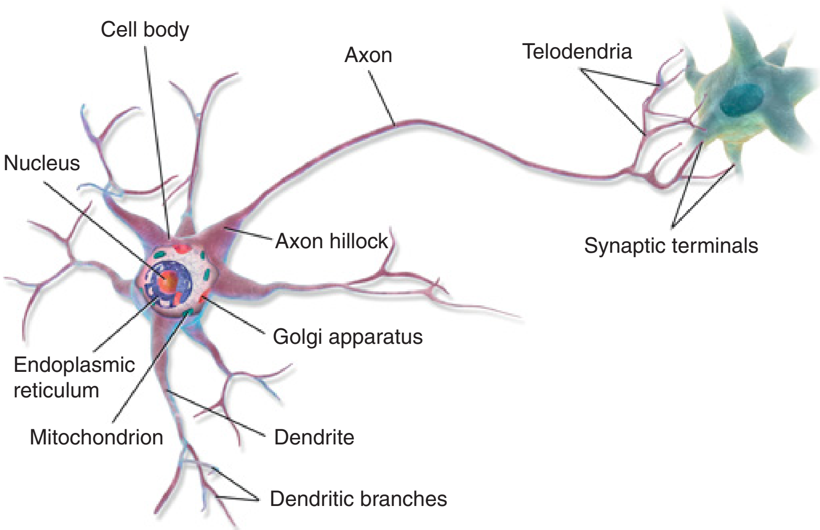

Figure 3 A prototypical neuron. Figure by Bruce Blaus, distributed under CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=28761830.

to slow down. In 1920, Otto Loewi provided compelling evidence that acetylcholine was released by nerves by bathing the heart of one frog in liquid and stimulating the vagus nerve. Once its heart contractions slowed, he transferred the liquid to the heart of another frog whose vagus nerve had been removed. Heart contractions in that frog also slowed.

Although this provided compelling evidence that chemicals released by neurons act on internal muscles such as those in the heart, many resisted the idea that chemicals transmit signals between two neurons or between neurons and skeletal muscles. Chemical signaling, it was thought, was too slow. This ensuing conflict came to be known as the war of the soups (advocates of chemical transmission) and the sparks (advocates of direct electrical transmission). (For an engaging analysis of the conflict, see Valenstein, 2005.) In the wake of the victory by the soups, hundreds of chemicals, referred to as neurotransmitters, have been discovered and neuroscientists have developed an understanding of how they are synthesized and released from one neuron and, by binding to receptors, generate changes, including action potentials, in other neurons.

In most cases, neurotransmitters bind to a receptor in the postsynaptic cell and serve either to depolarize it (thereby increasing the likelihood that it will generate an action potential) or further polarize it (thereby inhibiting it). Any excess is typically quickly broken down and the components recycled. Some neurotransmitters, referred to as *volume transmitters* or *neuromodulators*, disperse widely and serve to modulate the behavior of neurons that have the appropriate receptors. We noted that neurons come in a huge variety. An important type of variation involves the neurotransmitters that they release or to which they respond.

2.2 Foundational Neural Structures: Nerve Networks

As important as individual neurons are, they typically carry out their activities as parts of collectives. Hence, in this and the following sections, we focus on some important ways in which neurons combine into larger structures.

Given that dendrites receive signals and axons send out signals, it is plausible to view neurons as having evolved to connect sensory and motor processes. Philosophers Keijzer, van Duijn, and Lyon (2013) have challenged that view, arguing instead that the first function of neurons was to coordinate muscles. They refer to their proposal as the *skin–brain hypothesis* and appeal to jellyfish to illustrate it. Jellyfish belong to the phylum *Cnidarian*, which differentiated from other animals between 500 and 700 million years ago and is thought to be representative of early evolved animals. A prominent feature of jellyfish is the