5. What has neuroscience learned?

In previous sections we introduced some of the tools used in neuroscience and the organisms which it investigates. At this point it will be helpful to introduce some examples of what neuroscience has learned about vertebrate brains, including our own. We will make use of these examples in subsequent sections as we engage in philosophical discussions about neuroscience.

5.1 Keeping track of time of day in the suprachiasmatic nucleus

We start with a nucleus within the hypothalamus (labeled in Figure 1), the suprachiasmatic nucleus (SCN). The hypothalamus is a collection of nuclei that play critical roles in regulating fundamental activities such as eating, maintaining wakefulness or going to sleep, and reproduction (Leng, 2018). Individual nuclei receive inputs and send outputs to various regions of the body but also to regions elsewhere in the brain. Most release neuropeptides and volume transmitters that diffuse broadly, modulating activity of other neurons as well as controlling physiological processes. For example, the arcuate nucleus contains neurons that respond to peptides released in the intestinal tract that signal fat concentration or whether food is being digested. The outputs of these neurons in turn regulate eating behavior (Sohn, 2015). A nucleus in the lateral zone of the hypothalamus contains hypocretin-producing neurons that play critical roles in maintaining wakefulness or transitioning to sleep—silencing these neurons induces slow-wave sleep (Burk & Fadel, 2019). The role of the hypothalamus is sometimes minimized as merely engaged in bodily maintenance. A useful corrective is to reflect on how much of our behavior is focused on activities such as eating, sleeping, and reproducing.

Coordinating our activities with the light-dark cycle on our planet is of fundamental importance. Although artificial lighting allows us to carry on our activities around the clock, our physiological and cognitive activities are affected by endogenously produced rhythms of approximately 24 hours (named, *circadian* from *circa*, approximately + *dies*, day). These become apparent to us as jetlag when we travel across multiple time zones, but they are also manifest in the increased rates of obesity, cancer, etc., in shift workers. Enzymes responsible for the activities of nearly every organ in our bodies exhibit oscillating expression over the course of a day, thereby resulting in varying performance. This includes regions of the brain that are involved in higher cognitive activities such as reasoning and decision making. Our capacity to perform these activities varies over the course of the day.

Research on fruit flies provided the first clues to how circadian rhythms are generated. (Fruit flies do not have an SCN; rhythms are instead maintained by a small collection of neurons in their brain.) A search of genetic mutations revealed one gene, named *period* (or *per*), in which mutations altered the period of rhythms or eliminated them altogether (Konopka & Benzer, 1971). Genes are transcribed into messenger RNAs (mRNAs) and translated into proteins. Still working in fruit flies, Hardin, Hall, and Rosbash (1990) established that concentrations of both the *per* mRNA and the protein Per oscillated over a 24-hour period, with the protein lagging a few hours behind the mRNA. Since a negative feedback loop is a common mechanism for

generating oscillations (see section 6.3 for further discussion), they proposed that rhythms resulted from a feedback process: as the protein Per accumulated, it inhibited the expression of the *per* gene, resulting in the concentration of the protein subsequently diminishing, only to increase again when Per itself degraded.

Studies lesioning the SCN or recording from SCN neurons revealed that circadian rhythms are generated in much the same manner in vertebrates (Takahashi, 2017). While the oscillations occur within individual SCN neurons, the connections between them turn out to be important. Individual neurons generate rhythms with different periods; only as a result of each neuron modulating the activity of others does a regular oscillation of approximately 24 hours arise (Welsh, Takahashi, & Kay, 2010). As that oscillation is still only approximately 24 hours, the SCN, like an old-fashioned watch, will drift gradually from the correct time and frequently has to be reset by external cues. The effects of not doing so were shown in classic experiments in which humans lived in enclosures without external time cues. Their circadian rhythms were somewhat longer than 24 hours, leading them to rise several minutes later each day. Keeping the oscillations in the SCN synchronized with the day-night cycle on our planet is achieved through daylight, which in humans is only processed by the eyes.[1] Daylight also plays a central role in the human ability to overcome the effects of jetlag. One of the best ways to adjust to a new time zone is to time exposure to daylight appropriately.[2] Conversely, avoiding daylight after nightshift work can minimize the ill effects of rotating shiftwork, which stem from constant resetting of the clock as a result of continuing changes in timing of light exposure.

5.2 Mapping location in space in the hippocampus

Knowing one's location in space and how to navigate to other locations is an extremely important behavioral capacity. Animals often depend on being able to figure out routes to food or their nests. Tolman (1948) demonstrated this ability in rats running mazes. When a new, more efficient route became available or when the route previously used was blocked, rats flexibly adjusted the routes they took to a food location. He inferred that the rats were using *cognitive maps* to determine their routes much in the manner humans use physical maps. Tolman, however, did not have the research tools to identify where these maps are in the brain and how the rats could use them to guide behavior.

Clues to the location of cognitive maps in the brain came from research on deficits in rats with damage to the hippocampus, a cortical structure adjacent to the temporal lobe of the neocortex. In experiments using the Morris Water Maze—an apparatus in which rats are forced to swim until they find a hidden platform on which they can stand—normal rats learn the

[1] In other animals such as birds, the pineal gland, which Descartes thought was the locus of interactions between the mind and body, is sensitive to light that passes through the skull. Although it is not the locus of mind-brain interactions, the pineal gland is important in mammals, including us: it releases melatonin, which serves to reset the SCN neurons (hence, the popular use of melatonin to counter jetlag).

[2] The key is to target exposure to the period just prior to normal first light exposure in the location from which one started when traveling eastbound. Since the clock can only adjust about one hour per day, it is important to advance this exposure about an hour each day.

location quickly and swim directly to the platform from wherever they are in the maze. Rats with damage to the hippocampus fail to learn, continuing to swim erratically on subsequent trials (O'Keefe & Nadel, 1978).[3] Little was known at the time about the function of hippocampus. The hippocampus had primarily been a source of neurons to study individual neuron behavior. Typically, neurons cease to respond to stimuli over time, but researchers discovered that if they applied a brief but intense sequence of electrical stimulations to hippocampal neurons, they would continue to respond. This phenomenon, now known as *long-term potentiation*, indicated that response properties of neurons could be altered and offered a model of how neurons in the hippocampus (and elsewhere) can quickly alter their response properties (a feature important for learning). (See Craver, 2003, for a philosophical examination of this history.)

To figure out how the hippocampus could constitute a cognitive map, John O'Keefe recorded from neurons in the hippocampus while a rodent occupied different locations in an enclosure. He found different neurons that would increase their firing rate when the rodent was in different specific regions of its enclosure. O'Keefe and Conway (1978) named these neurons *place cells*. Since different cells functioned as place cells in different environments, numerous researchers began to examine how place cells would respond as they morphed one environment into another. Some changes, such as rotating the enclosure or shrouding it with a curtain, did not affect the activity of place cells, but others, such as altering cue cards placed on the walls of the enclosure or significantly increasing the size of the enclosure, did. Studying how neurons altered their response patterns with environmental changes, such as gradual transformation of a square enclosure into a circular one, provided an avenue for studying how these maps are established (Colgin, Moser, & Moser, 2008).

In section 3.3 above, we described how EEG identified ongoing oscillations of subthreshold electrical activity in many brain areas. These oscillations can also be detected intracranially. In the hippocampus there is an ongoing theta oscillation (6-12 Hz). Comparing the timing of action potentials in place cells with this oscillation, O'Keefe and Recce (1993) revealed that when the rat first entered the area to which a place cell would respond (its place field), the place cell would fire at the trough of the theta cycle. On each subsequent oscillation, it would fire at a slightly earlier phase of the theta-cycle (Figure 12). Comparing the timing of action potentials in neurons for nearby regions provided the rodent a means to track its location along a path (place cells that fired earlier in the theta cycle represented locations arrived at earlier.) An even more striking finding was that when rodents were allowed to run down a runway, the activity of place cells after they finished the run would reflect the sequence of place cell activity during the run, but in reverse. If the animals were delayed in starting the run, place cells would fire in the

[3] During the same period, a surgery to remove the hippocampus in order to reduce epileptic seizures in Henry Molaison (HM; see section 4.1) resulted in his total inability to develop new explicit memories, leading to the hypothesis that the hippocampus is the locus in which new explicit memories are initially encoded. Since HM could remember events from years before the surgery, researchers have proposed that over time memories are transferred to areas in the neocortex. Initially the rodent researchers and the human researchers viewed themselves as studying different phenomena. More recently, though, there have been attempts to connect the two abilities that are lost with damage to the hippocampus (Eichenbaum, 2002).

same sequence as when they subsequently ran (Diba & Buzsaki, 2007). This indicated that place cell activity serves to replay or anticipate future routes.

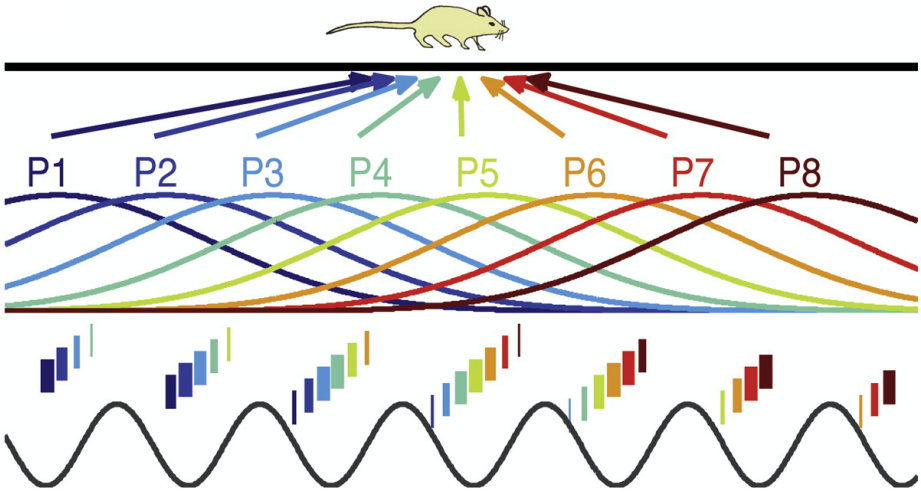

Figure 12. Neurons responding to place fields P1 through P8 fire in relation to the ongoing theta cycle. Those representing the current location produce more spikes (indicated by the width of the bar) at the trough of the theta cycle. Those representing earlier locations issue fewer spikes and are earlier in the theta cycle. Those representing future locations also issue fewer spikes, but later in the theta cycle. Reprinted from Buzsáki (2010) with permission from Elsevier.

5.3 Seeing the world with the visual neocortex

In section 3.3 we described how Hubel and Wiesel (1959) presented visual stimuli while recording from neurons in BA17/V1 and determined that these neurons responded to edges in the visual scene. These researchers noted that detecting edges was only an early step in seeing the world and initiated a project of recording from areas in front of V1. They determined that neurons in a region of BA18 that came to be known as V2 are able to respond to illusory contours (Figure 13). Moving progressively forward in the brain, researchers such as Semir Zeki (1971) showed that V4 neurons responded to shapes and that V5/MT neurons, as we discussed in section 3.4, respond to motion. (The adjacency relations between these areas are indicated in Figure 14A, a map developed to show, as well as possible, these relations in two dimensions.)

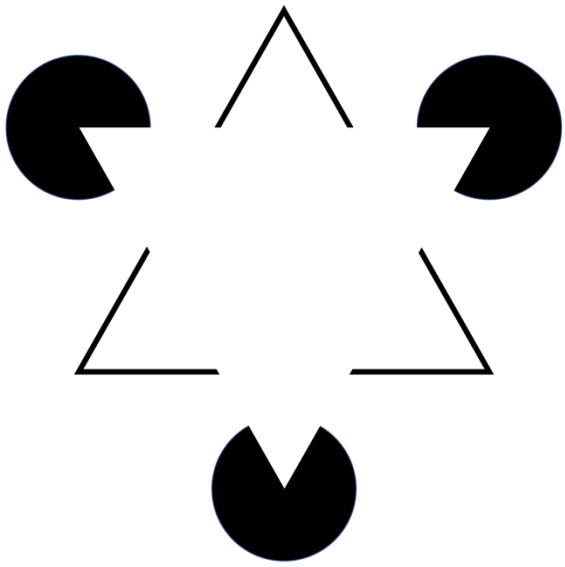

Figure 13. In this figure, originally developed by Kanizsa (1976), the white triangle appears to have boundaries, but these are illusory. Image by Fibonacci - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=1788215

Based on the different effects of lesions to regions in the temporal and parietal lobes in monkeys, Mishkin, Ungerleider, and Macko (1983) advanced the hypothesis that visual processing proceeds along two separate pathways from the occipital cortex: a dorsal pathway to the parietal lobe in which neurons respond to information about where the stimulus is in the visual field (the *where* pathway) and a ventral pathway to the temporal lobe in which neurons respond to identity of the object serving as the stimulus (the *what* pathway). Recording from neurons at the end of the what pathway in the inferotemporal cortex revealed that they respond to the identity of objects wherever they appear in the visual field. For example, Gross, Rocha-Miranda, and Bender (1972) identified neurons that responded to the human hand however it was oriented. Neuroimaging studies on humans revealed that one widely discussed area in the pathway, the *fusiform face area*, responds to faces of specific people (Kanwisher, McDermott, & Chun, 1997). Individuals with lesions in this area experience prosopagnosia (the inability to recognize the faces of people they know). It is in this area that Quiroga et al. (2005) identified neurons that responded selectively to pictures of Jennifer Aniston or Bill Clinton (see section 3.3). There is ongoing controversy as to whether these neurons respond only to faces or also to specific individuals in categories such as trees.

Neurons at the top of the where pathway in the posterior parietal cortex respond to location information but in terms of coordinates fixed by one's head (Andersen, Essick, & Siegel, 1985). van Essen and Gallant (1994) synthesize much of the information about different visual processing regions into an account of two processing streams (Figure 14), referring to streams rather than pathways to recognize that there are various points of crossover from one stream to the other.

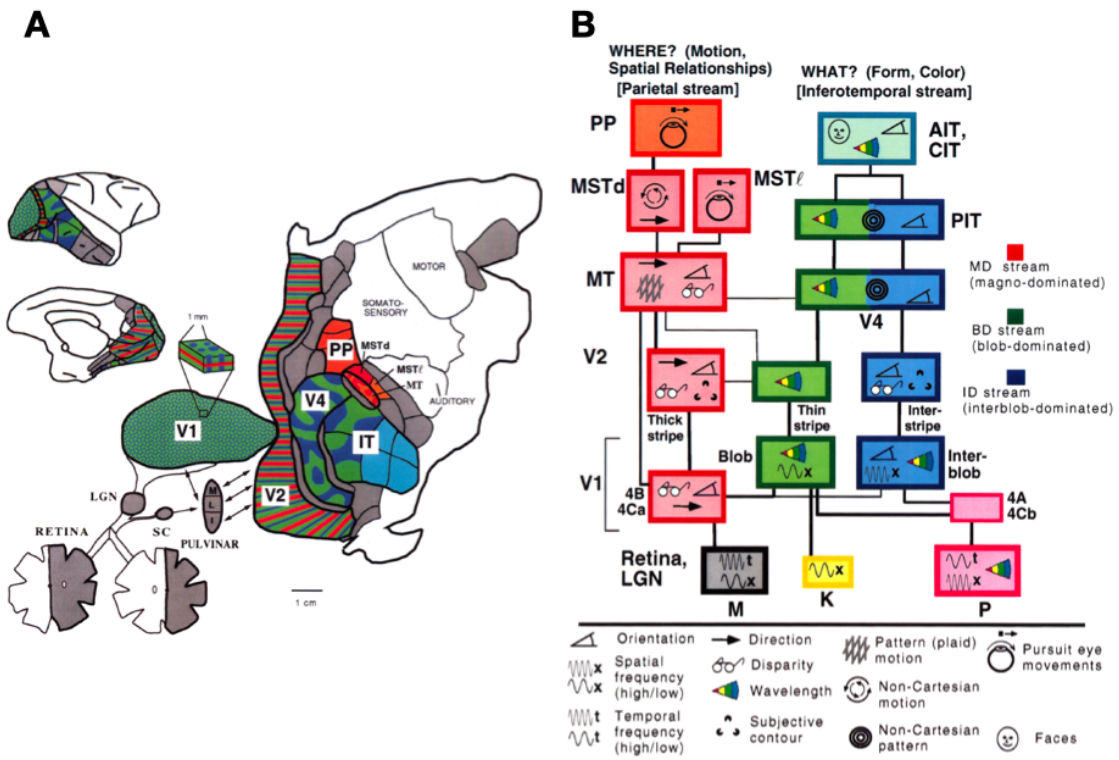

Figure 14. A. A flat representation of the neocortex, with areas known to be involved in vision shown in color or dark grey. B. A schematic diagram of the processing pathways involved in vision. Reprinted from van Essen and Gallant (1994) with permission from Elsevier

Accounts of just what is being processed in a brain area are often controversial. Milner and Goodale (1995) challenged the what/where distinction, arguing instead, on the basis of deficits in patients with brain damage, for a distinction between vision for perception and vision for action. One patient they studied, with temporal lobe damage, could see features of objects but was unable to recognize what type of object it was. She could only draw them laboriously. She also could not describe the orientation of a slot but nonetheless was able to correctly insert a letter into it. Goodale and Milner concluded that the areas spared in the parietal stream were not generically involved in processing location information but rather served to coordinate visual information with actions. The disagreement between Miskin et al. and Milner and Goodale illustrate the challenge of settling what information a brain area is actually processing.

5.4 Making decisions in the basal ganglia

Decision making is a fundamental activity for all organisms as they can only perform some of the activities available to them at a given time. This is clear if one considers locomotion—if an organism moves, it cannot remain still. The network in *C. elegans* we described in section 4.2 makes decisions between forwards and backwards movement. The basal ganglia, several interconnected nuclei found in all vertebrates, including those without a neocortex, play a particularly important role in decision.

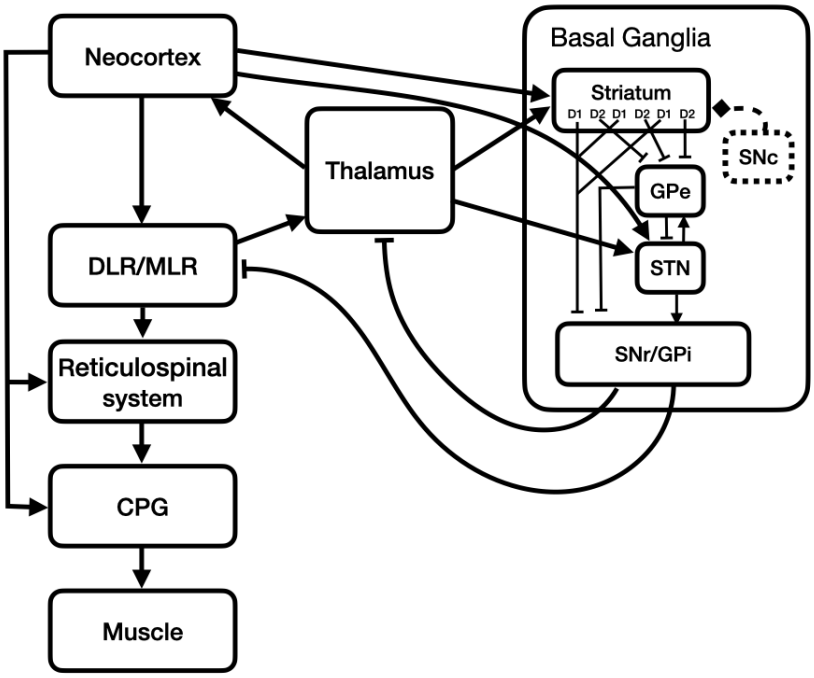

Figure 15 Neural centers operating on muscles. Arrows represent excitation, edge-ended lines inhibition. Dotted line with diamond end indicates dopamine modulation for D1 and D2 neurons in the striatum. See text for abbreviations.

The critical role the basal ganglia play in selecting behaviors is illustrated by mid-20th century research involving cats and other animals in which neural pathways were destroyed at various points between the CPGs (discussed in section 2.2) and the neocortex (Figure 15). Stimulating the CPGs in organisms in which the CPGs were cut off from the rest of the brain resulted in specific movements that adjusted only to feedback from the muscles themselves. When lesions were between the mesencephalic locomotor region (MLR)/the diencephalon locomotor region (DLR) and the reticulospinal system, electrical stimuli to the reticulospinal system generated activity in multiple CPGs but not overall coherent activities such as walking or running. However, if the lesion is above the MLR/DLR, severing the connections with the thalamus, basal ganglia, and neocortex, stimulation of MRL neurons resulted in coherent movement patterns such as walking or running. The animals also avoided obstacles and elicited appropriate metabolic activity to meet the energy demands of muscle activity (Grillner & El Manira, 2019). What these results point to is a hierarchy of areas involved in ever larger-scale coordination of activity.

With lesions below the level of the basal ganglia and thalamus, animals only perform activities when stimulated; they do not initiate them. If the thalamus and basal ganglia are spared, however, animals are able to initiate and choose between activities. For example, cats whose neocortex was removed in infancy but retained an intact thalamus and basal ganglia walk, explore their environment, clean themselves, seek and eat food, etc. They can live and fend for themselves for several years in the protected environment of the laboratory (Bjursten, Norrsell,

& Norrsell, 1976). To do this, they must regularly make decisions about which actions to perform.

The structure of the basal ganglia suggests its role in decision making. As shown in the upper right of Figure 15, the basal ganglia are an interconnected collection of subcortical nuclei. It is easiest to understand how they function by starting with the output nuclei: the substantia nigra pars reticulata (SNr) and the globus pallidus internal (GPi). Projections from neurons in these areas to other parts of the brain are all inhibitory. Moreover, they generate action potentials without external stimulation. Hence, by default the basal ganglia inhibit or shut down activity elsewhere in the brain. What processing elsewhere in the basal ganglia does is selectively remove this inhibition, thereby allowing target areas to perform their activities.

Decisions as to whether to release other brain regions from inhibition are made through the collective operation of three pathways within the basal ganglia: the direct, indirect, and hyperdirect. Here we discuss just the direct and indirect pathways. Both originate with the striatum, which receives inputs from elsewhere in the brain and is organized so that nearby regions mostly receive inputs from adjacent brain areas. Striatal neurons are distinguished by whether they have D1 or D2 dopamine receptors (we will return to the significance of dopamine below). These give rise to the direct and indirect pathways respectively. D1 neurons send inhibitory outputs directly to neurons in the output nuclei (hence, the name *direct*). The overall effect of D1 neurons is to decrease the inhibition generated by these output neurons, thereby activating the cortical and subcortical areas they target. As an illustration, by stimulating appropriate D1 neurons, Roseberry et al. (2016) were able to activate neurons in the MLR that in turn generate locomotive activities.

The pathways starting with D2 neurons are referred to as *indirect* as they involve intermediate nuclei. D2 neurons send inhibitory projections to the globus palladus external (GPe). Since GPe neurons also send out inhibitory projections, in this case to the sub-thalamic nucleus (STN), the net effect (inhibiting inhibition) is to increase activity of STN neurons. These STN neurons then send excitatory signals to neurons in the output nuclei, enhancing their inhibitory action. Overall, the effect of the indirect pathway is to strengthen the basal ganglia's inhibitory outputs. Accordingly, Roseberry et al. also activated D2 neurons, and demonstrated reduced MLR activity, which prevented specific forms of locomotion.

Due to their opposite effects of activating and inhibiting target areas, the direct and indirect pathways are thought to carry respectively Go and NoGo signals for specific actions (Hazy, Frank, & O'Reilly, 2007). Whether an action is selected depends on the relative activations of D1 and D2 neurons: roughly, if D1 activation (Go signal) is stronger, an action is selected whereas if D2 (NoGo signal) is stronger, an action is inhibited. To perform this function, neither D1 nor D2 neurons need to encode any detailed information about specific actions; they only need to encode in the strength of the Go and NoGo signals the evaluation of the action generated elsewhere in the brain. What the basal ganglia do is effectively promote the action which receives the more positive evaluation and inhibits others (Bogacz & Gurney, 2007).

Among the brain areas affected by decisions made in the basal ganglia are areas in the neocortex. Indeed, neurons in the neocortex, basal ganglia, and thalamus are typically organized into loops in which output from the basal ganglia are directed back to the same locations from which inputs arise. The result is that the basal ganglia determine which among potentially competing processes in the neocortex are permitted to proceed (we will return to this role of the basal ganglia in determining the flow of processing in the neocortex in section 9.3).

Above we noted that the input neurons in the striatum are distinguished by the type of dopamine receptor they contain. Dopamine, as we have discussed in section 2.3 and 4.2, is a volume transmitter that acts as a neuromodulator. That is, it alters the response properties of these neurons to their inputs. The importance of dopamine in the basal ganglia is illustrated by Parkinson's Disease in which low dopamine levels results in tremors and difficulty in initiating voluntary movement. Although claims about what modulatory role dopamine performs are still contested, we describe two of the more generally accepted hypotheses, one involving phasic (fast and temporary) and the other tonic (slow and continuous) dopamine signals.

Phasic dopamine signaling is widely viewed as enabling the basal ganglia to learn to make better decisions by representing reward prediction error—the difference between the predicted reward and actual reward following an action. Reinforcement learning is a type of learning that reinforces choices that result in rewards beyond those expected. The basic idea is that if the selected action generates the expected reward, no further learning is needed. If it generates more reward than expected (a positive reward prediction error), then that action should be reinforced. If it leads to less reward, then the selection should be attenuated. A phasic dopamine increase indicates a positive reward prediction error. It strengthens responsiveness of D1 neurons (the source of the direct pathway that carries a Go signal) and weakens the responsiveness of the D2 neurons (the source of the indirect pathway that carries a NoGo signals). As a result, the action is more likely to be selected in the future. A phasic dopamine decrease results in the opposite effect.

In contrast, tonic dopamine levels are thought to control how an organism addresses the tradeoff between exploiting the current situation or exploring elsewhere for potentially better opportunities (Chakravarthy & Balasubramani, 2018). The example discussed in section 2.3 illustrated this tradeoff in the context for seeking foods. The choice arises more generally: in selecting new music to listen to, a person must decide whether to continue with other compositions by the same artist (often successful when one enjoyed the current composition) or to explore those of other artists. One computational model suggests that lower tonic dopamine levels heighten the random fluctuation in the NoGo signals for the competing options, such that alternatives to the option represented as better will sometimes be selected, resulting in exploration. Higher level of dopamine dampens the fluctuation, resulting in pursuit of the choice currently judged to be best (exploitation).

5.5 Summary

In this section we have presented brief accounts of four areas in which neuroscience has provided an understanding of activity in the vertebrate brain. We will have occasion to return to each of these in subsequent sections.

6. How do neuroscientists explain activities of the nervous system?

A major goal of any science is to explain phenomena—regularities in the world that researchers take to be important to account for (Woodward, 2019). Some phenomena are immediately obvious: organisms seek and eat food. But many phenomena are only discovered through extensive research. It took over a hundred years of exploring electrical activity in animals to recognize that most neurons transmit action potentials (section 2.1). Detailed studies using mazes were required for Tolman to conclude that rodents navigate using cognitive maps (section 5.2). Sometimes researchers conclude that what were taken to be phenomena don't actually occur and so don't require explanation. For example, despite the claims of several researchers that worms could transfer what they learned to those that ate them, researchers concluded that this doesn't actually happen (for a philosophical analysis of this case, see Colaço, 2018).

In this section we explore different accounts philosophers have offered of how scientists explain phenomena. Proponents of each account present them as characterizing explanations advanced in neuroscience. An important question is whether these accounts are competing or can be integrated into a common account. To the degree they are competing, further questions arise: does one need to choose between them? or might there be multiple modes of explanation?

6.1 Mechanistic explanation

Many biologists since the 17th century have viewed functional components within biological organisms as comparable to machines that humans construct. A key feature of machines is that they are composed of parts that carry out different activities. These activities are coordinated so that the whole machine generates a phenomenon that none of its parts alone can produce. Comparably, mechanistic explanations appeal to the composition and organization of a mechanism to explain a phenomenon. (For further discussion of mechanistic explanations, see Machamer, Darden, & Craver, 2000; Bechtel & Abrahamsen, 2005; Craver & Tabery, 2019).

Much of the philosophical discussion of mechanistic explanation has focused on how researchers develop such explanations (Bechtel & Richardson, 1993/2010; Craver & Darden, 2013). The first step is to localize the phenomenon in a particular system—a mechanism—that is taken to be primarily responsible for producing it. For example, research on circadian rhythms in mammals (section 5.1) identified the SCN as the responsible mechanism. Researchers studying spatial navigation in rodents identified the hippocampus as the locus of cognitive maps (section 5.2). In these cases, deficits that resulted when the putative mechanism was damaged provided evidence for identifying the mechanism. In other cases, other strategies

provided the basis for localization. Recording from neurons while presenting visual stimuli played a central role in determining the components of the mechanism of visual processing (section 5.3), while the architecture of the basal ganglia motivated viewing it as a decision-making mechanism (section 5.4).

After picking out the responsible mechanism, the key step in advancing a mechanistic explanation is to take it apart—to decompose it. There are two ways to do this—by identifying the mechanism's physical parts or by identifying the operations required to produce the phenomenon. Although identifying the physical parts is relatively straightforward with human built machines, this can be challenging with biological mechanisms, as we saw in section 2. Parts don't necessarily have well-defined boundaries. Golgi and Cajal debated whether neurons were separate units. Brodmann contested the traditional differentiation of cortical areas in terms of gyri and sulci, proposing instead that we use as boundaries the points where layers he identified in the neocortex changed thicknesses. One reason that identifying parts is challenging is that the goal is to pick out the parts that perform operations relevant to the phenomenon (what Craver, 2007, refers to as working parts). Not any way of cutting up the mechanism (e.g., chopping it into cubes) will be informative.

Differentiating operations requires different strategies than used to identify parts. Consider, for example, the difference between distinguishing people and distinguishing occupations. One can distinguish people by their physical traits, but to distinguish occupations one must determine what jobs the people perform. To identify the operations involved in generating a phenomenon, one needs both to reason about the operations that could produce the phenomenon and to find ways to intervene on the mechanism to see that they are indeed all carried out in the mechanism.

Reasoning from the phenomenon to the operations required is challenging. One of Gall's shortcomings that we didn't emphasize in section 3 is that he simply focused on traits on which people differ. He did not try to decompose these traits into the operations needed to realize them. Likewise, Broca characterized what the area that bears his name does in terms of the overall phenomenon—producing articulate speech—not the operations required to generate speech. The challenge in identifying operations is that, in most cases, operations are not described in the same vocabulary as the whole mechanism. Sometimes there is already a well-developed language for describing the operations performed by the parts of mechanisms (if individual neurons are the relevant parts, there is a rich vocabulary for describing their electrical and chemical activities). In many cases such a vocabulary does not exist and researchers must create it. This often requires proposing task-analyses for the activity of the whole mechanism (if the mechanism performed operations A, B, and C then, by doing them in sequence, it would perform the overall activity) and then seek evidence that the mechanism actually performs those operations.

As difficult as it is to find one decomposition of a task, there are typically multiple task decompositions that would suffice to generate the phenomenon (this is why, once one designer creates a machine, competitors can often come up with their own designs that do not violate

the original designer's patent). To figure out which is actually implemented in a given organism, researchers must try to localize the operations in different parts and then empirically investigate whether the parts actually perform those operations, using research strategies such as we described in section 3.

In real science, there is often a prolonged period of revision both in the characterization of the parts and of the operations. For example, in developing the account of visual processing we described in section 5.3, researchers subdivided Brodmann's areas 18 and 19 into V2, V3, V4, and MT/V5. And once researchers had advanced the hypothesis that MT/V5 is involved in motion detection, other researchers began to identify other operations it performs in addition to motion detection. In the end, though, the goal is to be able to map parts onto operations.

Beyond decomposing a mechanism into its parts and operations, there is a further step in characterizing the mechanism. Sometimes when we buy a product, the box will say "some assembly required." Until the parts are assembled, the product won't do what we bought it to do. Likewise, even if researchers had the correct account of the working parts of a mechanism and what operation each performs, they wouldn't have a complete mechanistic explanation. Researchers must determine how they are organized so that the products produced in one operation can be further acted upon by other operations. To represent possible modes of organization, scientists often construct diagrams, such as Figure 14B. In this diagram, rectangles represent brain regions while the icons in them represent the operations each is thought to perform. The lines connecting the rectangles, commonly referred to as *edges*, represent pathways between regions, with thickness indicating the prominence of the pathway. In other mechanism diagrams, yet other conventions are used (Abrahamsen, Sheredos, & Bechtel, 2018). While the diagram is static, it provides a basis for humans to reason about how the whole mechanism generates a phenomenon (in this case, recognizing an object or locating it in space) by performing the operations portrayed in the order shown. In this manner, researchers can mentally simulate the operation of the mechanism.

In thinking about the organization of a mechanism, humans start by thinking sequentially. The edges in Figure 14B are thought to carry activity from inputs at the bottom upwards to higher processing areas. But the researchers who developed the diagram were very much aware that in the brain there are as many recurrent projections (neural projections from areas viewed as later in a pathway to those viewed as earlier) and that each of these areas sends and receives projections from regions of the thalamus and the basal ganglia (see section 5.4). One can add additional rectangles and arrows to represent these, but it quickly becomes impossible to simulate the mechanism mentally. Instead, researchers often supplement a verbal and diagrammatic representation of a mechanism with a mathematical one, developing a computational model (section 3.5). We will illustrate this in section 6.3, but first we turn to an account of explanation that proposes using computational models to supplant the need for mechanistic accounts.