In introducing representations, we used linguistic phrases as examples. How does a word like "neuron" represent neurons? There is nothing about the word itself that determines what it represents. Waldeyer, who invented this term, might have coined a different term. Rather, its meaning depends on the conditions in which language users insert it into sentences and, especially, how they respond to it when they encounter it being used by others. This raises the question as to whether a similar account applies to neural activities characterized as constituting representations. Viewing neural activities as states in control mechanisms already suggests how a similar account can be developed: When they are generated and used in controlling other mechanisms, neural mechanisms treat them in ways similar to how language users treat linguistic phrases.

Further support for interpreting neural processes as representations is that working scientists investigate how they possess the ability to represent. When you encounter a word or phrase you do not know, you investigate when people use it and how others respond to it. Similarly, after O'Keefe characterized neurons in regions of the hippocampus as place cells, he and others set out to identify how they come to respond to particular locations and how they figure in the animal's behavior. As discussed in Section 5.2, researchers manipulated environments to see which would produce changes in a neuron's response. They also showed how activation of these neurons before and after rodents ran on paths occurred in anticipation and recall. At a minimum, the researchers are not just glossing an already described neural mechanism, but taking seriously the hypothesis that it functions as a map-like representation and investigating how it does so. That is, they, treat these neurons as actually representing places (Bechtel, 2016).

8.3 Do Neural Processes Accurately Represent the World?

If one accepts that neural processes do represent the world, a further question is: Do they *accurately* represent the world? Many of the approaches to ascribing content to neural processes assume that that is what they are doing. In characterizing what different visual areas do, researchers presented stimuli with a given feature and treated the neurons that responded most strongly as representing that feature. One might also wonder what would be the point of a perceptual system that did not accurately represent the world since that would seem to defeat the goal of successful interaction with the world.

Using temperature perception as an example, Akins (1996) argues that our sensory systems do not accurately represent the world. Rather, she characterizes our temperature receptors as narcissistic – they respond when the

https://doi.org/10.1017/9781108946964 Published online by Cambridge University Press

temperature of objects we are touching is too hot or too cold. Moreover, they do so in a nonlinear fashion: as temperature increases, the hot receptor will first generate spikes at a greatly increased frequency before gradually dropping to a frequency slightly above what it was at the outset. When the stimulus terminates, it stops spiking altogether before gradually increasing to its default rate. The cold receptor operates similarly. Such a system signals major changes in temperature, but unlike a thermometer, it does not give a readout that corresponds to the temperature of the object touched. Moreover, responses are contextually sensitive to immediately preceding experiences, as illustrated in a familiar demonstration: put one hand in hot water, the other in cold, and then move both to water of intermediate temperature. It will feel like the hand previously in cold water is in warmer water than the one originally in hot water.

While such a sensory system does not generate accurate representations of temperature, Akins argues that it does alert the organism to what it needs to know – is something too hot, such that I should drop it or avoid touching it? Receptors that acted like a thermometer would be less efficient – the nervous system would have to incorporate the temperature information into a plan for action rather than responding directly. Does Akins' argument support treating the senses and the neural processing systems downstream from them as generating narcissistic representations or alternatively, as "nonrepresentational systems"? In fact, Akins has argued for the stronger, nonrepresentationalist, conclusion: sensory systems are not only narcissistic but function through nonrepresentational feedback processes. However, it is not clear why one cannot view the activity of the sensors as constituting context-sensitive narcissistic representations. Consider again the Watt governor – one may argue that it does not represent the speed of the flywheel per se, but only represents, narcissistically, that it is moving too quickly or too slowly.

8.4 Summary

We have characterized different views about whether brain processes count as representations. What is worth highlighting is how these views tie to views of explanation discussed in Section 6. If one adopts the view that the nervous system embodies control mechanisms, one needs to characterize it as making measurements, and hence talk of representations seems motivated. If one adopts a dynamical system view that eschews mechanisms, then one can view these processes as elements of dynamical systems without treating them as representations.

https://doi.org/10.1017/9781108946964 Published online by Cambridge University Press

9 What Is Distinctive about the Neocortex?

As we noted at the outset, the neocortex is the brain area that has expanded the most in primates, including humans. It is clearly important for human life, especially for those activities that humans distinctively perform. But, as we have stressed through this Element, other brain structures are also important. With few exceptions, the neocortex does not take over their activities but supplements them. The relevant question is: What is the distinctive type of processing that occurs in the neocortex? One suggestion comes from the studies of decorticate cats discussed in Section 5.4. While cats in which the neocortex is removed can live in protected environments, they would be unlikely to fare well in the world in which they confront variable conditions, including predators. Based on these studies, Buchwald and Brown (1973) proposed that the neocortex serves for detailed analysis of stimuli, extracting and representing complex and subtle information about an organism's environment and identifying relations between different bits of information. Such information is extremely useful in solving problems posed by a variable environment. In this section, we investigate how the neocortex can perform these tasks.

9.1 (Artificial) Neural Networks and Pattern Extraction

The neocortex is organized in a distinctive manner that supports the hypothesis that it acts to extract subtle and complex information from sensory inputs. While many brain areas, including both the basal ganglia and the hypothalamus, are organized as interconnected nuclei, the neocortex is laid out much more systematically. As we discussed in Section 2.4, Brodmann (1909/1994) differentiated areas within the neocortex based on the thickness of layers identified in stained cortical tissue. Tracing axons from neurons in one area reveals that they mostly project to selected neurons in other specific areas, resulting in relatively orderly anatomical hierarchies such as shown in Figure 14(b). At the top of the figure are areas in the temporal and parietal lobe. Both streams, however, continue into the frontal cortex, reaching the far frontal area known as the *prefrontal cortex*, on which we will focus in Section 9.4.

To see how such a (anatomically) hierarchically organized network could enable the extraction of information, consider *artificial neural networks* (ANNs) – computational systems that were inspired by the architecture of the neocortex. As illustrated in Figure 23, these networks consist of layers of artificial neurons, commonly referred to as *units*. A weighted connection links a unit in one layer to units in the next higher layer; in processing, the weight is multiplied by the activity value of the unit in the lower layer to determine an input to the higher-level unit. Each higher-level unit accumulates these inputs

https://doi.org/10.1017/9781108946964 Published online by Cambridge University Press