

Table 1.1
(Continued)

	Propositional	Depictive
	Abstract: Can refer to non-picturable entities	Cannot be abstract in any sense: Refers to picturable entities
	Can refer to classes of objects	Refers to exemplars
	Not tied to any specific sensory modality	Specific to a particular sensory modality
Rules of Combination		
	Arrangement of symbols can assert a true or false state of affairs	As spatial relations among points are distorted, the depiction becomes increasingly less accurate

the right leg, and so on. Such spatial relations are immediately accessible in a depictive representation, and thus depictive representations may be more efficient in representing the spatial layout of objects. On the other hand, the fact that there are three “lines” (i.e., strokes) is explicit in the descriptive representation but not in the depictive one. In the depiction, all portions of the pattern are both explicit and accessible, but they are not semantically interpreted; thus, to access the concept of “line,” a process must interpret the pattern. In contrast, the propositional representation makes this interpretation of the pattern explicit and accessible.

In sum, depictive representations make explicit and accessible all aspects of shape and the relations between shape and other perceptual qualities (such as color and texture), as well as the spatial relations among each point. In contrast, propositional representations make explicit and accessible semantic interpretations, which can include aspects of shape and other perceptual qualities. Depictive representations of shape must also incidentally specify size and orientation; propositional representations only specify what was explicitly included when the representation was created. Depending on the precise task at hand, one or the other format may be most useful.

Depictions in the Brain

As we saw in the computer analogy described above, there need be no actual picture in the brain to have a depiction: all that is needed is a “functional

space” in which distance can be defined vis-à-vis how information is processed. (However, as we discuss in chapter 2, a functional space is sufficient as a depictive form of representation only if the geometric properties of the representation emerge because there are fixed, hard-wired processes that interpret the representation as if it were a space; if the processes are not fixed, then the representation is not necessarily a depiction.)

Nevertheless, even though all that is required in order to have a depiction is a functional space, there is good evidence that the brain depicts representations literally, using space on the cortex to represent space in the world. To be specific, we will argue in the following chapters that images rely in part on areas in the brain that are specifically designed to depict patterns. These areas are *topographically organized*—they preserve (roughly) the geometric structure of the retina. Such areas use space on the cortex to represent space in the world (e.g., see Felleman & Van Essen, 1991; Fox et al., 1986; Heeger, 1999; Sengpiel & Huebener, 1999; Sereno et al., 1995; Tootell, Hadjikhani, Mendola, Marrett, & Dale, 1998; Tootell, Silverman, Switkes, & De Valois, 1982; Van Essen, 1985). For example, figure 1.2 illustrates the results of an experiment reported by Tootell, Silverman, Switkes, and De Valois in 1982. They trained a monkey to stare at the pattern shown on the left, which consisted of a set of blinking lights arranged as shown. The animal was injected with a radioactive form of sugar, which was taken up into brain cells in proportion to how active the cells were while the animal observed the pattern; the more active the brain cell, the more sugar it took up. This particular isotope gets lodged in the neurons and is not quickly broken down by metabolic processes. The animal was then sacrificed, and its brain was removed. Figure 1.2 illustrates the first cortical area to receive input from the eyes, known variously (but synonymously) as area V1, area 17, area OC, the striate cortex, and the primary visual cortex. The dark bands in the right part of figure 1.2 label brain cells that took up a lot of the radioactively tagged sugar. As is clear, the geometric structure of the stimulus is physically laid out on the cortex!

These areas do not simply have a topographically organized physical structure; they *function* to depict information. If a patch of cortex in one of these areas is damaged (for example, because a tumor had to be removed), this damage will produce a scotoma (i.e., a “blind spot”) in the corresponding part of the visual field. The scotomas that arise when the topographically organized visual cortex is damaged demonstrate conclusively that these areas function to depict information; crucially, the closer two damaged regions of the topographically organized visual cortex are, the closer in the visual field the corresponding scotomas will be. And this result is not simply about the effects of chronic damage: transcranial magnetic

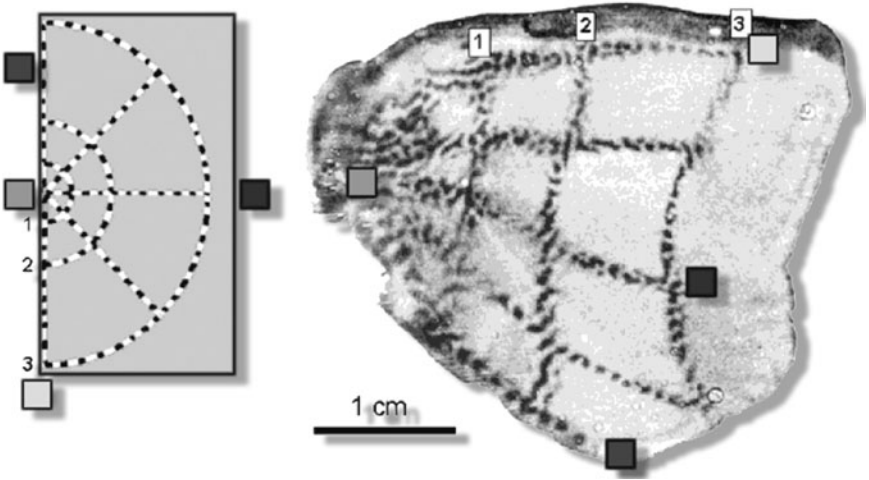


Figure 1.2. A geometric pattern of flashing lights (*left*) shown to a monkey after it was injected with a radioactive sugar, which is taken up into the brain cells in proportion to their level of activity. The animal was trained to stare at the stimulus, and then was sacrificed so that its brain could be examined. The occipital cortex of the monkey (*right*) is shown with dark bands revealing the neurons that were most activated while the animal viewed the pattern. A pattern of activation corresponding to the geometric structure of the stimulus can be seen clearly. This result demonstrates that visual stimuli are represented topographically in the occipital cortex of the monkey brain. Shaded squares are used to demarcate the points on the visual stimulus that are represented in corresponding space on the visual cortex. The numbers 1, 2, and 3 indicate the inner, middle, and outer bands of flashing lights that are clearly represented on the cortex. Note that for simplicity, only half the circular light stimulus is shown here, along with a section of V1 from a single hemisphere of the brain. (Adapted with permission from Tootell et al. [1982]. *Science*, 218, 902–904; and Tootell et al. [1988]. *Journal of Neuroscience*, 8[5], 1531–1568.)

stimulation has been used to stimulate occipital cortical sites transiently to produce phosphenes (i.e., bright flashes of light that are not produced by sensory input); when nearby sites are stimulated, phosphenes appear in nearby locations in space, and when far apart sites are stimulated, phosphenes appear in far apart locations in space (e.g., Kastner, Demmer, & Ziemann, 1998).

The appearance of figure 1.2 notwithstanding, the pattern of activation is not literally a picture of what the animal is seeing. From the point of view of an outside observer (such as the reader, observing that illustration), metric information on the cortex is distorted. However, as we discuss at length in

the following pages, the key is how processes that receive outputs from this cortex interpret distance on the cortex—not how a human observer armed with a ruler would interpret such distances. If we were to treat the cortex like a piece of paper, we would conclude that the depiction is distorted, with the central part amplified—but the brain processes that access this representation correct for such distortions. Within the context of the brain as a whole, these topographically organized areas truly depict information.

Why does the brain use space on the cortex to represent space in the world? Although the ease of genetic coding or other factors may play a role, the best current guess is that this structure has been retained through evolution for a simple reason: this trick makes explicit and accessible information needed for the tasks at hand.⁵ For example, the first cortical visual areas to receive input from the eyes are confronted with the task of organizing figure from ground. In order to do so, they must delineate edges. This task is facilitated by the fact that many (perhaps most) of the connections among neurons in a topographically organized area are both very short and inhibitory. This means that if one neuron is stimulated, it attempts to inhibit its neighbors—which represent adjacent points in space. The effect of this is to exaggerate differences in neural activation across the boundaries of edges. For instance, if an object reflects a lot of light and the background does not, then neurons that are near the edges of objects will inhibit those that register the background, which will amplify the difference in their activation—thereby helping to identify edges of objects.

For mental imagery, purely functional depictions may have been sufficient; there is no obvious reason why physical depictions are required. However, because the imagery system draws on mechanisms used in like-modality perception (as we demonstrate throughout the remainder of this book), it relies on such physically topographically organized structures. As we argue, many of the properties of imagery arise because of this simple fact. For instance, because the input to early topographically organized areas changes every time the eyes move, the patterns of activation within them cannot linger for long; if they did, the world would seem smeared as we moved our eyes. But what is a virtue for perception is a drawback for imagery: as we argue in chapter 5, it is difficult to maintain images for long, in part because they rely on neural machinery also used in perception.

In short, each location on the topographically organized cortex corresponds to a specific location in space, and distance between the locations on cortex corresponds to distance between the corresponding locations in space. The brain has numerous such representations, but we shall emphasize two classes of them in this book. First, many of these depictive

representations are involved in processing shapes, particularly in the occipital lobe. We group the topographically organized areas in the occipital lobe into a single functional structure, which we refer to as the *visual buffer*. Patterns of activation within the visual buffer depict shapes, according to the definition of *depiction* offered here. The visual buffer, in essence, is the canvas upon which images are painted; it is the medium that supports depictive representations. In later chapters we provide strong evidence that topographically organized brain areas are in fact used in visual mental imagery, but first we will use the idea that imagery relies on patterns of activation within such areas to show that the arguments levied against depictive representations do not hold water.

Second, the brain also uses depictive representations to specify information about the locations of objects in space. These representations are primarily in the posterior parietal lobes (e.g., Sereno, Pitzalis, & Martinez, 2001). Although a depictive format is used both to represent shape and location, we shall see that the contents of these representations differ markedly. As we discuss in later chapters, because of the differences in content, we will distinguish between *object images*, which represent shape (and shape-related properties, such as color and texture), versus *spatial images*, which represent relative locations in space.

In the most recent round of the debate, at stake is the very idea that turning to the brain can inform theories of cognition. We will argue that the function of the brain—which is, after all, the organ of thought—has evolved in tandem with its structure, and vice versa: the brain is not a “general purpose computer,” like a von Neumann machine (such as your personal computer), which acts very differently depending on the program it has in memory. Instead, the brain is largely a special-purpose machine, which is tailored to function in specific ways. Because different formats make different information explicit and accessible, different formats are more or less useful for performing different tasks. Thus, an efficient strategy for the brain is to use different formats in different situations, and animals that could do this may have had an adaptive advantage over their less specialized brethren.

Hybrid Depictive Representations

Consider again the gripping illustration in figure 1.2. Those dark lines indicate where neurons were particularly active while the monkey was observing the pattern. If we take this finding to indicate that a depictive representation is being used (as we should, given our characterization of such representation), we are led to conclude that mental images are not like

of visual mental imagery that can be empirically tested and evaluated in terms of generality, parsimony, and so on. As it now stands, there are simply too many possible varieties of depictive and propositional theories, theories that have appreciably different properties, to allow any sort of evaluation of relative efficacy on purely rational grounds. In chapter 5 we present one example of such a theory.

Finally, we must confront Anderson's (1978) demonstration that one can always produce a new theory by changing the representation of one theory and then compensating for that change by altering the theory of processing. The problem is that nothing is nailed down in advance: there are no a priori constraints. Anderson also pointed out, however, that the brain can serve to constrain this wanton arbitrariness. A theorist *cannot* make up properties of the brain—such as functions of specific brain areas and the anatomical connectivity among areas—simply to account for data. Thus, cognitive theories can be *constrained* by such facts about the brain; the facts do not dictate the theories, but they limit the range of what can be posited. Facts about the brain anchor theories in such a way that theorists cannot invent alternative accounts for sets of data by changing theories of representations and processes at their convenience.

Notes

1. Most of the arguments against depictive representations are addressed to the topic of visual imagery. We will follow suit, although the reader should keep in mind that parallel arguments also may be levied for and against imagery in other modalities.

2. The anti-imagery arguments are gleaned primarily from Anderson and Bower (1973), Clark and Chase (1972), Dennett (1969, 2002), Pylyshyn (1973, 1975, 1981, 2002, 2003), and Reid (1974). Unless otherwise noted, the arguments against depictive representation are drawn from Pylyshyn's original paper (1973) and his follow-up paper in 1981.

3. Much of this chapter is built around chapter 2 of Kosslyn (1980), which in turn was based on the Kosslyn and Pomerantz (1977) paper. James Pomerantz deserves credit for helping develop the arguments offered here.

4. The following section is adapted from Kosslyn (1983), with permission of the publisher.

5. Unless otherwise specified, we will use the terms *image* and *imagery* to refer to depictive representations. After all, it is only such representations that are truly "images."

6. Millar (2002) notes that Pylyshyn's assertion is "odd in the context of evolutionary biology" (p. 202). We will not go down this path here, but it is likely to lead to yet another line of disagreement between the camps.

7. Most initial distinctions in science—especially in biology—are qualitative; the quantitative characterization of relations only comes later, after a phenomenon is relatively well understood. We acknowledge that there is much work left to be done, but this is not to say that the distinctions we have already drawn and supported cannot accomplish much explanatory work.

8. Note that even current measures of the amount of information contained in single neuron spike trains (e.g., Osborne, Bialek, & Lisberger, 2004) say very little about overall capacity limitations of the human brain as a collection of interacting neural networks.

9. The point being made here is simply that people can be aware of the phenomenology of the image itself, not simply the information being conveyed. However, the fact that imagery vividness ratings are reliable (Marks, 1973) does not imply that they necessarily reflect specific aspects of information processing—this is a separate issue.

10. Anderson (1978) claims that a map will have additional information indicating absolute orientation. Usually this is indicated explicitly by a compass or direction arrow; however, if a “north point” is placed top and center and this point is included in the intercity distances, a multidimensional scaling solution will then also recover this orientation information. In any event, it is not clear whether the absolute orientation information is a property of the map itself or of how the reader can process it.

11. Although it is possible that measurements derived from information theory (e.g., Osborne, Bialek, & Lisberger, 2004) may one day provide realistic estimates of brain storage and processing capacity, these measures would still address only the content, not the format, of representations. By analogy, one could characterize the processing capacity of a computer but not know which high-level programming language is running on that computer.

4

Depictive Representations in the Brain

The imagery debate hinges on the distinction between the format and the content of internal representations, the format being the type of code used to represent the shape of an object, and the content being the information conveyed. The issue is whether mental imagery makes use of representations that depict (perhaps in addition to those that also describe). Following many years of behavioral research, this debate reached an impasse: the results from these studies could be explained by theories that posited only propositional representations or by theories that also posited depictive representations (e.g., see Anderson, 1978). Modern neuroimaging techniques now provide an opportunity to press this issue forward, for two reasons.

First, researchers have documented that approximately thirty-two areas of the monkey cerebral cortex contain neurons that respond selectively to visual input; moreover, about half of these areas are topographically organized—they preserve (roughly) the geometric structure of the retina. Such areas use space on the cortex to represent space in the world (e.g., see Felleman & Van Essen, 1991; Fox et al., 1986; Heeger, 1999; Sengpiel & Huebener, 1999; Sereno et al., 2001; Tootell, Hadjikhani, et al., 1998; Tootell, Silverman, et al., 1982; Van Essen, 1985). As discussed in chapter 1, these areas are not simply physically topographically organized—they function to depict information. The representations in topographically organized areas are depictions, not propositional descriptions.

Positron emission tomography (PET; e.g., Fox et al., 1986) and functional magnetic resonance imaging (fMRI; e.g., Sereno et al., 1995) have shown that such depictive areas are used during visual perception in the human brain. If these areas are activated when one visualizes, and disruption of these areas impairs the ability to visualize, this is strong evidence that the representations underlying visual mental images are not entirely propositional; rather, such evidence would show that at least some of these representations depict information and that such representations play a role in information processing.

The second reason neuroimaging results offer a way to break the impasse in the debate about imagery is that such findings cannot be explained by appeal to tacit knowledge. As we discussed in the previous chapter, some researchers have claimed that the results from studies of imagery arise because task demands are built into the experiments, which in turn activate tacit knowledge about perceived events (a phenomenon that is largely unconscious, such as our knowledge of how the surface of water will appear when a glass of water is tilted), which in turn leads participants to produce responses like those they believe (perhaps unconsciously) would occur in the corresponding perceptual situation. Findings from the brain would be truly definitive evidence against this position: no ordinary person has knowledge of how perception (and hence, according to the tacit knowledge account, also imagery) is processed in the brain—and thus people are not in a position to manipulate their brain activation (even if they had that ability) in accordance with tacit knowledge.

In the first part of this chapter, we summarize the arguments against the idea that topographically organized areas are used in visual mental imagery and the arguments that even if they are used, they play no essential role. In the second part, we address each argument in turn, considering counterarguments and evidence to the contrary. In the third part, we discuss a meta-analysis of studies of visual mental imagery (summaries of the studies are provided in the appendix). This meta-analysis untangles what might at first blush appear to be inconsistencies in the literature and provides strong support for the claim that topographically organized areas support depictive representations during visual mental imagery.

Do Topographically Organized Areas Depict Information?

Some researchers have tried to dismiss the evidence that depictive representations are functional in the brain. Such arguments can be divided into

three classes, summarized below. We first summarize the arguments, and then consider them in detail.

Arguments against Neural Depiction

Some critics of the recent neuroimaging research have claimed that topographically organized areas in the human brain are not truly depictive. For example, area 17 (also known as area V1, the striate cortex, and the primary visual cortex) is organized so that objects that subtend small visual angles activate posterior portions of the area, and objects that subtend larger visual angles activate more anterior portions of the area (e.g., Fox et al., 1986). Larger objects do not necessarily occupy a larger amount of the area, so how can we infer that the area is truly depicting information?

In addition, the idea that topographically organized regions in the brain support depictive representations purportedly is flawed because vision does not accumulate information in an internal display to produce a panoramic view of the world; this is a subjective illusion. Every time the eyes move, the information in early visual structures is replaced by the new input. So, how can theorists claim that neural structures used in vision also support depictive representations in mental imagery? In criticizing our claim that topographically organized visual areas depict information, Dennett (2002, p. 190) asserts: “It has yet to be established when and how *vision* utilizes image processing!” He continues:

Vision isn't television. The product of vision is not a picture on the screen in the Cartesian Theater (Dennett 1991). The fleeting retinal images punctuated by saccades are the first images, and they are not the last, as Julesz (1971) demonstrated by showing perception of depth in random dot stereograms that requires image-processing after the optic chiasma. But which subsequent cortical processes also exploit any of the informational properties of images? The *eventual* “products” of vision are such things as guided hand and finger motions, involuntary ducking, exclamations of surprise, triggering of ancient memories, sexual arousal . . . and none of these is imagistic in any sense, so assuming that the events in their proximal causal ancestry are imagistic is rather like assuming that power from a hydroelectric plant is apt to be wetter and less radioactive than power from a nuclear plant. The raw retinal data are cooked in many ways betwixt eyeball and verbal report (for instance). How cooked are the processes involved

in (deliberate or voluntary) mental imagery? (Dennett, 2002, p. 190)

Moreover, continuing this line of argument, critics have claimed that the very idea that the brain would create a copy of the visual world is flawed. This idea is nicely captured in the following quotation: “There is no evidence of the mental construction of images to be looked at or maps to be followed. The body responds to the world, at the point of contact; making copies would be a waste of time.” These words were written by B. F. Skinner (1977, p. 6) but could have been written by many other critics of depictive theories of imagery. Pylyshyn (2002, p. 182, n. 14) asserts that vision “is not for turning one topographically organized pattern into another.” Echoing this point, Dennett (2002) emphasizes that vision is used to identify objects and their properties and to guide and track movements—and none of these eventual goals directly relies on topographic maps. Dennett (2002) notes that electricity generated from hydroelectric plants is not wetter than electricity generated by other means—and thus, by analogy, topographically mapped areas may have no bearing on the representations actually used in vision.

Clarifying the Nature of Depictive Representation

Here we discuss why the arguments raised against depictive representations in the brain miss their intended target.

Representation within Topographically Organized Areas

Many neuroimaging findings have documented topographic organization in several areas of human visual cortex during perception (e.g., DeYoe, Bandettini, Neitz, Miller, & Winans, 1994; Engel et al., 1994; Hasnain, Fox, & Woldorff, 1998; Sereno et al., 1995; Van Essen et al., 2001). The high-resolution central (foveal) portions of these areas represent the central portions of the visual field, whereas low-resolution, peripheral portions of these areas represent more peripheral portions of the visual field. Space on the cortex is literally used to represent space in the world (more precisely, space on the cortex is used to represent a planar projection of space in the world, because the cortical areas are two-dimensional; information about depth is represented in a different manner, e.g., Tsao et al., 2003). Propositional representations have the property of being arbitrarily related to the objects they represent, whereas the relationship between an external visual

stimulus and the activation it engenders in topographically organized areas is not arbitrary: points that are relatively close in space in the external world are represented by activation in nearby portions of these areas, and points that are relatively far in space are represented by activation in relatively separated portions within each of these areas.¹

What about the question of why larger objects do not produce larger swathes of activation within these brain areas? This concern is based on a misunderstanding of how information is represented in topographically organized areas. First, neurons in the anterior portions of area 17 have larger receptive fields, which receive inputs from a larger region of space than do neurons in more posterior portions. Thus, activation in more anterior portions of this structure indicates that parts of the image are spaced farther apart and are less resolved than is the case when activation is in more posterior portions. This point is worth repeating, since it is often not appreciated: activation toward more anterior portions of this structure not only depicts larger extents but does so with poorer resolution. However, keep in mind that what is important is not how the area looks to an external observer but how it is interpreted by processes that operate on it in the brain. From this perspective, when neurons in the anterior portions of area 17 are activated, this is interpreted as specifying a larger portion of an object's surface.

The spatial organization of topographically organized areas in the human brain is complex, in part because numerous topographically organized areas must be connected to preserve the overall topography. Adjacent topographically organized areas meet at the horizontal meridian (i.e., the horizontal line through the fixation point in the frontal plane of one's visual field) or at the vertical meridian (i.e., the vertical line through the fixation point in the frontal plane). The most common way to establish that areas are topographically organized takes advantage of these facts. In this method, participants observe a narrow flashing wedge as it rotates slowly around a pivot, starting from the vertical. fMRI is used to track changes in activation over time as the stimulus moves. If the stimulus is vertical at the outset, activation will initially be detected along the vertical meridian, and as the stimulus rotates, this activation literally shifts over the cortex, toward the horizontal meridian. Because adjacent visual cortical areas abut at the horizontal (or vertical) meridian, this means that the wave of activation in one area will literally shift until it meets an incoming wave of activation from an abutting area—and researchers use the places where waves meet to establish the boundaries between topographically organized areas. We wish to emphasize that these waves reflect the precise position of the stimulus in

space: as the stimulus moves, the activation is mapped, point-for-point, into the topographically organized areas. Each neuron in these areas registers a part of space, and the distance between neurons reflects the distance between the corresponding parts of space. However, the physical distance on cortex also reflects the resolution—with the same extent on an object's being represented by greater distances in the foveal region of cortex than in peripheral regions.²

And this brings us to the crucial point: connections to other areas compensate for the distortions in the actual (physical) representation—and thus the result is that the topographic areas function to depict shape accurately. This is a variant of the “Modigliani effect,” named after the Italian artist Amedeo Clemente Modigliani (1884–1920), who famously painted and sculpted elongated figures. The question was once raised whether the painter actually saw the world the way he portrayed it. However, if Modigliani had in fact had distorted vision, which led him to see objects as elongated, the same distortions would operate when he painted, thereby compensating and leading him to produce a veridical depiction. But in the case of the brain, later brain areas compensate for the distortions introduced into earlier visual areas. Topographically organized areas depict, but do so within the context of the system as a whole.

The Role of Topographically Organized Areas in Perception

We must distinguish two separate issues. First, are depictive representations actually used in vision at all, no matter whether it is for a single fixation or only a portion of the visual field? Second, if so, how are such representations combined successively to produce the impression of seeing the entire visual field?

Regarding the first question, Dennett's (2002, p. 190) bald claim that we do not know whether vision uses depictive representations is a vast overstatement. The neuroanatomy reveals that these areas play a role in the earlier phases of visual processing, and neurological data indicate that these areas play a special role in the parsing of shapes into their component parts. For example, apperceptive agnosia results when there is diffuse neural death in these areas (usually due to carbon monoxide poisoning), which in turn characteristically disrupts the ability to organize input into objects and their parts (Vecera & Gilds, 1998).

The mere fact that, as noted in previous chapters, local damage to topographic areas produces localized scotoma demonstrates that the topographic properties of these areas play a role in vision. Moreover, studies

in the macaque monkey, such as those by Roelfsema and Spekreijse (2001), clearly demonstrate that the visual topography of area V1 has a functional role in visual cognition. In this and related studies, monkeys had to trace a curve mentally (out of a set of two), without moving their eyes. During a random subset of trials, the monkey made a mistake and traced the wrong curve. These mistakes could be predicted by looking at the pattern of neural activation corresponding to the topographic representation of the two curves!

Certainly, we do not yet know everything there is to know about the functioning of topographically organized visual brain areas—but this is not necessary for present purposes. The depictive properties of such areas clearly play a role in visual perception.

Regarding the second question, what about the issue of how information from successive fixations is accumulated? Ingle (2002) notes that every time the eyes move, the material in the visual buffer (i.e., the set of topographically organized areas in the occipital lobe) is written over. Therefore, the buffer itself cannot integrate information from successive fixations. This fact is, however, completely irrelevant to the issue of whether visual processing exploits depictive representations. That said, we note that every time one's eyes move, the location of the stimulus that gives rise to the perceptual image is registered in the posterior parietal lobes, which keep track of where in space one is looking when a particular part of the field is encoded. Thus, one can store successive images (presumably in the temporal lobes) plus the loci of corresponding objects (in the parietal lobes), and the two sorts of information can be integrated downstream (as we discuss in chapter 5). Moreover, the parietal representation accounts for how we see three-dimensional images when two-dimensional depictive representations are used in vision; based on a variety of cues (such as retinal disparity, texture gradients, and occlusion) processed in earlier areas, the parietal lobes index each location in space with a distance. This representation allows one later to locate and fixate again on a particular object, if needed.

The Necessity of Interpretation

Is nothing in fact accomplished by “turning one topographically organized pattern into another”? To the contrary, there are good reasons for such mappings. For example, although the retina and area 17 have a similar spatial organization, the pattern of connections among neurons is radically different, because different information is processed in each structure. Indeed, area 17 not only has numerous short inhibitory connections, but also has

neurons that compare input from two eyes (which obviously cannot be accomplished in a single retina); such information is important for segregating figure from ground and organizing perceptual units—which are not the tasks of the retina. A series of topographically mapped areas is computationally useful because their spatial organization makes explicit and accessible information needed to accomplish the initial phases of visual processing, and each phase takes advantage of transformations accomplished in earlier phases.

The fact that visual processing is not accomplished in one fell swoop was stressed by Marr (1982); vision is accomplished in a succession of small steps. We now know that visual input is processed in a set of successive brain areas, each of which abstracts specific types of information from processes that operate earlier in the sequence (e.g., Miyashita & Hayashi, 2000). The initial visual areas in cortex are topographically organized, but progressively later ones are increasingly less well organized topographically, until the latest ones in the temporal lobe no longer preserve the organization of the retina, or do so very loosely (Malach, Levy, & Hasson, 2002). Although not all visual processing relies on depictive representations, the evidence strongly indicates that depictive representations are in fact used during the early phases of visual processing.

Dennett's (2002) quip that electricity generated from hydroelectric plants is not wetter than electricity generated by other means is cute, but misses the mark. The topographical organization of early visual areas is not irrelevant in the end, after input has been fully processed. Dennett's notion is a little like saying that gasoline would only be important for automobiles if they moved ahead on a path of flaming fluids! Gasoline plays a key role in the sequence of events that allows an automobile to move, even if in the end it does not directly turn the wheels. Moreover, many features of an automobile, such as room for a gas tank and placement of an exhaust pipe so as not to allow fumes into the passenger compartment, emerge from use of an internal combustion engine. Similarly, the depictive properties of topographically organized areas play crucial roles in the early phases of visual perception, allowing the process to begin by making the spatial structure of an object explicit and accessible (with respect to processes that will operate on this representation). Furthermore, there is much evidence that the visual system is wired to allow information that is only implicit in the higher areas to be made explicit and accessible by reconstructing the shape in earlier areas. If so, then the stored representations are formed in part to have the ability to recreate such earlier, depictive representations—and this capacity is one factor that determines how the stored representations are organized.

Conclusions

The virtues of turning to the brain should now be clear, and we press this approach forward in the following chapter. We showed in this chapter that the claim that visual mental images depict is consistent with key facts about the brain and the relevant behavioral findings. Moreover, neither the anatomical properties of topographically organized areas nor their role in imagery can be explained by appeal to tacit knowledge theories.

However, although many neuroimaging studies have documented that the topographically organized cortex is activated when people visualize, many other studies have failed to find such activation. The meta-analyses summarized here made sense of what at first glance might appear to be a chaotic set of empirical findings (reviewed in the appendix). When people visualize shapes with high resolution, depictive representations are in fact created. The theory summarized in the following chapter allows us to understand these findings, and also to understand why spatial imagery tasks do not activate topographically organized areas.

One strong prediction that emerges from our interpretation of the results is that if the topographically organized visual cortex represents the geometric properties of surfaces of shapes during imagery, then the precise pattern of activation should reflect the shape of the object in the same way it does during perception. Moreover, if a small part of topographically organized visual cortex is impaired, either from a very localized lesion (as might occur after a small tumor is removed) or from very focused TMS, only a small portion of an imaged object should be obscured—and the precise portion should depend on the size, orientation, and position of the imaged object. As the methodologies used in cognitive neuroscience continue to improve, it should be possible to test these strong predictions.

Notes

1. We must note that in the topographically organized areas there are some cases in which points close by in external space are not close by on the cortical surface. For example, if one takes two dots, one just above the horizontal meridian, one just below, their representation in area V2 will be far away because of how the upper and lower visual fields are represented in V2. Topographic areas typically represent only the upper or lower half of a single hemifield. As we would expect, perceptual effects arising from this neuroanatomical configuration can be measured psychophysically (Rubin, Nakayama, & Shapley, 1996). Nonetheless, the system interprets this pattern correctly. We again emphasize that the depictive properties of these areas

emerge from the way they are connected to areas downstream. The fact that these areas function to depict information within the context of the system is clear: the location of damage in topographically organized brain areas is directly related to the loci of blind spots in the visual field, and magnetic stimulation of different parts of such cortex produces phosphenes in the corresponding locations in space (see also chapter 1).

2. One might argue that the crucial variable is not the actual physical distance on cortex, but rather a kind of “functional distance” defined by connectivity among neurons in a brain area. Start by taking the extreme case, where physically distant neurons are connected directly to each other and thus function as if they were adjacent, whereas physically nearby neurons are connected indirectly to each other via intermediate neurons and thus function as if they were very far apart. If the neurons in an area were so arranged, would the area in fact depict? Now take a less extreme case, where physically distant neurons tend to have more intermediate neurons between them—and thus, the crucial variable is not actual distance on the cortex but rather the number of connections that intervene among neurons. Would such an area still depict?

We have two responses to this concern (which was expressed to us in conversation by the distinguished philosopher Ned Block): (1) In point of fact, most connections between neurons in topographically organized areas are short and inhibitory (at least in the monkey brain, about which we know considerably more than we do about the human brain). This is a good example of how neuroanatomy constrains theory. We cannot simply posit an arrangement of connections by fiat; there is a physical device, and there are known facts about how it is constructed (and how it operates). (2) The fact that topographically organized areas are physically depictive is irrelevant for present purposes. The neurons in these areas could be interconnected arbitrarily, but as long as *fixed* connections to areas farther downstream “unscramble” the activity in earlier areas appropriately, the earlier areas will functionally depict. We know that such an orderly mapping to later areas occurs because of the systematic relations between loci that are damaged in an area and loci in the visual field where blind spots appear.

3. Slotnick et al. also found that such imagery induced activation in the motion area MT. This finding is of interest because motion-related activity in this area has been directly linked to conscious awareness of a moving stimulus (Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996; Britten, Shadlen, Newsome, & Movshon, 1992; Shadlen & Newsome, 1996; Zeki & fytche, 1998).

4. Moreover, many other researchers have found activation in the topographically organized visual cortex when images are formed on the basis of information stored in LTM—specifically, Chen et al. (1998), Goldenberg and colleagues (Goldenberg, Podreka, Steiner, et al., 1989; Goldenberg, Steiner, Podreka, & Deecke, 1992; with high-imagery sentences); Handy et al. (2004); Klein, Paradis, et al. (2000); Kosslyn, Alpert, Thompson, Maljkovic, et al. (1993, Experiment 3); and Lambert, Sampaio, Scheiber, and Mauss (2002). On the other hand, some studies that required participants to form images based on recently encoded stimuli failed to find activation in topographically organized visual cortex (Ghaëm et al., 1997;

tanks. Although the matches that represent the tanks may not themselves be iconic with respect to the tanks, the spatial attributive structure into which they are fitted is certainly iconic. Referential applications are applied within a structure that is iconically related to the array of particulars that the referential applications function to pick out. For perceptual representation, iconicity is centrally a feature of the attributives—property, kind, and relational attributives.

Since being iconic is not in itself being representational, perceptual representation's being iconic depends on there being non-representational aspects of perceptual attribution. What are the non-representational features that make perception's iconic representation possible? The primary features are geometrical or topographical, temporal, and qualitative.

Iconic Aspects of Perceptual Spatial Representation

Spatial structure is by far the most thoroughly studied non-representational natural feature used in iconic perceptual representation. Many philosophers have ridiculed the idea of a picture in the mind. They base the ridicule partly on the correct point that perceptual state kinds do not literally have spatial extension. Modern perceptual theory does not attribute the literal spatiality of pictures to perception. Representational contents are abstract psychological kinds. They are not laid out in space. Yet perceptual theory does take visual and tactile perception to be spatially iconic.³⁰⁵ Aspects of matrix-like or picture-like format of perceptual content are attributed in models of perceptual representation. Such aspects can be mapped geometrically to spatial structures in the environment. The mapping preserves, in the format of perceptual representational content, analogs of environmental spatial structure. The pure geometry that applies to perceptual representational content derives from the geometry of physical space. Geometrical and topological structures are natural, non-representational aspects of perceptual representation. The structure of normal perceptual representational content is topological, even if not specifically spatial. The structure of both the spatial subject matter and the perceptual representational content can be described in pure geometry or topology. Thus an edge element in a perceptual content represents and can be mapped to an edge in the environment. Parts of the edge element represent parts of the edge in the environment. Directional—say, left-right—relations in the percept correspond to directional relations in the scene.

Such mapping occurs in conscious experience. One can block light from the lower half of one's eye. One thereby blocks representation of the lower part of the scene. Similarly, for the upper, the left, the right, and smaller parts of one's eye. There is a systematic correspondence between spatial aspects of one's receptive apparatus, spatial aspects of the perceived environment, and spatial aspects of one's perceptual representation of the

³⁰⁵ For useful discussions of why postulating iconic representations is not subject to the common ridicule of postulating literal pictures in the head, see S. Kosslyn, S. Pinker, G. Smith, and S. Schwartz, 'On the Demystification of Mental Imagery', *The Behavioral and Brain Sciences* 2 (1979), 535–581; R. Finke, S. Pinker, and M. Farah, 'Reinterpreting Visual Patterns in Mental Imagery', *Cognitive Science* 13 (1987), 73–74; M. Rescorla, 'Cognitive Maps and the Language of Thought', *British Journal of the Philosophy of Science* 60 (2009), 377–407.

environment. No similar mappings occur between words in a non-iconic sentence and spatial subject matters of the words.³⁰⁶

The intuitive point is elaborated in perceptual theory. The abstract topology of perceptual representational content has derivative mathematical association with literal topologies in the brain, and thence to spatial positions of areas of surfaces in the environment. For example, neural activity in the cortical area that connects to retinal receptors that receive light from the upper right corner of the retinal image correlates with aspects of conscious perception that represent the upper right corner of the visual perception of space (and represents entities in the upper right corner of the spatial layout of the scene). Even if consciousness and perceptual content are not literally laid out in space, there is a geometrical or topological structure-preserving mapping between aspects of space and aspects of conscious visual representation of space.

The iconic nature of visual spatial *representation* and *representational content* is grounded in the iconic nature of visual *sensory information registration* and in the spatial layout of aspects of the central brain that correspond to space-mapping areas in visual perception.

Four factors—visual representational content, visual non-representational information registration, cortical central brain areas that underlie spatial perception, and visual receptors—figure in the discussion. All ground either iconic information registration or iconic spatial representation. All are structurally related. The physical layout of the visual receptors renders the registration of information from the retinal image iconic. The iconic registration of the retinal image has mappings into the central brain. These registrational mappings are also iconic. Both mappings help make visual representation in perceptual states (and their representational contents) iconic. The mapping of geometric or topological structures common to physical space, visual receptors, visual information registration by banks of retinal and cortical neurons, and the format of visual representational content lies at the base of the iconic nature of visual spatial perception.

I start with the spatial structure of sensory receptors (at least in vision and touch). Such receptors form a 2-D array in physical space. In vision, each receptor is associated with a certain spatial part of the retinal image—the 2-D array of light that constitutes the proximal stimulus.³⁰⁷

This 2-D layout is preserved in neural layouts in the cortex, layouts specialized for vision. Areas of the primary visual cortex (V1) and other brain areas associated with early stages of visual processing are specialized to respond to corresponding areas of the retinal receptors and retinal image (the proximal stimulus).³⁰⁸ Thus, the *sensory information registration* of proximal stimuli in the early stages of vision has a literal spatial arrangement that corresponds, in vision, to the spatial arrangement of the retinal image. The spatial information registered by the sensors and by the further registrational states in the cortex that derive from them is laid out in spatial matrixes. The matrixes bear literal,

³⁰⁶ For an intuitive rendering of the idea that perceptual representation has a matrix-like format, see Peacocke, *A Study of Concepts*, chapter 3.

³⁰⁷ D. Hubel and T. Wiesel, 'Receptive Fields and Functional Architecture in Two Nonstriate Visual Areas (18 and 19) of the Cat', *Journal of Neurophysiology* 28 (1965), 229–289; S. Engel, G. Glover, and B. Wandell, 'Retinotopic Organization in Human Visual Cortex and the Spatial Precision of Functional MRI', *Cerebral Cortex* 7 (1997), 181–192; B. Wandell, S. Dumoulin, and A. Brewer, 'Visual Field Maps in Human Cortex', *Neuron* 56 (2007), 366–383.

³⁰⁸ Frisby and Stone, *Seeing: The Computational Approach to Biological Vision*, chapters 3, 8.

systematic relations to positions in the retinal image—the area of light hitting the retinae. These non-perceptual registrations are unquestionably iconic.

The notion of a *visual field map* is a technical notion in vision science. It applies to layouts of neural receptors in retinal, sub-cortical, and cortical areas of visual processing. A visual field map must meet three conditions.

First, a region of neural receptors should register, or ground representation of, a substantial portion of the retinal image or an environmental layout. The qualification ‘substantial portion’ is entered to allow for the fact that visual field maps do not register each part of whatever they register with equal surface area. Some visual field maps in V1, the earliest cortical area specialized for visual processing, are oriented primarily to foveated parts of the retinal image.

Second, a visual field map must register a substantial part of what it registers in a way that preserves spatial structure. The map should be roughly continuous. Distance from the center of the map should vary roughly continuously with distance (the eccentricity) from the center of the registered entity—for example, the retinal image. Directions from the center of the map (the angle dimensions) should vary roughly continuously from directions from the center of the registered entity. ‘Roughly’ allows for discontinuities. For example, it allows for discontinuities between the two hemispheres of the human brain that divide registration of the left and right sides of the retinal image. Many discontinuities bear on the border of a map. Some occur within maps.

Third, the basic features of a field map must be roughly consistent across individual subjects—though maps often vary across individuals in size and precise anatomical location.³⁰⁹

There is massive evidence for mapping from the spatial layout of registrational sensors that underlie the early stages of visual processing to areas of the retinal image. The retinal receptors, layouts in the lateral geniculate nucleus (LGN), and many layouts in all the cortical areas (V1–V3) that support early visual processing, all constitute visual field maps.³¹⁰ Since there are hundreds of V1 neurons for every ganglion cell or LGN relay neuron, there is room in V1 for many specialized maps, each mapping the retinal image, each functioning to carry its own task. Numerous species have multiple retinotopic maps in each early processing area, such as V1.³¹¹

³⁰⁹ The conditions are laid out in Wandell, Dumoulin, and Brewer, ‘Visual Field Maps in Human Cortex’, 370. I have generalized their formulations slightly.

³¹⁰ Engel, Glover, and Wandell, ‘Retinotopic Organization in Human Visual Cortex and the Spatial Precision of Functional MRI’; R. Dougherty, V. Koch, A. Brewer, B. Fischer, J. Modersitzki, and B. Wandell, ‘Visual Field Representations and Locations of Visual Areas V1/2/3 in Human Visual Cortex’, *Journal of Vision* 3 (2003), 586–598; Wandell, Dumoulin, and Brewer, ‘Visual Field Maps in Human Cortex’; M. Silver and S. Kastner, ‘Topographic Maps in Human Frontal and Parietal Cortex’, *Trends in Cognitive Sciences* 13 (2009), 488–495; Wandell and Winawer, ‘Imaging Retinotopic Maps in the Human Brain’.

³¹¹ A. Cowey, ‘Projection of the Retina on to Striate and Prestriate Cortex in the Squirrel Monkey, *Saimiri sciureus*’, *Journal of Neurophysiology* 17 (1964), 366–393; A. Cowey, ‘Aspects of Cortical Organization Related to Selective Attention and Selective Impairments of Visual Perception: A Tutorial Review’, in M. Posner and O. Marin eds., *Attention and Performance XI* (Hillsdale, New Jersey: Erlbaum, 1985); R. Gattass, A. Sousa, and M. Rosa, ‘Visual Topography of V1 in the Cebus Monkey’, *Journal of Comparative Neurology* 259 (1987), 529–548; R. Gattass, S. Nascimento-Silva, J. Soares, B. Lima, A. Jansen, A. Diogo, M. Farias, M. Botelho, O. Mariani, J. Azzi, and M. Fiorani, ‘Cortical Visual Areas in Monkeys: Location, Topography, Connections, Columns, Plasticity and Cortical Dynamics’, *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 360 (2005), 709–731; G. Wang and A. Burkhalter, ‘Area Map of Mouse Visual Cortex’, *Journal of Comparative Neurology* 502 (2007), 339–357; W. Geisler, ‘Visual Perception and the Statistical Properties of Natural Scenes’, *American Review of Psychology* 59 (2008), 167–192, 183; B. Wang, H. Yamamoto, J. Wu, and Y. Ejima, ‘Visual Field (cont.)

Until the last 30 years, it was widely thought that visual field maps were located entirely in early visual areas. This view depended partly on inadequate technology for testing the later visual areas, for example, from V4 onward. It also depended on reasoning by psychologists that since the higher areas supported spatial constancies, visual representation was “position invariant”. So, went the reasoning, one should not expect field maps in those areas.

The view also probably depended on not taking into account discoveries in the late 1970s of neural maps in the barn owl that correlate with spatial positions in *distal* environmental space, and that also correlate with the owl’s spatial perceptual competencies.³¹²

Better technology found many maps in the higher visual areas in humans. Many of these maps play a role in mapping the environmental positions of perceptually represented entities, as contrasted with maps that merely register positions in the retinal image.³¹³

The reasoning by psychologists, cited four paragraphs back, was unsound. A given area of a brain map can be active in response to a position in distal environmental space, or can otherwise aid in representation of a position in space. It can support a representation of something in that position via an egocentrically (or even allocentrically) anchored map. The discovery of maps that figure in *perceptually* mapping the distal environment is perhaps most important development in neuroscience’s visual-field mapping in the last 30 or so years.³¹⁴

Sensitivity to *retinal* position is not lost among maps that figure in representing spatial structures in the distal environment. That is, even where a map figures in representing the distal position of a perceived entity, there is evidence that retinal position of the effect of the perceived entity is kept track of, either in that map or in some other map with which that map coordinates.³¹⁵

Maps of the Human Visual Cortex for Central and Peripheral Vision, *Neuroscience and Biomedical Engineering* 1 (2013), 102–110; Wandell and Winawer, ‘Imaging Retinotopic Maps in the Human Brain’.

³¹² E. Knudsen, and M. Konishi, ‘A Neural Map of Auditory Space in the Owl’, *Science* 200 (1978), 795–797; E. Knudsen, ‘Auditory and Visual Maps of Space in the Optic Tectum of the Owl’, *Journal of Neuroscience* 2 (1982), 1177–1194. The owl has auditory and spatial maps of positions in distal environmental space, each of which feeds a further bimodal map.

³¹³ C. Galletti, P. Battaglini, and P. Fattori, ‘Parietal Neurons Recording Spatial Locations in Craniotopic Coordinates’, *Experimental Brain Research* 96 (1993), 221–229, reported an early discovery of neurons that correlate with real positions in space, determined in perceptual spatial constancies. See also N. Logothetis, J. Pauls, and T. Poggio, ‘Shape Representation in the Inferior Temporal Cortex of Monkeys’, *Current Biology* 5 (1995), 552–563; DeAngelis and Newsome, ‘Organization of Disparity-Selective Neurons in Macaque Area MT’; N. Swindale, D. Shoham, A. Grinvald, T. Bonhoeffer, and M. Hübener, ‘Visual Cortex Maps are Optimized for Uniform Coverage’, *Nature Neuroscience* 3 (2000), 750–752; D. Chklovskii and A. Koulakov, ‘Maps in the Brain: What Can We Learn from Them?’, *Annual Review of Neuroscience* 27 (2004), 369–392; S. Murray, H. Boyaci, and D. Kersten, ‘The Representation of Perceived Angular Size in Human Primary Visual Cortex’, *Nature Neuroscience* 9 (2006), 429–434; S. MacEvoy and R. Epstein, ‘Position Selectivity in Scene- and Object-Response Occipitotemporal Regions’, *Journal of Neurophysiology* 98 (2007), 2089–2098; A. McKyton and E. Zohary, ‘Beyond Retinotopic Mapping: The Spatial Representation of Objects in the Human Lateral Occipital Complex’, *Cerebral Cortex* 17 (2007), 1164–1172; Wandell, Dumoulin, and Brewer, ‘Visual Field Maps in Human Cortex’; Silver and Kastner, ‘Topographic Maps in Human Frontal and Parietal Cortex’; T. Carlson, H. Hogendoorn, H. Fonteijn, and F. Verstraten, ‘Spatial Coding and Invariance in Object-Selective Cortex’, *Cortex* 47 (2011), 14–22; Carlson, Hogendoorn, Kanai, Mesik, and Turret, ‘High Temporal Resolution Decoding of Object Position and Category’; B. Barton and A. Brewer, ‘Visual Field Map Clusters in High-Order Visual Processing: Organization of V3A/V3B and a New Cloverleaf Cluster in the Posterior Superior Temporal Sulcus’, *Frontiers in Integrative Neuroscience* 11 (2017), <https://doi.org/10.3389/fnint.2017.00004>.

³¹⁴ Wandell and Winawer, ‘Imaging Retinotopic Maps in the Human Brain’.

³¹⁵ Z. Kourtzi and N. Kanwisher, ‘Representation of Perceived Object Shape by the Human Lateral Occipital Complex’, *Science* 293 (2001), 1506–1509; J. DiCarlo and D. Cox, ‘Untangling Invariant Object Recognition’, *Trends in Cognitive Sciences* 11 (2007), 333–341; MacEvoy and Epstein, ‘Position Selectivity in Scene- and Object-Response Occipitotemporal Regions’; Carlson, Hogendoorn, Fonteijn, and Verstraten, ‘Spatial Coding and

Visual perception is always from a perspective that is marked by the way perceived objects affect our sensors. Retinal position almost inevitably affects mode of presentation.

These points support the double-indexing method that I employed in developing the form of perceptual representational contents. (See Chapter 7, the section *THE FORM OF PERCEPTUAL ATTRIBUTIVES IN LINKAGES*.) An attributive is constitutively as of an attribute. Different attributives can indicate, and attribute, the same attribute. The way in which the attribute is represented, which depends largely on the retinal source of the representation, is constitutive of a representation. Different attributives that indicate the same entity do so through a perspective that is marked by features—prominently spatial features—of the entity's effect on the retina.

Spatial and feature-based attention, both their initiations or commands and their effects—are grounded in visual field maps. Not surprisingly, the mechanisms of visual attention coordinate with the visual field maps that support perceptual processing.³¹⁶

So far, I have discussed visual maps at the neural level. What bases are there for taking neural maps, which are not themselves representational, to bear on the format of visual representational contents—that is, with kinds of perceptual representational competencies? Of course, it is immensely plausible, antecedent to specific evidence, that the functions and causal relations involved in visual perception capitalize upon, or show the marks of, the underlying neural maps. Specific evidence backs this natural expectation.

There are systematic correlations between activity in neural areas and perceptual functions. Such correlations are specific enough to connect features of the visual field maps with corresponding features of perceptual representation.

Stimulation that yields directional change in a neural visual map, or that affects a position in a neural map, corresponds with perceptual representations (and behavioral responses) in the corresponding direction or position.³¹⁷ (I assume here, what I have repeatedly emphasized in earlier chapters, that direction is not *represented* perceptually unless it is combined with ordinal depth or with distance in a 3-D mapping.) Similarly, perception of motion and experience of motion aftereffects correlate with direction of changes in underlying firings across neural visual field maps.³¹⁸ In binocular rivalry, changes in perceptual dominance from one eye to the other commonly correspond to directional waves of firing across a V1 map.³¹⁹

Crowding is a degradation of perceptual representation by nearby distal stimuli. Crowding appears to be determined by spacing between stimuli on visual field maps. So

Invariance in Object-Selective Cortex'; Cichy, Chen, and Haynes, 'Encoding the Identity and Location of Objects in Human LOC.'

³¹⁶ S. Astafiev, G. Shulman, C. Stanley, A. Snyder, D. Van Essen, and M. Corbetta, 'Functional Organization of Human Intraparietal and Frontal Cortex for Attending, Looking, and Pointing', *Journal of Neuroscience* 23 (2003), 4689–4699; J. Maunsell and S. Treue, 'Feature-Based Attention in Visual Cortex', *Trends in Neurosciences* 29 (2006), 317–322.

³¹⁷ T. Albright, 'Direction and Orientation Selectivity of Neurons in Visual Area MT of the Macaque', *Journal of Neurophysiology* 52 (1984), 1106–1130; C. Salzman, K. Britten, and W. Newsome, 'Cortical Microstimulation Influences Perceptual Judgements of Motion Direction', *Nature* 346 (1990), 174–177; C. Salzman, C. Murasugi, K. Britten, and W. Newsome, 'Microstimulation in Visual Area MT: Effects on Direction Discrimination Performance', *Journal of Neuroscience* 12 (1992), 2331–2355.

³¹⁸ A. Ezzati, A. Golzar, and A. Afraz, 'Topography of the Motion After-Effect With and Without Eye Movements', *Journal of Vision* 8 (2008), 1–16.

³¹⁹ S. Lee, R. Blake, and D. Heeger, 'Traveling Waves of Activity in Primary Visual Cortex During Binocular Rivalry', *Nature Neuroscience* 8 (2005), 22–23.

there is a relatively direct correspondence between proximity in the neural maps and perceptual spatial representation of close-by stimuli, retinal and distal.³²⁰

Shifts of attention correlate with such shifts in receptive fields of neurons in visual field maps. These shifts have been shown to have systematic effects on perceptual representation of size, such as an increase of the perceived size of attended stimuli. The shifts in size are of entities in positions that relevant neurons respond to. Such systematic relations between functions of neurons that constitute spatially specific components in map-like neural registration of spatial properties (distal as well as proximal), on one hand, and perceptual representation of spatial properties, on the other, indicates that the format and representational capacities in the visual representation of spatial structures parallel spatial structures in the visual field maps.³²¹

Damage to areas of a neural map produces deficits in spatial representation that correlate with the damaged areas. Given that the brain can produce remappings after such damage and given that there are often many “back up” maps that can compensate for local damage to a map, there is no simple inference from an area of neural map damage to a spatial area for which visual perception shows a deficit. However, allowing for such complexities, there are mappings between areas in a neural map where there is damage and corresponding spatial perceptual deficits. The mappings indicate a topographical format for visual perception.³²²

Psychologists working in these areas assume that these mappings between visual field maps in the brain and perceptual spatial representation are vastly more extensive than the cases that are now known, of which I have cited just a small sample. The fact that there are so many topographically organized neural maps and the limitations of current techniques in identifying and measuring them make investigating these relationships complex and slow-moving.³²³

The role of spatial structure in determining iconic aspects of visual perception goes much further than a natural correspondence between the geometrical or topological structure of perceptual content and the layout of perceptual objects in space. Accounts of perceptual representation of specific spatial attributes in 3-D space use geometrical characteristics of edges and lines that are imprinted, 2-D, in retinal registration.³²⁴ The accounts are explicitly in geometrical or topological form. Patterns of local lengths and orientations of registration of edges in the (literally) spatial layout of the retinal image are used by the visual system in making estimates of larger patterns in physical space. The repeatable, ability-general representational content is taken explicitly, in the science, to have a geometrical or topological structure.

More specifically, neural receptors are sensitive to the directional orientation of bars, edges, and other patterns in the retinal image, at various spatial scales. Recall from

³²⁰ D. Pelli, ‘Crowding: A Cortical Constraint on Object Recognition’, *Current Opinion in Neurobiology* 18 (2008), 445–451.

³²¹ K. Anton-Erxleben, C. Henrich, and S. Trueue, ‘Attention Changes Perceived Size of Moving Visual Patterns’, *Journal of Vision* 7 (2007), 1–9. I discuss philosophical implications of attention’s distorting effects in Chapter 15, the section ATTENTION AND ACCURACY.

³²² J. Kaas, ‘Topographic Maps are Fundamental to Sensory Processing’, *Brain Research Bulletin* 44 (1997), 107–112.

³²³ Wandell and Winawer, ‘Imaging Retinotopic Maps in the Human Brain’, 725.

³²⁴ One can get a sense of the power of conversion from patterns among local surface shapes to representation of 3-D shapes from the op-art pictures of Bridget Riley. See image search results for “Bridget Riley”.

Chapter 3, the section *RETINAL IMAGE CONTOUR REGISTRATION AND SURFACE CONTOUR PERCEPTION*, that an edge in the registration of the retinal image is marked by a line of neurons that fire at a much higher intensity than a line of adjacent neurons. Among central cues to representation of spatial attributes of environmental edges, shadows, and textural features of surfaces—attributes such as slant and tilt—are patterns among lengths and orientations of edges in registrations of the 2-D retinal image.

For example, a pattern of lines in the initial registration of the retinal image that grows progressively smaller at the top of a pattern, in comparison to the bottom, might be a cue to foreshortening that would indicate a backward slant of a surface. Or a pattern of lines that grows progressively shorter in a pattern from left to right might be a cue to surface tilt.

The patterns are present in a perceptual representation. At *perceptual* stages of processing they are given content that represents features of entities in 3-D physical space. For example, the lines might represent textural features on a surface with an orientation. Or they may represent the contours of the facing part of a body. Other cues are relevant—shadows, consistency with occurring within surface boundaries, binocular disparity. Patterns of geometrical characteristics in registrations that come from pre-representational stages of vision (and that can be directly correlated with spatial patterns among neuronal firings) are a *central* cue to computations of perceptual representational content that attributes 3-D volume shapes in the environment.

Roughly, a pattern of lines in a plane takes on 3-D meaning, in something like the way in which the lines in a drawing are used to represent shapes in 3-D space. Again, science takes the percept—hence its representational content—to have a geometrical structure that has systematic, natural, iconic, mappings to structures in physical space.

The science's theory and practice take the structure and these aspects of the content of visual representation to be systematically analogous to structure and content in drawings or other pictures.³²⁵ Again, the representation's being iconic does not hinge on its being in space. It is enough that it have a geometrical or topological structure that has a natural mapping to spatial structures. The idea that visual spatial representation occurs in a map-like format that bears systematic relations to underlying neural visual field maps is well supported. This is the format of spatial and temporal frameworks described in Chapter 7, the section *MINIMALISM AND ICONIC REPRESENTATION IN A SPATIAL COORDINATE SYSTEM* and in Chapter 8, the section *EGOCENTRIC SPATIAL INDEXES IN PERCEPTION*. Work in the psychophysics of vision repeatedly relies on the idea.

So the iconic nature of visual spatial representational content is grounded in the layout of visual receptors and in the layout of the environment. It is also grounded in the iconic nature of visual information registration in the central brain that corresponds to space-mapping areas in visual perception. All five factors—visual spatial representation, visual information registration in brain areas that underlie spatial perception, information

³²⁵ E. Mingolla and J. Todd, 'Perception of Solid Shape from Shading', *Biological Cybernetics* 53 (1986), 137–151; J. Koenderink, A. Van Doorn, and A. Kappers, 'Surface Perception in Pictures', *Perception and Psychophysics* 52 (1992), 487–496; Knill, 'Surface Orientation from Texture: Ideal Observers, Generic Observers, and the Information Content of Texture Cues'; Murray, Boyaci, and Kersten, 'The Representation of Perceived Angular Size in Human Primary Visual Cortex'; Silver and Kastner, 'Topographic Maps in Human Frontal and Parietal Cortex'; R. Fleming, D. Holtzman-Rice, and H. Bülthoff, 'Estimation of 3D Shape from Image Orientations', *Proceedings of the National Academy of Sciences USA* 108 (2011), 20438–20443.

registration in pre-perceptual areas (including visual receptors), the retinal image, and spatial aspects of the distal environment—are structurally related. The physical layout of the visual receptors renders information registration of the retinal image iconic. The registration of the retinal image has mappings into registration states in the central brain. Both bear mapping relations to visual perceptual representation. These natural mappings of geometric or topological structures among levels of information registration and spatial perception are each mappable to spatial structures in the retinal image and the distal environment. These mappings are natural, not representational.

Showing that picture-like or map-like structures correctly model visual states hinges on these structures' providing insight into perceptual competencies. Causal relations among perceptual states and between such states and behavior have been naturally and fruitfully explained by taking these states to have geometrical or topological structures that map to geometrical or topological structures in the environment that are literally spatial. Such models are fruitful because they suggest new hypotheses that a list of the relevant causal relations would never suggest.

For example, suppose that visual perception were represented simply by a set of abstract propositions, with the form of the meanings of ordinary sentences. Then one would expect that size, shape, location, orientation would normally be factored into separate propositions one for each relevant attribute. In iconic, fully formed perceptual representations, such attributes are represented together as a package, although some may be processed in different sequences. This feature is not a consequence of being iconic. It is a prominent feature of iconic *perceptual* representation. It is widely assumed from many sources of evidence that spatial features of a scene are represented together, not separately.³²⁶ Not only perception, but imagery, various types of short-term perceptual memory, and long-term perceptual memory represent in iconic representational structures. All these capacities derive their iconic structure from perception.

Evidence for this fact derives from the way different aspects of a perceptual or perception-derived representation are responded to. Many experiments show the normal interdependence of representation of size, orientation, shape, location. For example, subjects are asked to say whether an arrow points to an imagined dot. The time required to answer is proportional to how far the arrow is from the dot.³²⁷ A view that imagery is stored as non-iconic representations does not naturally predict interlocking representations of different attributes. It must add postulates that purport to express relevant background principles to explain the results. The principles are not motivated by any fruitful

³²⁶ For an almost arbitrary example, see A. Oliva, S. Park, and T. Konkle, 'Representing, Perceiving, and Remembering the Shape of Visual Space', in L. Harris and M. Jenkin eds., *Vision in 3D Environments* (Cambridge: Cambridge University Press, 2011).

³²⁷ S. Pinker, P. Choate, and R. Finke, 'Mental Extrapolation in Patterns Reconstructed from Memory', *Memory and Cognition* 12 (1984), 207–218. The famous rotation experiments by Shepard suggest that orientation and shape are interlocked in mental images as well as in perception. Shepard reported that subjects could not determine whether two depicted 3-D objects had the same shape until they compared one mentally rotated one object into the orientation of the other. The rotations take proportionally more time the greater the difference in their angles of orientation. R. Shepard and J. Metzler, 'Mental Rotation of Three-Dimensional Objects', *Science* 171 (1971), 701–703. There are methodological questions about these experiments. Other explanations of the results are possible. See Chapter 18, the section PERCEPTUAL IMAGINING for more discussion. In any case, the close connection between representation of orientation and representation 3-D volume shape is present in imagery as well as perception.

general view of imagery. The evidence is naturally explained if perceptual imagery is taken to have the geometrical or topological iconic format of perception.³²⁸

Similar evidence supports the widespread view that some aspects of the iconic form of visual perceptual representation are retained in visual working memory and visual anticipation.³²⁹ There is also evidence that iconic form remains in visual perceptual long-term memory.³³⁰ Retention often comes with reduced richness, in something like the way sketches are less rich than realist paintings. In long-term memory, especially in beings with linguistic capacities, there are surely both iconic and non-iconic types of representation.

The fact that perception, perceptual memory, perceptual anticipation, and imagery all share the format and much of the content of visual perception helps suggest, what I will argue in Part IV, that they operate at the same representational level in a single perceptual-level system.

Temporal, Qualitative, and Packaging Iconic Aspects of Visual Perception

Spatial representation is most refined and complex in vision and touch. Temporal representation is most refined in vision and hearing. All three of these systems, and the proprioceptive system, are iconic in temporal representation as well as spatial representation.

The iconic nature of temporal perceptual representation is even more obvious than that of spatial representation. Perceptual states are extended in time. Attributional applications (hence contents) can be temporally extended. Natural correlations between temporal relations in vision and in what is perceived render temporal perceptual attributions in vision iconic.

For example, a temporally extended perception may have a sequence of referential applications that apply attributives to different events that occur sequentially. The temporal order of the referential applications matches the temporal sequence of events. Temporal order is a natural property. Mapping between perception and represented events

³²⁸ See also Shepard and Cooper, *Mental Images and Their Transformations* and Kosslyn, *Image and Mind*. The view that imagery is iconic and non-propositional is disputed in Z. Pylyshyn, 'Return of the Mental Image: Are There Pictures in the Brain?', *Trends in Cognitive Sciences* 7 (2003), 113–118; Pylyshyn, *Seeing and Visualizing: It's Not What You Think*. I believe that Pylyshyn's criticism of some of the proposed neural evidence for an iconic view of mental imagery is correct. However, I think that the most natural explanation of psychophysical evidence is that the formats or structures of the representational content of many instances of perceptual imagery, produced either in short-term memory or endogenously produced on demand, share iconic structure with perceptions. Such imagery is typically less rich and differs from perceptual representation in other ways. I do not go into the issues in depth. Even Pylyshyn does not dispute the standard view that perceptual representational content itself is iconic. I think it fair to say that the mainstream in the science of visual imagery and of relations between perception and imagery has not followed Pylyshyn. Further evidence for relations between visual perceptual imagery and visual perception: L. Cooper and R. Shepard, 'Chronometric Studies of the Rotation of Mental Images', in W. Chase ed., *Visual Information Processing* (New York: Academic Press, 1973); Shepard and Cooper, *Mental Images and Their Transformations*; M. Farah, 'Psychophysical Evidence for a Shared Representational Medium for Visual Images and Percepts', *Journal of Experimental Psychology: General* 114 (1985), 91–103.

³²⁹ For neural and psychophysical evidence that supports taking visual working memory as iconic, see M. Esposito, 'From Cognitive to Neural Models of Working Memory', *Philosophical Transactions of the Royal Society B* 362 (2007), 771–782; N. Rose, F. Craik, and B. Buchsbaum, 'Levels of Processing in Working Memory: Differential Involvement of Fronto-Temporal Networks', *Journal of Cognitive Neuroscience* 27 (2015), 522–532.

³³⁰ M. Schurgin, C. Cunningham, H. Egeth, and T. Brady, 'Visual Long-Term Memory Can Replace Active Maintenance in Visual Working Memory', *BioRxiv* (2018), <https://doi.org/10.1101/381848>.