

## 4 What Do Brain Areas Do?

In the past chapters, we encountered several brain areas and some of their functions. But what exactly is an area? Are brain areas more or less like discrete computational units that have well-defined functions? Because this question is so central to our understanding of the brain, we need to slow down here and discuss it in greater depth. We will see that the idea of “one area, one function” largely hinges on assuming a brain organization that is *modular* in essence. In contrast, if a given brain area is always involved in multiple functions, as advocated throughout the book, how should we revise our thinking?

### Tan, Tan

When Tan was admitted to the hospital at the age of 21, he had lost the use of speech for some time.<sup>1</sup> He could no longer pronounce more than a single syllable. Whenever a question was asked, he would always reply “tan, tan,” accompanied by varied expressive gestures. In fact, throughout the hospital, he was known only by his nickname, Tan. Despite his impediment, at the time of admission, he was perfectly able-bodied and intelligent and appeared to comprehend almost everything that was said to him. (In Brazilian Portuguese, “tan-tan” is colloquially, and pejoratively, used to denote someone who is “crazy.” I wonder if the origin has something to do with Tan’s predicament.)

In April 12, 1861, about 10 years after his initial admission, and rapidly deteriorating in health, Tan was seen by Paul Broca, a surgeon with an unusual background—he was one of the founders of the field of anthropology in France. Five days later, Tan would die of a severe case of gangrene. The brain was removed and preserved in a fixation fluid that made the tissue harden with time. Based on his examination, Broca concluded that Tan’s speech deficit was due to a lesion of the left frontal lobe. Broca

concluded his short case report published in the *Bulletin de la Société Anthropologique*, a mere page and a half in length, with the following momentous conclusion: “All this permits, however, the belief that, in the present case, the lesion of the frontal lobe was the cause of the loss of speech.” More than 150 years later, Broca’s report is the most important paper in the history of brain function localization.

### Discovering the Function of Brain Areas

Historically, lesions have played a major role in trying to infer the function of brain subparts. Two types of lesion have been considered: in humans, naturally occurring damage from tumors and vascular accidents; in animals, more precisely delineated lesions produced surgically. Broca’s paper, capitalizing on the first type of injury, catapulted forward the idea that a mental function can indeed be *localized*. At the time, very little was known about how the convoluted mass of gray and white matter inside the head supports mental faculties. Consider that Broca’s observations took place not long after the heyday of the much-maligned phrenology movement espoused by Franz Gall and his disciples, which was particularly influential between 1810 and 1840. Phrenologists would observe and feel the skull of individuals to determine their psychological propensities such as “philoprogenitiveness” (that is, the love of offspring or children in general), which was located centrally at the back of the head (more or less where we now know the visual cortex to be!).

An early series of lesion studies in animals was performed by Eduard Hitzig and Gustav Fritsch on dogs (published in 1870).<sup>2</sup> Hitzig was a psychiatrist interested in the potential applications of weak electrical currents to ameliorate certain medical conditions. By the mid-1860s, he had developed an apparatus to deliver electrical stimulation to human patients and observed that current applied to the back of a patient’s head reliably elicited eye movements, prompting him to investigate the use of the technique (also called galvanization) further. Hitzig thus invited Fritsch, an anatomist, to join him in studies to be conducted in dogs. Their most famous experiment was performed not in a well-equipped university laboratory but on a dressing table in a bedroom of Hitzig’s house in Berlin. Initially, they electrically stimulated the canine cortex with weak electrical currents. By systematically varying the site that was excited, they uncovered locations that elicited

muscular responses of the forepaw, hindpaw, face, and neck (in all cases, on the side of the body opposite to the stimulation, also called the contralateral side). With a scalpel, they then removed the area that led to, say, forepaw movement upon galvanization. Although this did not abolish all movement from the contralateral paw, movement was impaired, and abnormal postures were observed. Notably, sensation appeared to be normal, as the animals' responses to stimuli were unaltered. That is to say, the observed deficit caused by the lesion was relatively *selective* for motor production (and not sensory perception), and it was even linked to a specific body part.

The importance of the study by Hitzig and Fritsch, like the observations by Broca, cannot be overestimated. This was as much due to their results as to their *conclusions* based on combining electrical stimulation and lesions. As immodestly stated by them: "Some psychological functions, and perhaps all of them, in order to enter matter or originate from it, need circumscribed centers of the cortex." That is to say, according to them, the cortex contained processing *centers*. Hitzig and Fritsch therefore suggested that it would be worthwhile for researchers to search for areas concerned with sensation and even regions involved with intelligence. The time was ripe to explore the locations where mental functions *reside*.

### Dissociating Mental Functions

It is definitely possible to study the brain at its most elementary sense: mechanisms of neuronal spike generation and signal propagation along axons; molecular mechanisms along the gap, or synapse, between two neurons; and so on. But, often one studies brain mechanisms, even the most basic ones, to understand the neural basis of mental functions—seeing a sunset, hearing a screech, speaking a sentence, remembering a childhood memory, feeling uncertain about the future.

A chief goal in the sciences of the mind and brain is to explicitly unravel the *functional architecture* of the mind: to identify and characterize the *mental* processes underlying behavior. But mental processes are not directly observable. Rather, their existence must be inferred from external manifestations—what we call overt behaviors. Insight into mental functions can be gained based on the by-products of brain damage on carefully chosen tasks—much like the approach of Broca and of Hitzig and Fritsch. In broad terms, the existence and general contours of mental processes are

then inferred from the manner in which task performance changes from manipulation to manipulation, involving different levels of an experimental variable and different forms of brain damage. In trying to delineate mental processes, lesions are extremely valuable even to cognitive scientists not inherently interested in the brain. The reason is that disturbances in behavioral performance when the brain is damaged can inform us about the organization of the *mind*. For example, to what extent is the processing of verbs and nouns separate? If one uncovered damage that affects the processing of words used as verbs (say, “he judges”) more than of the same words used as nouns (say, “the judges”), this would be extremely valuable in outlining the organization of the mental processes in question—what was called the functional architecture above.

To infer the existence of *separate* mental processes, researchers rely on the logic of *dissociations*.<sup>3</sup> Consider, first, a *single* dissociation. Let *A* and *B* be two tasks (say, one involving verbs, another involving nouns) and let *m* be a “manipulation.” A single dissociation is observed if *m* affects performance on *A* but not on *B*. The “manipulation” could correspond to a lesion of a region, and a dissociation would be established if the damage impaired performance of *A* but not *B*. In all, a dissociation invites the inference that there is an underlying mental function required by *A* but not by *B*.

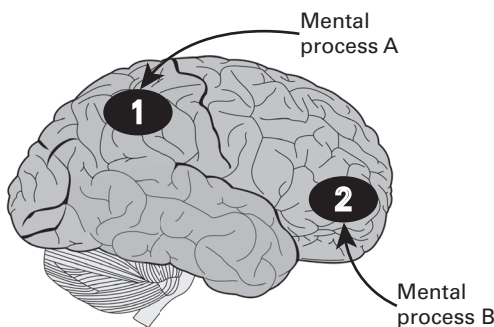
The logic of dissociation is central to neuroscience and has long been used to *localize* mental functions. The reasoning is analogous to what one would adopt to reverse-engineer a human-made device—say, remove the pistons in a car to try to “discover” that they are a key element in the combustion process that powers standard automobiles. But single dissociations are inferentially weak, and although they were frequently employed in research in the first decades of the twentieth century (and in many ways after that), in the 1950s investigators started to question the single dissociation’s application. For example, in some cases it may well be that *general* deficits following a lesion could explain the pattern of results; perhaps the lesion impairs most tasks that are difficult, and task *A* happens to be harder than task *B*.

Evidently, better experimental design and methodology, with careful choice of tasks *A* and *B* (and the dimensions along which *A* and *B* are matched), ameliorate the problems with guessing function. This is simply good experimental design, which is the cornerstone of solid experimental science. After all, if changes are observed in one experimental condition, the question is always, “Relative to what?” Control conditions are fundamental

in drawing reasonable inferences. Nevertheless, in trying to establish the anatomical underpinnings of mental functions, the single dissociation strategy is simply too weak. The inconclusiveness of the methodology motivates the *double* dissociation logic (figure 4.1). A single dissociation is observed if region 1 affects performance on task *A* but not on task *B*. A double dissociation is observed if, in addition, region 2 affects performance on *B* but not on *A*. Both single and double dissociations indicate that there is an underlying mental function required by *A* but not by *B*. In addition, a double dissociation invites the converse inference—namely, that there is an underlying mental function required by task *B* but not by task *A*. In addition, it is surmised that brain regions 1 and 2 carry out functions *that* are relatively *isolable* from each other. The power of the double dissociation logic lies in its specificity: lesions of regions 1 and 2 do not simply cause a series of impairments; instead, they impair *circumscribed* mental functions.

### What Do Double Dissociations Tell Us?

In humans, before the advent of modern neuroimaging techniques such as functional magnetic resonance imaging (MRI), researchers tried to uncover function by studying impaired and preserved abilities in brain-damaged patients. And the most powerful weapon in the neuropsychologist's armamentarium was the double dissociation approach just described.



**Figure 4.1**

The double dissociation logic. If mental process *A* is affected by a lesion to region 1 but not region 2, and vice versa for mental process *B*, we say that they are doubly dissociated. To many neuroscientists (but not all!) this pattern suggests that areas *A* and *B* are functionally specialized.

Here's a specific example. Although neurologically intact persons have no difficulty pronouncing written words regardless of whether they have a concrete meaning (say, "table") or an abstract meaning (say, "truth"), after a severe left-hemisphere stroke, patient PW correctly pronounced 67 percent of concrete words but only 13 percent of abstract words.<sup>4</sup>

The finding that concrete and abstract words were differentially susceptible to damage suggested that they are represented separately in neural tissue. An alternative view, however, is that PW's brain damage affected concrete and abstract representations equally, but that abstract words were more impaired because they are inherently more difficult to pronounce. On this latter account, one would not expect to see the opposite relationship: better reading of abstract than concrete words following brain damage. This is exactly what the neuropsychologist Elizabeth Warrington observed in patient CAV who had a left-hemisphere tumor: He read correctly 36 percent of concrete words but 55 percent of abstract words. Together, patients PW and CAV exemplify a double dissociation of concrete and abstract word reading. (Although both lesions were in the brain's left hemisphere, they were assumed to compromise different parts of the cortex. Moreover, the behavioral impairments of PW and CAV were observed in other patients, too.)

Double dissociations among brain-damaged patients, as well as animals with focal lesions, have been identified for many pairs of tasks, spanning perception, action, emotion, and motivation. To exemplify the typical interpretation of this form of relationship, consider the conclusion by Warrington, herself a hugely influential scientist, that "the only plausible interpretation of a double dissociation between abstract-word deficit and concrete-word deficit . . . is that the functional and structural organization of semantic representations of words is categorical" (Warrington 1981). That is to say, the semantics of concrete words and those of abstract words must be implemented separately—they rely on different functions *and* are carried out in separate parts of the brain. The reasoning supporting this interpretation dovetails nicely with the view that mental functions rely on a collection of relatively independent processing components or *modules*, each dedicated to performing a particular function, which is a view embraced by many influential researchers. In fact, double dissociations and modularity go together so naturally that the theoretical perspective of modularity has dominated several subfields of neuroscience. As another

neuropsychologist, Andrew Ellis, put it: “There can be no argument with the fact of modularity, only about its nature and extent” (Ellis 1987, 402).<sup>5</sup>

### What Is Modularity?

So, how isolable are the parts of the brain? Modularity can be conceptualized in multiple ways:<sup>6</sup>

M1: Two parts A and B of a system are defined as modules, if and only if they are separately modifiable.

M2: The process carried out in the subsystem, so modifiable, computes a particular type of input-output mapping.

M3: There exists a decomposition of the system such that the computational interactions *within* subsystems are much more complex than those *between* subsystems.

M4: Subsystems form into complex networks with other subsystems so that each is carrying out only a particular subfunction of a much more complex overall function.

M5: The subsystem needs to be relatively spatially localized in the brain.

Property M1 provides a generic description of the idea, with M2 further specifying that some input-output relationship should be computed by the module in question. But what are modules? Perhaps they are fairly well delimited parts—say, Brodmann’s area 17 in the back of the brain corresponding to the primary visual cortex. But acknowledging that modules might involve more than just a single area, property M3 tries to capture a less restrictive notion of modularity, where one can think of subsystems (themselves perhaps composed of more elementary components) that are relatively encapsulated from *other* subsystems. Property M4 goes a step further, admitting that subsystems themselves are fairly complex and might interact with other subsystems. Still, being a subsystem, it should compute an identifiable “elementary” function.

Properties M1 to M4 are general and could apply to any system, natural or human-made. Property M5 is specific to the brain, of course, and is key to how the notion is conceptualized in neuroscience. Unless some version of property M5 holds, the system would not be recognized as modular. For example, it is possible to imagine a brain whose functions follow properties

M1 to M4 but is physically implemented in a spatially distributed fashion (imagine an artificial brain yet to be produced; incidentally, philosophers and cognitive scientists love to concoct all sorts of such challenging possibilities). In such case, the system would be *functionally* modular but not in terms of how it is instantiated in a *physical* medium.

It is useful to recast modularity in terms of *decomposability* (chapter 1). A decomposable system is one in which each subsystem does its job independently of the others. In contrast, in a non-decomposable system, the components are so interrelated as to defy attempts to break them up. What kind of system is the brain? Where along the spectrum of decomposable to non-decomposable does it reside?

This question is not idle armchair musing. It is at the core of our strategy to investigate the architecture of the mind-brain. Even more so because we must confront head-on the following question: Are there kinds of systems for which a *reductionistic analysis*—that is, one in terms of simpler subcomponents—would fail (Bechtel and Richardson 2010)? Here, reductionism means the type of approach central to science, in which an organization of greater complexity is understood in terms of the contributions of its subparts, which when put together give rise to the behavior of the broader system.

Consider the case of an object in which the components, perhaps simple computational elements not unlike neurons, do not perform operations that are (too) distinct from one another and for which the *interactions* between elements within the system are chiefly responsible for generating its behavior. To add to the difficulties, imagine a scenario in which the interactions between components are nonlinear, where, say, more of an input does not necessarily translate into more of an output (see chapter 8). In such cases, I contend that insurmountable difficulties arise in trying to unravel the object's working by reducing it to that of putative subcomponents.

To be sure, posing the question in this manner may sound counterintuitive to readers (and scientists alike) accustomed to successes of mechanistic analyses—the very bedrock of science. Indeed, reductionism is the declared philosophy of most scientists. Reduce everything to the smallest parts, determine their properties, and you explain the whole system. As developed throughout the book, I believe such an approach provides at best an impoverished description of brain function, as most of the explanatory work needs to be done at the level of *interactions*. Unfortunately,

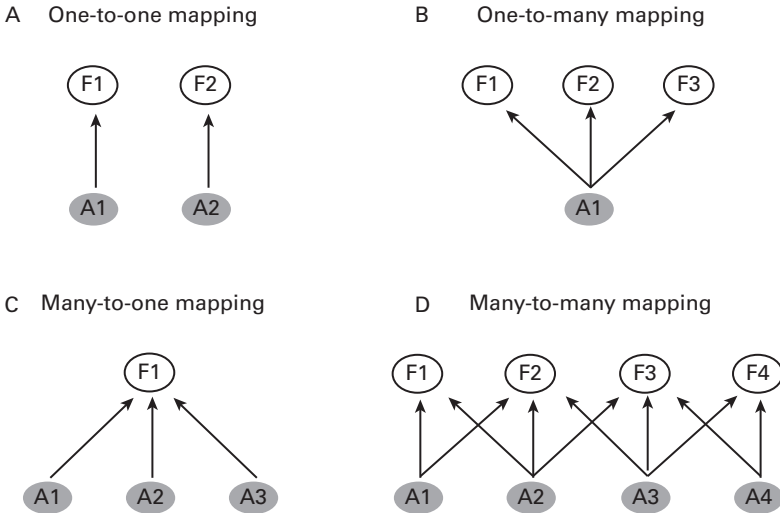


neuroscience as a discipline is all too reductionist. But rejecting the philosophy of reductionism is not an attack on scientific analysis (that is, the decomposition into parts and their analysis) and a concomitant embrace of some ill-specified holism. As stated by Ernst Mayr, sometimes hailed as the Darwin of the twentieth century, “No complex system can be understood except through careful analysis; however, the interactions of the components must be considered as much as the properties of the isolated components” (Mayr 2004, 34). We will have a lot more to say about these issues in chapter 8 when discussing *complex systems*.

### One Area, One Function?

Let’s go back to brain areas and consider, once more, their relationship to mental processes. We’ll start with the simplest formulation—namely, by assuming a one-to-one mapping between an area and its function. (We are assuming for the moment that we can come up with, and agree on, a set of criteria that defines what an area is. Maybe it’s what Brodmann defined early in the twentieth century, or perhaps it is as defined in the recent proposal discussed in chapter 1. For example, we could say that the function of the primary visual cortex is visual perception, or perhaps a more basic visual mechanism, such as detecting “edges” (sharp light-to-dark transitions) in images. The same type of description can be applied to other sensory (auditory, olfactory, and so on) and motor areas of the brain. This exercise becomes considerably less straightforward for brain areas that are not sensory or motor, as their workings become much more difficult to determine and describe. Nevertheless, in theory, we can imagine extending the idea to all parts of the brain. The result of this endeavor would be a list of area-function pairs:  $L = \{(A_1, F_1), (A_2, F_2), \dots, (A_n, F_n)\}$ , where areas  $A$  implement functions  $F$ .

To date, no such list has been systematically generated. However, current knowledge indicates that this strategy would *not* yield a simple area-function list. What may start as a simple  $(A_1, F_1)$  pair, as research progresses, gradually is revised and grows to include a list of functions, such that area  $A_1$  participates in a series of functions  $F_1, F_2, \dots, F_k$ . From initially proposing that the area implements a specific function, as additional studies accumulate, we come to see that the area *participates* in multiple functions. In other words, from a basic one-to-one  $A_1 \rightarrow F_1$  mapping the pictures evolves to a one-to-many mapping:  $A_1 \rightarrow \{F_1, F_2, \dots, F_k\}$  (figure 4.2a and 4.2b).

**Figure 4.2**

Structure-function mapping in the brain. (a) One area, A, might be involved in a single function, F, such as when people suggest that the amygdala is specialized for processing fear. (b) But we know that the amygdala carries out multiple functions. (c) From the standpoint of functions (such as aversive processing), multiple areas may be able to carry it out. (d) Elements of biological systems, like areas of the brain, exhibit the most complex mapping of them all: many-to-many.

Consider this example: Starting in the 1930s, lesion studies in monkeys suggested that the prefrontal cortex implements “working memory,” such as the ability to keep in mind a phone number for several seconds before dialing it. As research focusing on this part of the brain ramped up, the list of functions grew to include many cognitive operations, and the prefrontal cortex became central to our understanding of what is called *executive function* (see chapter 7). In fact, today, the list is not limited to cognitive processes but includes contributions to emotion and motivation. The prefrontal cortex is thus multifaceted. One may object that this sector is “too large” and that it naturally would be expected to participate in multiple processes. While this is a valid interjection, the argument holds for “small areas,” too. For example, take the amygdala, a region often associated with handling negative or aversive information. However, the amygdala also participates in the processing of appetitive items (and this multifunctionality applies even to amygdala subnuclei).

Let's consider the structure-function ( $A \rightarrow F$ ) mapping further from the perspective of the mental functions: Where in the brain is a given function  $F$  carried out? In experiments with functional MRI, tasks that impose cognitive challenges engage multiple areas of the frontal and parietal cortex. Examples are tasks requiring participants to selectively pay attention to certain stimuli among many and answer questions about the ones that are relevant (in a screen containing blue and red objects, are there more rectangles or circles that are blue?). These regions are important for paying attention and selecting information that may be further interrogated. Such *attentional control* regions are observed in circumscribed sectors of the frontal and parietal cortex. Thus, multiple individual regions are capable of carrying out a mental function, an instance of a *many-to-one* mapping:  $\{A_1 \text{ or } A_2, \dots, \text{ or } A_i\} \rightarrow F_1$ . (See figure 4.2c.) The explicit use of "or" here indicates that, say,  $A_1$  is capable of implementing  $F_1$ , but so are  $A_2$ , and so on.<sup>7</sup> Now, together, if brain regions participate in many functions and functions can be carried out by many regions, the ensuing structure-function mapping will be *many-to-many* (figure 4.2d). Needless to say, the study of systems with this property will be considerably more challenging than systems with a one-to-one organization. (For a related case, consider a situation where a gene contributes to many traits or physiological processes; conversely, traits or physiological processes depend on large sets of genes.)

Structure-function relationships can be defined at multiple levels, from the precise (for instance, primary visual cortex is concerned with detecting object borders) to the abstract (for instance, primary visual cortex is concerned with visual perception). Accordingly, structure-function relationships will depend on the granularity in question. Some researchers have suggested that, at some level of description, a brain region does *not* have more than one function; at the "proper" one, it will have a single function (Price and Friston 2005). In contrast, a central idea developed in this book is that the one-to-one framework, even if implicitly accepted or adopted by neuroscientists, is an oversimplification that hampers progress in understanding the mind and the brain.

### Brain Areas Are Multifaceted

If brain areas don't implement single processes, how should we characterize them? Instead of focusing on a single "summary function," it is better

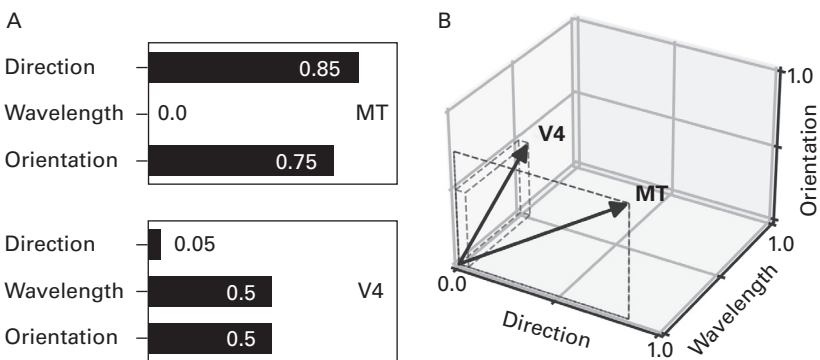
to describe an area's *functional repertoire*: Across a possibly large range of functions, to what extent does an area participate in each of them? No consensus has emerged about how to do this, but below we'll discuss some early results. The basic idea is simple, though. For example, coffee growers around the world think of flavor the same way: as a flavor profile or palette. Brazilian coffee is popular because it is very chocolaty and nutty with light acidity, to mention three attributes.

Research with animals uses electrophysiological recordings to measure neuronal responses to varied stimuli. The work is meticulous and painstaking because, until recently, the vast majority of studies are recorded from a single (or very few) electrode(s) in a single brain area. Setting up a project, a researcher thus decides what processes to investigate at what precise location—for example, probing classical conditioning in the amygdala. Having elected to do so, the electrode is inserted in multiple neighboring sites as the investigator determines the response characteristics of the cells in the area (newer techniques exist where grids of finely spaced electrodes can record from adjacent cells simultaneously; see chapter 12). For some regions, researchers have cataloged cell response properties for decades; considering the broader published literature thus allows them to have a fairly comprehensive view. In particular, the work of mapping cell responses has been the mainstay of perception and action research, given that the stimulus variables of interest can be manipulated systematically; it is easy to precisely change the physical properties of a visual stimulus, for example. In this manner, the visual properties of cells across more than a dozen areas in the occipital and temporal cortex have been studied. And several areas in the parietal and frontal cortex have been explored to determine neuronal responses during the preparation and elicitation of movements.

It is thus possible to summarize the proportions of functional cell types in a brain region.<sup>8</sup> Consider, for example, two brain regions in the visual cortex called V4 (visual area number 4) and MT (found in the middle temporal lobe). Approximately 85 percent of the cells in area MT show preference for the direction that a stimulus is moving (they respond more vigorously to rightward versus leftward motion, say), whereas only 5 percent of the cells in area V4 do so. In contrast, 50 percent of the cells in area V4 show a strong preference to the wavelength of the visual stimulus (related to a stimulus's color), whereas no cells in area MT appear to do so. Finally, 75 percent of the cells in area MT are tuned to the orientation of a visual stimulus (the

visual angle between the major elongation of a stimulus and a horizontal line), and 50 percent of the cells in area V4 do so, too. If we call these three properties  $DS$ ,  $ws$ , and  $os$  (for stimulus direction, wavelength, and orientation, respectively), we can summarize an area's responses by the triplet  $(DS, ws, os)$ , such that area MT can be described by  $(0.85, 0, 0.75)$  and area V4 by  $(0.05, 0.50, 0.50)$ , as shown in figure 4.3.

This type of summary description can be potentially very rich and immediately shifts the focus from thinking “this region computes  $X$ ” to “this region participates in multiple processes.” At the same time, the approach prompts us to consider several thorny questions. In the example, only three dimensions were used, each of which related to an attribute thought to be relevant—related to computing an object's movement, color, and shape, respectively. But why stop at three features? Sure, we can add properties, but there is no guarantee that we will cover all the “important” ones. In fact, at any given point in time, the attributes more likely reflect what researchers know and likely find interesting. This is one reason the framework becomes increasingly difficult for brain areas that aren't chiefly sensory or motor; whereas sensorimotor attributes may be more intuitive, cognitive, emotional, and motivational dimensions are much less so—in fact, they are constantly debated by researchers! So, what set of properties should we consider for the regions of the prefrontal cortex that are involved in an array of mental processes?



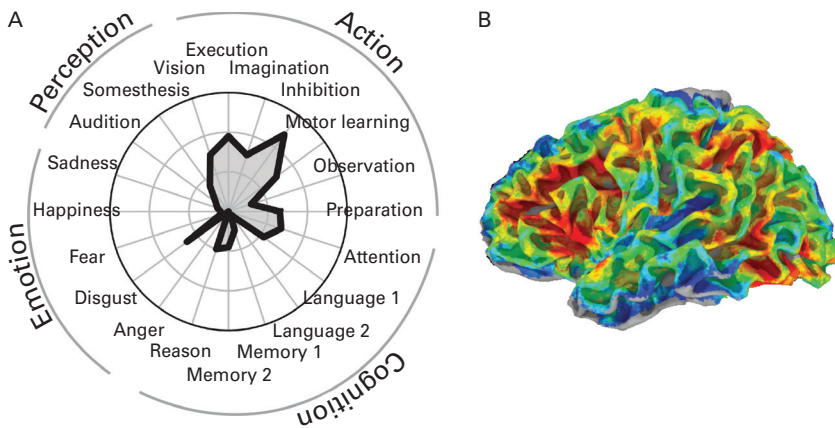
**Figure 4.3**

Multifunctional description of brain regions. (a) Visual areas MT and V4 can be described in terms of three attributes: direction, wavelength, and orientation. (b) Vector representation of the two areas.

More fundamentally, we would have to know, or have a good way of guessing, the appropriate *space of functions*. Is there a small set of functions that describes all of mentation? Are mental functions like phonemes in a language? English has approximately 42 phonemes, the basic sounds that make up spoken words. Are there 42 functions that define the entire “space” of mental processes? How about 420? Although we don’t have answers to these fundamental questions, some form of multifunction, multidimensional description of an area’s capabilities is needed. A single-function description is like a straitjacket that needs to be shed. (For readers with a mathematical background, an analogy to basic elements like phonemes is a “basis set” that spans a subspace, like in linear algebra, or “basis functions” that can be used to reconstruct arbitrary signals, like in Fourier or wavelet analysis.)

The multifunction approach can be illustrated by considering human neuroimaging research, including functional MRI. Despite the obvious limitations imposed by studying participants lying on their backs (many people will feel sleepy and may even momentarily doze off, not to mention that we can’t ask them to walk around and “produce behaviors”), the ability to probe the brain noninvasively and harmlessly means that we can scrutinize a staggering range of mental processes, from perception and action to problem solving and morality. With the growth of this literature, which accelerated in earnest after the publication in 1992 of the first functional MRI studies, several data repositories have been created that combine the results of thousands of studies in a single place.

In my laboratory, we capitalized on this treasure trove of results to characterize the “functional profile” of regions across the brain. We chose 20 “task domains” suggested to encompass a broad range of mental processes, including those linked to perception, action, emotion, and cognition. By considering the entire database of available published studies, at each brain location, we generated a 20-dimensional functional description indicating the relative degree of engagement of each of the 20 domain attributes (figure 4.4a). Essentially, we counted the number of times an activation was reported in that brain location, noting the task domain in question. For example, a study reporting stronger responses during a language task relative to a control task would count toward the “language” domain at the reported location. We found that brain regions are rather functionally



**Figure 4.4**

Multifunctionality of brain regions. (a) A radial plot shows the functional profile of a sample region. It includes 20 attributes across four classes (perception, action, cognition, emotion). The plot indicates the degree of engagement of the region for each attribute. For example, emotion studies involving disgust-related pictures or words engage the region strongly, but other emotion-related content does not. (b) The color insert shows the distribution of a measure of functional diversity across the cortex (warmer colors indicate higher diversity; cooler colors, less diversity).

Source: Panel B reproduced with permission from Anderson, Kinnison, and Pessoa (2013).

diverse and are engaged by tasks across many domains. But this didn't mean that they respond uniformly; they have preferences, which are at times more pronounced. To understand how multifunctionality varied across the brain, we computed a measure that summarized *functional diversity*. A brain region engaged by tasks across multiple domains would have high diversity, whereas those engaged by tasks in only a few domains would have low diversity. Functional diversity varied across the brain (figure 4.4b), with some brain regions being recruited by a very diverse range of experimental conditions.

The findings summarized in figure 4.4 paint a picture of brain regions as functionally diverse, each with a certain *style of computation*. The goal here was to illustrate the multidimensional approach rather than to present a more definitive picture. For one, conclusions were entirely based on a single technique, which has relatively low spatial resolution. (In functional

MRI, a signal at each location pools together processing related to a very large number of neurons; a typical location, called a “voxel,” can easily contain millions of neurons.) The approach also doesn’t account for the confirmation bias present in the literature. For example, researchers often associate amygdala activation with emotion and are thus more likely to publish results reflecting this association, a tendency that will increase the association between the amygdala and the domain “emotion” (not to mention that investigators might mean different things when they say “emotion”). Finally, the study makes the assumption that the 20-dimensional space of mental tasks is a reasonable decomposition. Many other breakdowns are possible, of course, and it might be even more informative to consider a collection of them at the same time (this would be like describing coffee in terms of a given set of attributes but then using separate groups of attributes).

### How Should We Think about Brain Areas?

Neuroscience has agonized over this question since its modern beginnings in the last few decades of the nineteenth century, as was forcefully summarized by Paul Morgane when describing the

... difficulty inherent in visualizing, or even conceptualizing, organizations of neurons distributed widely throughout cortical and subcortical structures, and somehow integrated into a functional unit, without recourse to an anatomically separate integrating center or system. ... (Morgane 1979, 14)

In other words, if not constrained to a spatially delimited area, then what?

Whereas science has proved exceedingly apt at describing modular systems, such as many found in physics and engineering, it has not made as much progress when the object of study is not as clearly decomposable. This is the case in both brain science and genetics. That both are within the realm of biology is not surprising, as biology has properties that are fairly unique and distinguish it from the physical sciences.<sup>9</sup>

The ideas charted in this chapter have implications to how we’ll *describe* brain regions in the remainder of the book. At first, a few of their important functions will be highlighted and the brain regions will be discussed mostly on their own; this is almost inevitable didactically. At times, therefore, it will appear as if the regions themselves are responsible for the functions



or behaviors described. Instead, the reader should bear in mind that brain regions participate in specific computations, functions, processes, or behaviors only when embedded within larger circuits comprised of multiple brain areas. Therefore, our understanding of the role of a specific region needs to be gradually bootstrapped so that eventually we will have a better appreciation for its functional contributions and repertoire, as we gain insight into how it interacts with other regions.



This is a section of [doi:10.7551/mitpress/14636.001.0001](https://doi.org/10.7551/mitpress/14636.001.0001)

# The Entangled Brain

## How Perception, Cognition, and Emotion Are Woven Together

By: Luiz Pessoa

### Citation:

*The Entangled Brain: How Perception, Cognition, and Emotion Are Woven Together*

By: Luiz Pessoa

DOI: [10.7551/mitpress/14636.001.0001](https://doi.org/10.7551/mitpress/14636.001.0001)

ISBN (electronic): 9780262372107

Publisher: The MIT Press

Published: 2022

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open



The MIT Press

© 2022 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC-BY-NC-ND license. Subject to such license, all rights are reserved.



The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Pessoa, Luiz, author.

Title: The entangled brain : how perception, cognition, and emotion are woven together / Luiz Pessoa.

Description: Cambridge, Massachusetts : The MIT Press, [2022] | Includes bibliographical references and index.

Identifiers: LCCN 2021061878 (print) | LCCN 2021061879 (ebook) | ISBN 9780262544603 (paperback) | ISBN 9780262372107 (pdf) | ISBN 9780262372114 (epub)

Subjects: LCSH: Perception. | Emotions and cognition. | Brain. | Neuropsychology.

Classification: LCC BF311 .P3767 2022 (print) | LCC BF311 (ebook) | DDC 153—dc23/eng/20220411

LC record available at <https://lcn.loc.gov/2021061878>

LC ebook record available at <https://lcn.loc.gov/2021061879>