Minds & Machines (2009) 19:255–271
DOI 10.1007/s11023-009-9154-6

Mechanisms and Functional Brain Areas

Gregory Johnson

Received: 12 November 2007 / Accepted: 18 May 2009 / Published online: 12 June 2009

Abstract Explanations of how psychological capacities are carried out often invoke functional brain areas. I argue that such explanations cannot succeed. Psychological capacities are carried out by identifiable entities and their activities in the brain, but functional brain areas are not the relevant entities. I proceed by assuming that if functional brain areas did carry out psychological capacities, then these brain areas could be included in descriptions of mechanisms. And if functional brain areas participate in mechanisms, then they must engage in activities. A number of ways in which we might understand the claim that functional brain areas engage in activities are examined. None are successful, and so one conclusion is that functional brain areas do not participate in mechanisms. Consequently, they are not the entities that carry out psychological capacities.

Keywords Functional brain area · Mechanism · Psychological capacity · Neurobiological explanation

Philosophers, cognitive scientists, and psychologists often invoke neurobiology when attempting to explain cognitive phenomena. In doing so, it is sometimes implied, if not stated outright, that talking about functional brain areas is one way of describing how the mind works. More specifically, there is a commitment to the idea that cognitive processes are carried out by functional brain areas such as the primary visual cortex, the hippocampus, or the amygdala. Anderson (2007) describes one procedure, what he calls "functional topography," that makes this idea attractive:

> I am following Fodor (2000) in the pragmatic definition of a (cognitive) function as whatever appears in one of the boxes in a psychologist's diagram

G. Johnson (✉)
Drexel University, Philadelphia, PA, USA
e-mail: gregory.s.johnson@drexel.edu

> of cognitive processing. The project of functional topography, simply put, is to map these boxes onto brain areas (p. 144).

A more specific example comes from Lewis (2005). He uses this sort of procedure when he identifies the brain areas involved in evaluating stimuli, although he does not describe (in what is quoted here) the actual processes by which these evaluations might occur:

> Evaluation is a slippery concept covering functions that extend from higher-level attentional systems such as the orbitofrontal cortex (OFC) and anterior cingulate cortex (ACC), to more basic pattern-matching functions of the amygdala (AM) and ventral striatum, to brainstem systems (e.g., the colliculi) that register innately defined events such as motion and rapid approach (p. 179).

I will argue that functional brain areas are not the entities that carry out psychological capacities. I share the commitment, expressed by Anderson and Lewis, that psychological capacities are carried out by identifiable entities and their activities in the brain, but I am skeptical of the claim that functional brain areas are those entities. But before going any further, the issue being addressed needs clarification.

One way to frame this problem is to ask: Do functional brain areas carry out psychological capacities? This, however, is a difficult question and open to being interpreted in a variety of different ways. Therefore, I will address a simpler question, namely, Can functional brain areas participate in mechanisms of the sort described by Machamer et al. (2000)? If we assume that psychological capacities are carried out by mechanisms of the type that Machamer et al. describe (or by a reasonably similar type), then establishing that brain areas do not participate in mechanisms will provide a negative answer to the original question. On the other hand, finding that brain areas do participate in mechanisms would not provide a definitive answer to the first question, but it would maintain functional brain areas as a possibility. In any case, I will defend the claim that functional brain areas do not carry out psychological capacities. To demonstrate this I will explain that functional brain areas do not participate in the type of mechanism described by Machamer et al.

There are several reasons for engaging this issue. The main one is that it has consequences for how we describe the working of the mind. To that end, the central goal of this paper is to make some progress on the issue of how neurobiological descriptions of psychological capacities should be constructed.

A second reason concerns the relationship between the psychological and the neurobiological. Insofar as functional brain areas are both large scale and neurobiological, mechanistic explanations using brain areas appear to provide a bridge between describing psychological capacities in computational terms and in terms of cellular and molecular activities. If, however, explanations that include functional brain areas are not available, then this idea of a bridge between the computational descriptions and the fine-grained neurobiological descriptions (i.e., the cellular and molecular) is, at least to an extent, undermined.

And third, many, if not most, phenomena in the natural world are uncontroversially located at particular levels of organization and carried out by specific types of entities and activities. For example, genes and the process of transcription are satisfactorily explained by referring to particular entities and activities found at the molecular level of organization, namely, DNA, RNA, and certain enzymes. Given that this view is acceptable for other biological phenomena and processes, it does not seem unreasonable to accept it in the case of psychological capacities. If, right now, we do not know what level is the appropriate one at which to locate psychological capacities, then we can make some progress by ruling out one set of entities as candidates: functional brain areas.[1]

Mechanisms

Machamer et al. (2000, p. 3) define mechanisms as "entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions". And describing entities and activities, they say, "Activities are the producers of change. Entities are the things that engage in activities."

For the purposes of this paper, the relevant entities are functional brain areas such as the primary visual cortex, the secondary visual cortex, the hippocampus, or even subdivisions of brain areas, for example, the lateral geniculate nucleus, which is a subdivision of the thalamus. The issue, then, is whether these entities participate in activities of the sort that are required in order for them to be considered parts of a mechanism. Clearly, these brain areas are, in some sense, entities, but merely saying that since they are entities, therefore, there must be activities associated with them is not a satisfying way to resolve this issue.

In more detail, Machamer et al. describe activities as follows:

> Activities…may be identified and individuated by their spatiotemporal location. They also may be individuated by their rate, duration, types of entities and types of properties that engage in them. More specific individuation conditions may include their mode of operation (e.g., contact action versus attraction at a distance), directionality (e.g., linear versus at right angles), polarity (attraction versus attraction and repulsion), energy requirements (e.g., how much energy is required to form or break a chemical bond), and the range of activity (e.g., electro-magnetic forces have a wider influence than do the strong and weak forces in the nucleus) (2000, p. 5).

In this explanation they focus on standard ways of describing a physical activity. A mechanism that is composed of functional brain areas would be in the domain of biology. Therefore, if there are such mechanisms, the activities will be some variety

[1] I will occasionally refer to levels in this paper. Since levels will not carry much of the explanatory burden, I will not spend too much time introducing them. By levels I mean *levels of organization* (see, for instance, Wimsatt 1976). And, in addition to a level occupied by functional brain areas, I will refer to a level occupied by neurons (the cellular level) and a level occupied by molecules (the molecular level). Furthermore, my view—which I do not defend in this paper—is that the cellular level and the molecular level are the levels at which the entities and activities that carry out psychological capacities occur.

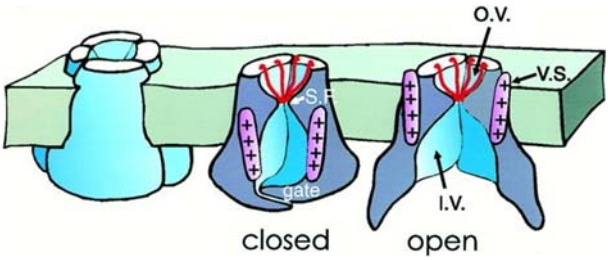

Fig. 1 Three views of a generic voltage-gated ion channel. The movement of the positively charged segments of the protein and the opening of the channel are shown in the *middle* and on the *right*. The segments labeled *S.F.* functions as a filter that allows sodium ions in, while excluding other ions. *V.S.* voltage sensor, *S.F.* selectivity filter, *O.V.* outer vestibule, *I.V.* inner vestibule. From Armstrong and Hille (1998, p. 372). Reprinted with permission from Elsevier

of the "standard" causal interaction (e.g., pushing or pulling), and not, for example, attraction at a distance.

A simple example of a mechanism is offered by Craver and Bechtel (2005). They describe a mousetrap as a mechanism composed of six entities: "a platform, a trigger, a latch, a catch, a spring, and an impact bar" (2005, p. 469). If we look at the activities that these entities engage in, we see that "*pressing* the trigger *releases* the catch, *allowing* the spring to *launch* the impact bar" (2005, p. 470). A more complex example, one example used by Machamer et al. (2000, p. 9) is the "mechanism by which [positively charged sodium ions] get inside the neuronal membrane"—in other words, one of the mechanisms for depolarization. Let us look at this example in detail.

The *set-up conditions* for any mechanism are the conditions present in the idealized "static time slice taken as the beginning of the mechanism" (Machamer et al. 2000, p. 11). For this depolarization mechanism, the main features of the set-up conditions are a closed voltage-gated sodium channel and a rising membrane potential.[2] The pore through which positively charged sodium ions flow is a long protein that is set into the cell membrane in such a way that it forms a channel (Fig. 1). In its resting state, when the membrane potential is below −40 mV, this channel is closed and neither the sodium ions, nor any other ions, can pass through it. However, parts of the molecule that form the channel are positively charged and sensitive to the rise in the membrane potential.

Machamer et al. call the activities that follow from the set-up conditions the *intermediate activities*. These are the activities of the mechanism. In this mechanism for depolarization the activities are the following. When the membrane potential reaches −40 mV the positively charged segments of the channel move outward

[2] The membrane potential is a measure of the electrical charge on the inside of the cell membrane relative to the electrical charge on the outside of the membrane. (The charge on the outside of the membrane is, by convention, set at zero, and so the membrane potential is a measure of the charge on the inside of the cell membrane.) At rest the membrane potential for a neuron is typically around −70 mV.

(relative to the membrane), which because of the helical structure of the protein is a rotating motion. This rotation causes changes in other parts of the protein that, in effect, open the channel. Positively charged sodium ions are then drawn into the neuron, which further raises the membrane potential.[3] Milliseconds after opening, the channel becomes inactive. A segment of the protein that forms the channel, but is not imbedded in the membrane itself, swings into the intracellular opening of the channel and blocks it. Thus, the flow of sodium ions through this channel is finished for the time being (Armstrong and Hille 1998; Yu and Catterall 2003).

The description of a mechanism ends with the *termination conditions*, "idealized states or parameters describing a privileged endpoint" for the mechanism (2000, p. 11). In the case of this mechanism, Machamer et al. (2000, p. 12) say that the termination condition is the "increase in intracellular Na^+ concentration and a corresponding increase in membrane voltage," which is the point when the channel closes.[4]

The entities that are part of this mechanism have been clearly specified; they are positively charged sodium ions and the long protein that forms the channel. The activities that these entities engage in have also been identified. These are (1) the protein forming the channel changes its conformation; (2) sodium ions move through the channel into the neuron; and (3) part of the protein swings into the mouth of the channel. The next section will address whether functional brain areas can participate in mechanisms in a similar, straightforward way.

But before moving on, we can note that the mechanism for depolarization is found at a fairly low level: the molecular level, or perhaps a level even lower—depending exactly on how we construe the conformational changes that mark the opening of the channel. Machamer et al. (2000, p. 13) do think that lower levels are privileged insofar as mechanistic descriptions "bottom-out" at a fairly low level where "the components… are accepted as relatively fundamental or taken to be unproblematic for the purpose of a given scientist, research group, or field." But mechanisms can be found at higher levels—as long as the requirements for a mechanism are satisfied. Moreover, they take lower level mechanisms to be components in higher level mechanisms. So, for instance, of the mechanism just reviewed they say, "the activation of the sodium channel is a component of the mechanism of depolarization, which is a component of the mechanism of chemical neurotransmission, which is a component of most higher-level mechanisms in the central nervous system" (2000, p. 13).

[3] The charged sodium ions (Na^+) are drawn into the neuron because of the concentration gradient that is created by the small amount of Na^+ inside the neuron compared to outside of it when in the resting state (i.e., when the membrane potential is −70 mV). There is also an electrical driving force that pulls Na^+ into the neuron until its equilibrium potential is reached (about 61.5 mV; Kingsley 2000).

[4] The inactivation of the sodium channel—when the channel is blocked by the portion of the protein that has swung into the mouth of the pore—is not the same as deactivation, which occurs when this part of the protein moves away and parts of the protein that form the channel itself are drawn together on the intracellular side. What should count as the termination conditions for this mechanism—inactivation or deactivation—is an interesting question, but it will not be pursued here.

Activities and Functional Brain Areas

Some functional brain areas that are candidates for being part of a single mechanism are those that are referred to when describing the first several stages of the visual process: the lateral geniculate nucleus (LGN), the primary visual cortex (V1), and the secondary visual cortex (V2). In the following sections several different ways of justifying the claim that these brain areas do participate in a mechanism will be examined. This investigation will focus on whether or not brain areas engage in activities that merit including them in a mechanism.

Interaction

The simplest way to begin is by seeking an interaction of some sort. Then, if need be, the type of interaction can be specified more carefully.[5] So, do these areas, for instance, the LGN and V1, interact in some way? The LGN and V1 do not touch each other, and merely saying that they interact by way of the axons of neurons is not an option, since axons are entities (or parts of entities) that participate in the mechanisms found at the cellular level.

Another option is to say that brain areas interact by way of *axon bundles*: structures such as the optic nerve and optic track, which connect the retina and the LGN,[6] or the optic radiations, which connect the LGN and V1. These are groups of axons that have their cell bodies in the same place and which terminate in the same general area. In order to pursue this option there has to be at least a minimal justification for thinking that axon bundles are different than the axons that are part of the cellular mechanisms found at a lower level.

One such justification might be that, at least in some cases, an axon bundle is a visible entity. The optic nerve, for instance, can be seen with the naked eye when the brain is removed from the skull. Presumably there is at least an intuitive desire to call something that can be seen, touched, and distinguish from other things an entity. This, therefore, might be one reason to say that the optic radiations are, along with the LGN and V1, part of a mechanism. The problem with this move, however, is that the only thing that it accomplishes is the introduction of another entity: the axon bundle. What sort of interactions these entities have with each other is still a question that has not been answered.

Let's take a moment to look at a different example. Take a neuron that has its cell body in one of the parvocellular layers of the LGN and sends its axon into V1 where it synapses on a spiny stellate cell in cortical layer 4Cβ. This neuron from the LGN interacts with the spiny stellate cell by way of an excitatory impulse, which under

[5] Machamer et al. are somewhat wary of the term *interaction*. They say, "Terms like 'cause' and 'interact' are abstract terms that need to be specified with a type of activity and are often so specified in typical scientific discourse. Anscombe (1981, p. 137) noted that the word 'cause' itself is highly general and only becomes meaningful when filled out by other, more specific, causal verbs" (2000, p. 6). Nevertheless, for our present purposes, *interaction* seems like a good place to start.

[6] The *optic nerves* designate the axons between the retinae and the optic chiasm (where some of the axons cross from the right eye to the left hemisphere of the brain, and from the left eye to the right hemisphere) and the *optic tracks* are the continuation of these axons from the optic chiasm to the LGN.

the right conditions will contribute to the spiny stellate cell generating an action potential and thereby exciting other neurons. In this case both the entities and the activities that they engage in are clearly identified. There are neurons, they generate action potentials, and multiple neurons that concurrently release excitatory neurotransmitter onto a spiny stellate cell in V1 will drive that cell to threshold, causing it to generate its own action potential. Thus, this interaction between the neuron that is in the parvocellular layer of the LGN and the spiny stellate cell is part of a mechanism that is found at the cellular level.

The cellular mechanism contrasts quite clearly with the functional brain areas. The brain areas do not push or pull each other, and there does not seem to be any other causal language to describe their interaction. Therefore, at this point we can say that there is not an obvious way of describing how brain areas interact with each other, even if we allow the axon bundles into our description. And so for two of the candidate brain areas, the LGN and V1, it is not clear what physical interaction they have that would merit including them in a description of a mechanism.

Additive Effects of the Activities of Neurons

The previous section was a first attempt at determining whether or not two brain areas such as the LGN and V1 interact, and the result was that there does not appear to be a straightforward way of describing an interaction between the brain areas themselves (i.e., qua brain areas). Another, somewhat less straightforward, way of understanding the interaction between functional brain areas will be examined in this section. Here I will focus on the additive effects of the individual neurons in the LGN interacting with neurons in V1 (i.e., all of the excitatory and inhibitory interactions). The interaction between these two brain areas could, perhaps, be understood as the total activity at any one time of the neurons whose axons project from the LGN to V1. Examining this case requires invoking details that are specific to the early part of the visual process. Consequently, the conclusion will only be about these brain areas. It is likely that the basic theme of the conclusion generalizes, but I will not pursue that here. In any case, if there is no way to construe a mechanism for vision that includes the LGN and V1, that is a significant problem for the proposal that brain areas participate in mechanisms.

First, let's look at the activity among neurons again. In the previous section the example of a neuron in the LGN interacting with a neuron in V1 was discussed. Broadening the example a bit, one way of understanding some of the interactions between the neurons in the LGN and those in V1 was proposed by Hubel and Wiesel (1962). The mechanism that they described is shown in Fig. 2. Retinal ganglion cells respond when light hits a particular location on the retina, and since each retinal ganglion cell projects to a single neuron in the LGN, the neurons in the LGN also respond when a spot of light hits a particular location on the retina. If a bar of light falls on the retina, the neurons in the LGN that are active are only responding to the specific patch of light in their receptive field. But since the bar of light stretches across the receptive fields of several of these neurons, several will be active. These neurons in the LGN may then project to a single neuron in V1. The collective excitatory input that the neuron in V1 receives when a bar of light hits the retina

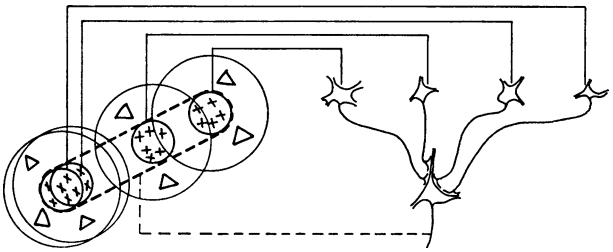

Fig. 2 A diagram illustrating a possible series of connections between retinal ganglion cells, neurons in the LGN (*top right*), and a neuron in V1 (*bottom right*). The *circles* on the *left* represent the receptive fields of the retinal ganglion cells (*one circle* for each cell). These retinal ganglion cells have "on-center" receptive fields, meaning that they will respond when light hits the center of the field (where the *X*s are in the diagram). The axon of each of the retinal ganglion cells contacts a single neuron in the LGN. The neurons in the LGN all project to the neuron in V1. Thus, the neuron in V1 is active when a bar of light, with the orientation outlined on the *left*, hits the retina. From Hubel and Wiesel (1962, p. 142)

causes this neuron to become active. Thus, the neuron in V1 responses to a bar of light that has a particular orientation and hits a particular location on the retina.

Information about the orientation of a bar of light is one example of the kind of information that passes from neurons in the LGN to the neurons in V1. The neurons in the LGN respond to spots of light hitting the retina, and, in virtue of the connections among these neurons, a neuron in V1 responds to a bar of light.

Now that this example has been explained, we can return to brain areas and the question of whether the interaction between the LGN and V1 can be understood as the total activity at any one time of the neurons whose axons project from the LGN to V1. Of course, there is such activity. But to be meaningful activity, there has to be some way of characterizing it; that is, we need an answer to the question: What is this activity doing?

One problem that is immediately encountered in answering this question is this: if we merely add together the activity of all of the neurons in the LGN that project to V1, then the way that information is characterized when talking about the activity among neurons is no longer available. Although we can observe (with the right tools) a neuron that responds to a bar of light at a particular orientation, there is no way to do the same with the brain area V1. Viewed as a brain area, V1 is active in the same general way in the presence of bars of light at any orientation.[7] And the same holds for other information that is encoded at the cellular level, for example, information about color and motion. Characterizing visual information with any specificity in terms of the orientation of bars of light, motion, or color is simply not possible when talking about the activity of brain areas.

But if there is information that moves, at the level of functional brain areas, from the LGN to V1, how can this information be characterized? One possibility is that the information at the brain areas level is less fine-grained than the information at the cellular level. If this is the case, then it is reasonable to characterize this

[7] That is to say, V1 (qua brain area) is equally active when a bar of light at a 70° orientation is hitting the retina and when a bar of light at a 45° orientation is hitting the retina. The relevant differences in activity only show up when we drop down to the level of neurons.

information as the visual scene (or something similar) that an individual sees. This proposal immediately runs into difficulty, however. It is wrong to say that the visual scene (out in the world) is encoded as a whole at the retina and then transferred from the LGN to V1. It has been well established that "bottom-up" processing (from the retina to the cortex) is, by itself, insufficient to account for what we see. Cavanagh (1999) explains the problem as follows:

> Perception represents the immediate present, what is happening around us as conveyed by the pattern of light falling on our retina. And yet the current pattern of light alone cannot explain the stable, rich experience we have of our surroundings. The problem is that each retinal image could have arisen from any of a vast number of possible 3-D scenes. That we rapidly perceive only one interpretation tells us that we see far more than the immediate information falling on our retina (p. 839).

That "bottom-up" processing is insufficient to account for what we see presents a number of difficulties for explaining our visual capacity. Right now, however, it just presents the single problem that if there is visual information that is passed from the LGN to V1, it cannot be defined simply as the information about the scene out there in the world.

To review, when we look at the activities of neurons, we have access to information about specific features of the visual scene; these features, it can be said, are what the neurons are responding to. But when we move up to brain areas, we do not have access to the all of the information about the scene at a less fine-grained level. Perhaps there is a way to capture the content of the specific information that passes from the LGN to V1 (qua brain areas), but it is not clear how this information can be characterized.

Before moving on I want to address a related way of construing the activity of brain areas, namely, treating the total activity of the neurons within a brain area as the activity of the brain area. (This section began with a slightly different scenario: the total activity of the neurons that project from one brain area to another.) Consider this analogy. By understanding the activity of a brain area as the collective activity of neurons within that brain area, the brain area is being treated like a building. A building is active if there is a sufficient amount of activity occurring within it. This way of talking is acceptable, but no one would suggest that just because a building is active that it interacts—in a standard (push-pull) causal way—with other buildings. But standard causal interactions would be needed if we were to propose a mechanism that included buildings. Thus, even when a building is active, it cannot be a part of a mechanism. Therefore, with regard to brain areas, the activity occurring "within" a brain area is not sufficient to demonstrate that brain areas participate in mechanisms.

Psychological Language

At this point, one might object by saying that there are examples in the literature of neuroscientists talking explicitly about the activities of brain areas. In this section I will look at and comment on one such example.

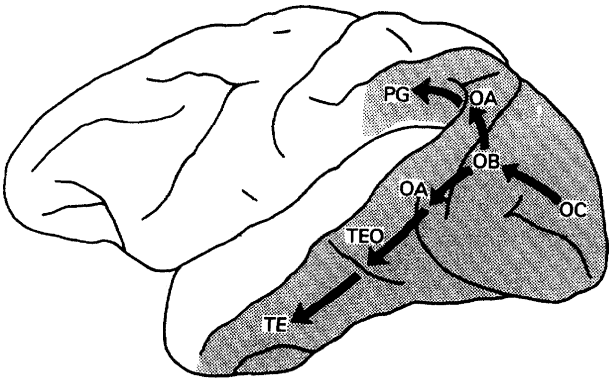

Fig. 3 The two visual pathways identified by Mishkin et al. (1983). The object identification pathway is the ventral (i.e., downward) pathway. The pathway that locates objects in space is the dorsal pathway. From Mishkin et al. (1983, p. 414). Reprinted with permission from Elsevier

Mishkin et al. (1983) identified two visual processing pathways and described them in terms of functional brain areas. Both of these pathways begin in V1 (referred to by them as the *striate cortex*) and then diverge (Fig. 3). The first pathway, which "interconnects the striate, prestriate, and inferior temporal areas, is crucial for the visual identification of objects," and the other, "which interconnects the sriate, prestriate, and inferior parietal areas, is critical for the visual location of objects" (1983, p. 414). Looking more closely at the object identification pathway, Mishkin et al. say the following:

> The anterior part of inferior temporal cortex, or area TE…, is the last exclusively visual area in the pathway that begins in the striate cortex, or area OC [i.e., V1], and continues through the prestriate and posterior temporal areas, OB [V2], OA [V3 and V4] and TEO [the posterior temporal cortex]. This ventrally directed chain of cortical visual areas appears to extract stimulus-quality information from the retinal input to the striate cortex, processing it for the purpose of identifying the visual stimulus and ultimately assigning it some meaning through the mediation of area TE's connections with the limbic and frontal-lobe systems. According to this view, the analysis of the physical properties of a visual object (such as its size, color, texture and shape) is performed in the multiple subdivisions of the prestriate-posterior temporal complex and may even be completed within this tissue (1983, p. 414).

If we examine this passage for an indication of the activities that these brain areas perform, we find that they *extract* information about the stimulus, *identify* this information, and *assign* it a meaning. Insofar as these are causal verbs, they are candidates for the activities that occur in a mechanism. In what follows I will argue against this proposal and will attempt to show that this passage can be a meaningful description of what is going on in the brain without being a description of a mechanism that is composed of functional brain areas.

An alternative possibility is that each of these activities is not the activity of a particular entity in a mechanism—i.e., a particular brain area—but rather is a description of the process that some mechanism carries out. Note the important difference between a description of the process that a mechanism carries out and the description of the activities that are part of a mechanism.[8] For clarity, I will adopt Craver's (2002) term 'role' for the former, and will use *activities* for the latter. A mechanism carries out (or performs) a role; the entities that compose the mechanism do not themselves perform that same role. For example, depolarization is the role that a certain mechanism performs. In this mechanism none of the entities depolarize (e.g., the sodium ions do not depolarize, they just enter the neuron).[9]

Therefore, one possibility is that taking stimulus information that has been encoded by the retina, processing it, and assigning some meaning to it is a role that is performed by a mechanism or series of mechanisms in this pathway that Mishkin et al. have identified. Establishing that this role is carried out in this particular area of the brain is important for understanding vision, even if it does not specify a mechanism.

Nevertheless, let us assume that one of these brain areas does perform one of these activities. For the sake of argument, let us say that "identifying the visual stimulus" is the activity performed by the posterior temporal cortex (TEO), in which case we have matched an activity with an entity. There are several difficulties with this proposal. I will examine three such difficulties, the first two of which are worrisome, but perhaps not debilitating. The third returns to the issue of participation in a mechanism and is a more direct problem.

To begin with, we have a biological entity (i.e., an entity that we identify in biological terms), area TEO, and a non-biological activity, identifying a stimulus. There are a couple of reasons why this could indicate a problem. One is that for a description of a mechanism, it seems natural to assume that biological entities participate in biological activities. This may not be a firm requirement, but it seems to be a safe way to ensure that we are describing the causal interactions that are needed for a mechanism. Second, we have to keep in mind the difference, mentioned above between describing the role that a mechanism performs and describing the activities that constitute the mechanism. We expect that the description of the role will be in psychological language because we are focusing on part of a psychological capacity. And unless we have good reasons to think otherwise, the simplest assumption is that *identifying a stimulus* is a psychological

[8] Machamer et al. (2000, p. 2) do not devote much attention to this distinction, but they do say, "Mechanisms are sought to explain how a phenomenon comes about or how some significant process works." Craver (2002) calls the process that the mechanism carries out the *role*, and Craver and Bechtel (2005) call it the *phenomenal aspect* of the mechanism.

[9] Bechtel (2005) has a very useful example concerning investigations into the mechanism that carries out fermentation. Researchers in the late nineteenth century were not in a position to make a distinction between the role that the mechanism performs and the activities of the mechanism itself. Without having the correct biochemical framework to discover the mechanism, researchers were, as Bechtel says, stuck asking "whether methylglyoxal, for example, would *ferment* as rapidly as sugar" (p. 318). But, Bechtel continues, "decomposing fermentation into fermentations simply invoked the vocabulary designed to explain the overall behavior to describe the operation of its components. It did not explain the process in terms of something more basic" (p. 318).

process carried out by a biological mechanism—perhaps one that is located in area TEO.

The third problem is that this activity, identification of the stimulus, has to have implications for some other brain area that is part of (what would be) this mechanism. Thus, some other brain area has to be capable of using or acting on the identification that is made by area TEO. And most importantly, we need to know what sort of interface is used for this interaction. Prima facie, this last issue does not have an obvious answer, which suggests that area TEO is not an entity that participates in a mechanism.

Techniques

Another objection to the argument presented so far is that some of the techniques used by neuroscientists in their investigations do, rather explicitly, invoke functional brain areas. The two most relevant techniques are lesion studies and functional imaging studies. Lesion studies demonstrate that the absence of, or damage to, a functional brain area can disrupt the operation of a psychological capacity, and functional imaging studies appear to show when a particular brain area is engaged in a specific task.

Lesion Studies

The logic behind using brain lesions to determine brain function is simple. When a brain area is destroyed or otherwise rendered inoperable, the behavior of the organism is studied in order to see what psychological capacities have been lost. There are a number of different lesioning methods, the most familiar of which are (1) the accidental damage to a particular part of the brain;[10] (2) the surgical removal of specific brain areas in animals so that controlled experiments can be performed;[11] and (3) temporary lesioning, which makes an area inactive for only a period of time.[12]

Mishkin et al. relied on a number of lesion studies to construct their model of the two visual pathways. In one study monkeys were shown an object once and then several minutes later given a choice between that object and a new object; a monkey was rewarded when it selected the new object (Mishkin 1982). Once trained on this task, normal monkeys are able to perform it easily, but monkeys with bilateral lesions of area TE cannot. As Mishkin et al. (1983, p. 415) say, “the area in which

[10] One of the most well known examples of this type of lesioning occurred to Phineas Gage who had a rod driven through the ventromedial prefrontal region of his brain (Damasio 1994). Another well known case is H.M. who underwent the bi-lateral removal of his medial temporal lobes in a procedure performed to reduce the occurrence of epileptic seizures (Scoville and Milner 1957; Corkin 2002).

[11] In these studies some animals have a particular brain area removed—often bilaterally—while control animals do not. All animals are then given tests of one type or another in order to measure the deficiency created by the lesioning. See, for example, Zola-Morgan and Squire (1984).

[12] Most types of temporary lesioning, for example, transcranial magnetic stimulation (Amassian et al. 1989), injection of muscimol (Robinson 2000), or injection of sodium amobarbital into the carotid artery (Binder et al. 1996), are better analyzed, not as lesioning studies, but as cellular or molecular interventions.

the neural trace [left by the previous viewing of the object] appears to be preferentially established is area TE, since lesions here—but not lesions elsewhere in the cortical visual system—nearly abolish the monkey's ability to perform the recognition task." One lesson that might be drawn, then, is that this brain area, area TE, is part of the mechanism for visual recall; area TE functions as a store of previously viewed stimuli and a "comparer" of the earlier viewed stimuli and the present one.

There are a couple of reasons to suggest that this conclusion is not warranted. First, a lesion of this sort clearly affects the entities and activities at lower levels of organization, and so the behavioral effects of the lesion could be attributed to destroying the entities or disrupting the activities at a lower level. But besides this possibility, a more important consideration has to do with the nature of lesion studies. In order to have a mechanism we need to observe the activity of these areas, not just the deficits that accrue when they are removed. Using the lesioning technique in order to investigate what behaviors rely on a particular brain area clearly does not provide direct information about how brain areas interact. Therefore, evidence from lesion work is not sufficient for grounding the claim that brain areas participate in mechanisms.

Imaging Studies

The other technique that, it might be claimed, measures the activity of brain areas is functional brain imaging. Two types of functional brain imaging will be examined here: functional magnetic resonance imaging (fMRI) and positron emission tomography (PET).

In order to understand how fMRI works, we first need to look at magnetic resonance imaging (MRI). MRI is a technique that measures the nuclear magnetic resonance that atomic nuclei give off after they have been subjected to a brief radio frequency pulse. When an individual's brain is being imaged, the MRI scanner's magnet causes the axis of some of the atomic nuclei to become aligned (along the north-south axis of the magnetic field). The radio frequency pulse then causes this alignment to shift and the proton(s) of these atoms to spin in phase with each other. When the radio frequency pulse is over, the nuclei return to their prior state. Nuclei shift back, and the proton spinning comes out of phase. This releases energy—a radio frequency pulse (the resonance)—which is given off by the atoms, and atoms in different tissue (e.g., intracellular, extracellular, in fat, in water, in the blood, or cerebral spinal fluid) give off differing amounts. The resonance is detected and from it an MRI image is generated (Kingsley 2000; Saper et al. 2000).

The result of an MRI is an image of a cross section (a slice) of the brain. The advantage of functional MRI (fMRI) over MRI is that fMRI produces an image that locates areas in the brain that are active. It does this by focusing on the hemoglobin in the blood supply. The supply of oxygenated blood to areas of the brain that are active is greater than is necessary. Therefore, the ratio of oxygenated blood (oxyhemoglobin) to deoxygenated blood (deoxyhemoglobin) actually increases in areas of activity. Oxyhemoglobin and deoxyhemoglobin also have different magnetic properties. After the radiofrequency pulse has ended, the protons in the

oxyhemoglobin stay in phase longer than the protons in the deoxyhemoglobin. Thus, the areas of the brain that have a higher ratio of oxyhemoglobin also have a stronger MRI signal (Saper et al. 2000).

In the other brain imaging technique considered here, PET, the subject is given (by injection to the bloodstream or inhalation) a small amount of a biological compound that is used by the brain, and which includes a short-lived radioactive isotope. For example, ^{18}F is an unstable isotope that can be used to create ^{18}F-deoxyglucose. The ^{18}F-deoxyglucose behaves like deoxyglucose except that it cannot be metabolized, and so it just builds up in neurons. The build up of ^{18}F-deoxyglucose indicates that those neurons are consuming glucose, which is an indication that they are active. Meanwhile, the unstable ^{18}F decays and emits a positron that eventually collides with an electron, destroying both and creating two gamma rays (photons). The gamma rays are detected by the PET scanner and this information is used to create an image highlighting the area where glucose consumption was occurring (or, more precisely, where the collision between the positron and the electron occurred).

For both of these imaging techniques, the information that is produced is reconstructed into an image of the brain. By looking at the image, one can see which brain areas are active. One might conclude from this that, since fMRI and PET studies demonstrate which brain areas are active when a particular task is being performed, brain areas must be able to participate in mechanisms. That is, however we characterize this activity, it is the activity of brain areas, and if the brain areas are active, then they can be parts of mechanisms.

This conclusion is hasty. We can separate two issues here. First, there are the entities that the researchers who are using these methods of investigation are trying to access (i.e., the entities that they want information about). Second, there are the entities and activities that these techniques are directly recording. PET records information about collisions between sub-atomic particles, and fMRI records information about the spinning of sub-atomic particles. These activities occur at a very low level, a level that we would normally think of as the domain of physics. But researchers studying the brain are not really trying to gather information about activity at this very low level. Rather, they are using the activity at this low level to make inferences about activity occurring at higher levels.

For example, when using fMRI, one infers the presence of oxyhemoglobin from the different rates of the dephasing of protons, and from the presence of oxyhemoglobin one infers the activity (and the oxygen needs) of neurons. We can stop here, which would mean that, when using fMRI, researchers are collecting information about the activities of neurons.

To make a claim about a brain area requires a further inference. This would be a simple inference, basically just the following: if the neurons in this particular area are active, then this brain area is active. This is simple enough, but it is very close to being an empty claim. The problem is that there is nothing that the brain area is doing that leads to the inference that the brain area is active except the statement just mentioned. For neurons, we take the fact that oxyhemoglobin is being delivered to neurons, include some facts about the consumption of oxygen, and conclude that these neurons are receiving inputs from other neurons and possibly generating their

own action potentials in order to pass a signal on to other neurons. There is no similar sort of claim that can be made about brain areas because there is nothing that they do except have the property of being composed of neurons.

In defense of the claim that it is brain areas that are active, one might point out that these imaging techniques are a fairly crude method of gaining access to cellular activity (although they have the practical, and ethical, benefit of being non-lethal to the subjects). Given the "grain" of these imagining techniques, a fairly large populations of neurons are identified, and so when we look at the results from an imaging study we cannot exactly say which neurons are active; all we know is that a lot of them in a certain area are active. We can, however, say (fairly) definitively that a particular brain area is active. Therefore, one might conclude that these techniques are measuring the activity of brain areas.

This is a reasonable argument, and may be the logic behind how people talk about the results of imaging studies. It does not, however, solve the problem of finding activities for a description of a mechanism.[13] In particular, it does not identify a causal interaction that one brain area has with another brain area.

Conclusion

The goal of this paper has been to demonstrate that psychological capacities are not carried out by functional brain areas. I have examined whether brain areas participate in mechanisms of the type that have been described by Machamer et al. (2000). Several ways that we might understand the claim that brain areas are entities that participate in mechanisms have been examined, but all of these possibilities have encountered problems. The primary consequence of this result is that when we seek the entities that carry out psychological capacities we have to look to a lower level—the level where neurons are found is probably the most promising option. To close, I will look at one more potential objection, and then I will expand on one of the comments that I made in the first section concerning the consequences of the view that brain areas do not participate in mechanisms.

My project might be construed as an eliminativist one insofar as I am attempting to eliminate functional brain areas from our ontology. It might, then, be pointed out that lower level entities such as neurons are equally susceptible to being eliminated. In response to this, let me clarify that I am not trying to eliminate brain areas as entities. I take it that brain areas are useful for some projects, for example, comparative neuroanatomy or developmental neuroanatomy. They are not, however, useful when the goal is a description of a mechanism. Furthermore, although it may be that at some time in the future neurons themselves will not be used to explain how psychological capacities are carried out, there is still an important difference between brain areas and neurons. For neurons, it is easy to describe how they interact, whereas for brain areas, describing their interactions is the fundamental problem.

[13] See also the argument made earlier about activity within a brain area (at the end of the section titled "Additive Effects of the Activity of Neurons").

Lastly, the work that has been done here is intended to have consequences for our understanding of the relationship between psychological descriptions and neurobiological descriptions. The idea that descriptions involving brain areas can serve as a bridge between computational descriptions and the fine-grained neurobiological has been undermined—at least to a degree. One account that relies on such a bridge is what can be called the decomposition view of the relationship between the psychological and neurobiological. This view has been developed and defended by Lycan (1981, 1987; see also Craver 2002). In short, Lycan's view is that a psychological capacity can be broken down into component parts that carry out the capacity, and then each of those components can be further decomposed into the entities and activities that are responsible for carrying out the components' functions. If this process is reiterated a number of times, there will be a smooth shift—via this sort of decomposition—from the psychological description that we started with to biological and chemical descriptions.

Churchland and Sejnowski (1990) have a related view. They say of the reduction of psychological descriptions to neurobiological ones:

> [I]t should be emphasized that the explanation of high level cognitive phenomena will not be achieved directly in terms of phenomena at the lowest level of nervous-system organization, such as synapses and individual neurons. Rather, the explanation will refer to properties at higher structural levels, such as networks or systems. Functional properties of networks and systems will be explained by reference to properties at the next level down, and so on (p. 351).

Although Lycan's and Churchland and Sejnowski's ultimate commitments differ (the former takes a non-reductive stance and the latter a reductive one), both dependent on there being a series of levels in place so that the move from the psychological description to the lower level neurobiological description is smooth (i.e., so that the decomposition from the psychological to the neurobiological will not involve any "jumps"). But this kind of smooth shifting may not be an option simply because there are not biological mechanisms at very many different scales. In particular, if the conclusions drawn here are correct, then there is not a level where brain areas participate in mechanisms. Therefore, the relationship between the psychological and the neurobiological that Lycan describes may not exist. And, contra Churchland and Sejnowski, it may be the case that descriptions of psychological capacities do have to "be achieved directly in terms of phenomena at the lowest level of nervous-system organization."

Acknowledgments Thanks to Tom Polger, John Bickle, Bob Richardson, and Jenefer Robinson for comments on earlier versions of this material.

References

Amassian, V. E., Cracco, R. Q., Maccabee, P. J., Cracco, J. B., Rudell, A., & Eberle, L. (1989). Suppression of visual perception by magnetic coil stimulation of human occipital cortex. *Electroencephalography and Clinical Neurophysiology, 74*, 458–462.

Anderson, M. (2007). The massive redeployment hypothesis and the functional topography of the brain. *Philosophical Psychology, 20*, 143–174.

Anscombe, G. E. M. (1981). *Metaphysics and the Philosophy of Mind*. Minneapolis, MN: University of Minnesota Press.

Armstrong, C. M., & Hille, B. (1998). Voltage-gated ion channels and electrical excitability. *Neuron, 20*, 371–380.

Bechtel, W. (2005). The challenge of characterizing operations in the mechanisms underlying behavior. *Journal of the Experimental Analysis of Behavior, 84*, 313–325.

Binder, J. R., Swanson, S. J., Hammeke, T. A., Morris, G. L., Mueller, W. M., Fischer, M., et al. (1996). Determination of language dominance using functional MRI: A comparison with the Wada test. *Neurology, 46*, 978–984.

Cavanagh, P. (1999). Top-down processing in vision. In R. A. Wilson & F. C. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 839–840). Cambridge, MA: The MIT Press.

Churchland, P. S., & Sejnowski, T. J. (1990). Neural representation and neural computation. *Philosophical Perspectives, 4*, 343–382.

Corkin, S. (2002). What's new with the amnesic patient H.M.? *Nature Reviews Neuroscience, 3*, 153–160.

Craver, C. F. (2002). Interlevel experiments and multilevel mechanisms in the neuroscience of memory. *Philosophy of Science, 69*, S83–S97.

Craver, C., & Bechtel, W. (2005). Mechanism. In S. Sarkar & J. Pfeifer (Eds.), *The philosophy of science: An encyclopedia* (pp. 469–478). New York: Routledge.

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam.

Fodor, J. (2000). *The mind doesn't work that way*. Cambridge, MA: MIT Press.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology, 160*, 106–154.

Kingsley, R. E. (2000). *Concise text of neuroscience*. Philadelphia: Lippincott Williams & Wilkins.

Lewis, M. D. (2005). Bridging emotion theory and neurobiology through dynamic systems modeling. *Behavioral and Brain Sciences, 28*, 169–194.

Lycan, W. G. (1981). Form, function, and feel. *The Journal of Philosophy, 78*, 24–50.

Lycan, W. G. (1987). *Consciousness*. Cambridge, MA: MIT Press.

Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1–25.

Mishkin, M. (1982). A memory system in the monkey. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 298*, 85–95.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences, 6*, 414–417.

Robinson, F. R. (2000). Role of the cerebellar posterior interpositus nucleus in saccades. I. Effect of temporary lesions. *Journal of Neurophysiology, 84*, 1289–1302.

Saper, C. B., Iverson, S., & Frackowiak, R. (2000). Integration of sensory and motor function: The association areas of the cerebral cortex and the cognitive capabilities of the brain. In E. R. Kandel, J. H. Schwartz, & T. M. Jessell (Eds.), *Principles of neural science* (pp. 349–380). New York: McGraw-Hill.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20*, 11–21.

Wimsatt, W. C. (1976). Reductionism, levels of organization, and the mind-body problem. In G. G. Globus, M. Grover, & I. Savodnik (Eds.), *Consciousness and the brain: A scientific and philosophical inquiry* (pp. 205–267). New York: Plenum Press.

Yu, F., & Catterall, W. (2003). Overview of the voltage-gated sodium channel family. *Genome Biology, 4*, 207.

Zola-Morgan, S., & Squire, L. (1984). Preserved learning in monkeys with medial temporal lesions: Sparing of motor and cognitive skills. *Journal of Neuroscience, 4*, 1072–1085.