MARK JOHNSON

# *The Meaning of the Body*

## AESTHETICS OF HUMAN UNDERSTANDING

*The University of Chicago Press* Chicago & London

One of the greatest obstacles to a general acknowledgment of the embodiment of mind, meaning, and thought is the persistence—in commonsense, scientific, and philosophical models alike—of the representational theory of mind. As we will see in the next chapter, the representational theory claims that the "mind" operates on "internal representations" (ideas, concepts, images) that can re-present (and thereby "be about") external objects and events. The representational theory had its source in dualistic metaphysical views that mind is separate from and different in kind from body, that what is inner is different in kind from what is outer, and that we have a direct access to the inner that is not available to us for the outer. Even today, some proponents of nondualistic views nevertheless allow a dualistic ontology to creep back into their theory whenever they accept some version of the representational theory.

The naturalistic theory of embodied meaning, mind, and language that I am developing is thus directly at odds with classical representational theories of mind. One major theme of part 2 of this book is that we can provide a philosophically and scientifically realistic alternative to the representational theory of mind. What is required is an outline of how we might begin to explain conceptualization and reasoning without assuming a representational perspective. Chapters 6 and 7 place human cognition in a broader context of animal cognition, drawing on research in the biological sciences about the embodied, situated, goal-directed nature of animal and human cognition. The key point is to see how organism-environment couplings do not require an internal mind that manipulates symbols referring to external things and events. I argue for the continuity of human thinking with these less sophisticated, less complex engagements of animals with their world. Chapter 8 investigates some of the neural bases for our ability to process both concrete and abstract concepts. I draw on recent research in cognitive neuroscience about the role of neural ensembles in conceptualization. Then, in chapter 9, I make a case for the embodied character of abstract concepts, drawing evidence from cognitive linguistics and neuroscience concerning the central role of image schemas and metaphor in abstract thinking.

In short, part 2 fleshes out some of my central claims about the body's role in conceptualization and reasoning by showing that there are neural mechanisms that might plausibly underlie and make possible the cognitive acts I have described in earlier chapters. Part 2 presents a challenge to the representational theory of mind by sketching key parts of an alternative view of cognition and meaning—one that is embodied, nondualistic, and naturalistic.

CHAPTER 6

# *The Origin of Meaning in Organism-Environment Coupling*

## A Nonrepresentational View of Mind

A naturalistic theory of mind, thought, and language must, as we have seen, explain conceptual thinking without introducing immaterial mind or a transcendent ego. According to the embodiment view I am developing here, meaning and thought emerge from our capacities for perception, object manipulation, and bodily movement. The chief challenge is to explain the phenomena of thought and symbolic interaction without resorting to a dualistic mind/body ontology that would violate Dewey's principle of continuity, insofar as it would deny continuity between so-called "bodily" processes and "mental" acts. What must be avoided, to paraphrase Hilary Putnam (1987), is the Kantian view that an adequate account of human mind and thought requires the keeping of two sets of books—one for the phenomenal world of things as appearances to us, and the other for the mysterious, noumenal world of things in themselves.

An embodied cognition view must avoid one of the most dangerous dualistic traps of Western philosophy, namely, asking how something inside the "mind" (i.e., ideas, thoughts, mathematical symbols) can represent the outside (i.e., the world). This trap is a consequence of the mistaken view that mind and body are two ontologically different kinds of entities, substances, or events. And this dualism then defines the problem of meaning as that of explaining how disembodied "internal" ideas can possibly represent "external" physical objects and events.

## THE REPRESENTATIONAL THEORY OF MIND

Mind/body dualism often generates what are known as representational theories of mind and cognition. Representationalism in its most general sense (as I am using the term here) is the view that cognition (i.e., perceiving, conceptualizing, imagining, reasoning, planning, willing) operates via internal mental "representations" (e.g., ideas, concepts, images, propositions) that are capable of being "about" or "directed to" other representations and to states of affairs in the external world. The technical term for this "aboutness" relation is *intentionality*. In commonsense language, we would simply say that we think with ideas and that our ideas can be about things in the world—things that are past, present, or future. To most people, this view of mind will seem self-evident. For example, if right now I'm thinking about fishing for rainbow trout, I must be entertaining ideas that have as their representational content *fishing* and *rainbow trout*. My idea of *fishing* picks out certain kinds of activities in the world that involve fishing rods, fishing lines, hooks, casting, and various other aspects of angling for fish. My idea of *rainbow trout* picks out certain members of the finny tribe that populate cold, freshwater streams and lakes. What could be more obvious?

Nevertheless, in spite of the widespread and long-standing appeal of the representational view of mind and its apparent obviousness, in its stronger versions it is an extremely problematic view of meaning and thought. Several centuries of struggling with the problem of how "internal" ideas can be related to "external" things should suggest that once you assume a radical mind/body dichotomy, there is no way to bridge the gap between the inner and the outer. When "mind" and "body" are regarded as two fundamentally different kinds of being, no third, mediating thing can exist that possesses both the metaphysical character of inner, mental things and simultaneously the character of the outer, physical things. Otherwise, we wouldn't even need a distinction between two radically different metaphysical kinds, since there would exist something (the mediating third) that could account for both what we attribute to inner mind and what we attribute to the external world (including our bodies).

At the end of this chapter, I will acknowledge some proper uses of the term *representation*, uses that support an extremely modest representational view. However, it is important to get clear about what is false and very misleading in strong representationalist theories that posit an inner world of ideas that get their meaning from their relation to external objects, events, and states of affairs. I am going to urge that whenever possible, we should avoid using "representation" talk, because it tends to foster the illusion of inner mental space populated by mental quasi-entities (such as concepts, propositions, and functions).



As an example of how easily we can be led astray, I am going to argue that the most popular contemporary version of the representational theory of mind—the very strong version articulated by Jerry Fodor and his supporters—is a false and misleading view of human cognition and meaning. Fodor is well known for his spirited defense of a technical version of what he calls folk psychology, which is, roughly, the commonsensical view that we can explain human thought and behavior by attributing mental states (such as beliefs and desires) to people. For example, I might say that the reason Scott is at this moment casting an artificial caddisfly on the surface of Otter Lake is that Scott *desires* to catch a trout and he *believes* (among other things) that there are rainbow trout in Otter Lake, that they often feed on caddisflies, and that an artificial caddisfly pattern is a decent mimic of an actual caddisfly on the surface of the water. Fodor's special version of this belief-desire psychology assumes the operation of what he calls a "language of thought" that provides the meaning of particular tokens of sentences in a person's natural language. Fodor describes his representational theory of mind as follows: "At the heart of the theory is the postulation of a language of thought: an infinite set of 'mental representations' which function both as the immediate objects of propositional attitudes and as the domains of mental processes" (Fodor 1987, 16–17). A propositional attitude is a mental state, such as belief or desire, directed toward a propositional content. For example, to think the thought (described in English) "The rainbows are rising to caddisflies" is to activate in one's language of thought a set of mental symbols or representations that constitute the appropriate proposition, *the rainbows are rising to caddisflies*. Now, it is an essential part of Fodor's view that the alleged mental representations that provide the structure and content of our thinking are not themselves intrinsically meaningful. Rather, they are symbols that can be related and arranged according to their syntactic (or formal) characteristics. Fodor explains:

> Mental states are relations between organisms and internal representations, and causally interrelated mental states succeed one another according to computational principles which apply formally *to the representations*. This is the sense in which internal representations provide the domains for such data processes as inform the mental life. It is, in short, the essence of cognitive theories that they seek to interpret physical (causal) transformations as trans-

> formations of information, with the effect of exhibiting the rationality of mental processes. (Fodor 1975, 198)

As Fodor sees them, mental processes operate like a computational program that performs formal operations on symbols. A major problem for this kind of view is to explain how this inner language of thought (or "mentalese," as he sometimes calls it) can get meaning by being related to things in the world. A simplified version of Fodor's answer is that symbols in the language of thought get their meanings by virtue of their law-like relations to events in the world that cause the appropriate symbols to be "tokened" in the language of thought. As Fodor says, "In such cases the symbol tokenings denote their causes, and the symbol types express the property whose instantiations reliably cause their tokenings. So, in the paradigm case, my utterance of 'horse' says *of* a horse that it *is* one" (Fodor 1987, 99).

William Bechtel, Adele Abrahamsen, and George Graham (1998) have concisely characterized Fodor's position as follows:

> To be a cognizer is to possess a system of syntactically structured symbols-in-the-head (mind/brain) which undergo processing that is sensitive to that structure. Cognition, in all of its forms, from the simplest perception of a physical stimulus to the most complex judgment concerning the grammaticality of an utterance, consists of manipulating innate symbols-in-the-head in accord with that syntax. The system of primitive, innate symbols-in-the-head and their syntactic combination in sentence-like structures is sometimes called "mentalese." (Bechtel, Abrahamsen, and Graham 1998, 63–64)

I am specifically addressing Fodor's theory for four reasons: First, it is the most influential contemporary version of the representational theory of mind. Second, it is a paradigmatic functionalist theory of mind, since it treats mind as a computational program that can be run on any suitable "hardware." Third, it thus regards the "hardware" of mind as relatively incidental to the content of thought. Fourth, it therefore denies any intrinsic meaning to mental states (such as meaning that might arise from the nature of human embodiment). In this and subsequent chapters, I do not intend to criticize Fodor's version of the representational theory directly.[1] Instead, I will attempt to elaborate an alternative, naturalistic account of mind, concepts, and meaning that is more compatible with the sciences of the mind and that does not presuppose supernatural or nonembodied entities or processes.



## EMBODIMENT THEORY'S CHALLENGE TO THE REPRESENTATIONAL THEORY OF MIND

Embodiment theory, in contrast to representationalist theories, requires a radical reevaluation of dualistic metaphysics and epistemology, and it challenges Fodor's representational view that cognition and thought consist of symbolic representations inside an organism's mind-brain that refer to an outside world. According to Dewey's principle of continuity, what we call "body" and "mind" are simply convenient abstractions—shorthand ways of identifying aspects of ongoing organism-environment interactions—and so cognition, thought, and symbolic interaction (such as language use) must be understood as arising from organic processes. I want to trace the rejection of mind/body dualism from the philosopher-psychologists known as the early American pragmatists (especially James and Dewey) forward through contemporary cognitive scientists (e.g., Francisco Varela, Humberto Maturana, Gerald Edelman, Edwin Hutchins, George Lakoff and Vittorio Gallese, and Rodney Brooks). The key to this reconceiving of mind is to stop treating percepts, concepts, propositions, and thoughts as quasi-objects (mental entities or abstract structures) and to see them instead as patterns of experiential interaction. They are aspects or dimensions or structures of the patterns of organism-environment coupling (or integrated interaction) that constitute experience. The only sense in which they are "inner" is that my thoughts are mine (and not yours), but they are not mental objects locked up in the theater of the mind, trying desperately to make contact with the outside world. As we will see, thoughts are just modes of interaction and action. They are *in* and *of* the world (rather than just being *about* the world) because they are processes of experience.

In the remainder of this chapter and in subsequent chapters, I argue for a nondualistic, nonrepresentational view of mind as a process of organism-environment interactions. My *interactionist* (or *transactional* or *enactionist*) view will provide the beginnings of an account of the bodily grounding of meaning and concepts in sensorimotor experience, including the role of

1. In a series of essays over several years, but especially in *Representation and Reality* (1988), Hilary Putnam has shown precisely why language of thought views and other forms of functionalism will not work. He has also shown why so-called causal theories of meaning and reference are misguided attempts to connect words to things.

neural maps.[2] In this chapter, I cite evidence from comparative neurobiology of organism-environment coupling for nonhuman animals; then, in chapters 7 and 8, I turn to parallels in human cognition, based on research in recent cognitive science, cognitive neuroscience, and computational neural modeling that supports a theory of embodied cognition and an interactionist view of mind. In chapter 9, I then suggest how this conception of embodied meaning makes it possible to explain even abstract concepts, concepts that have traditionally been thought to be completely disembodied.

## WHAT DIFFERENCE DOES EMBODIMENT THEORY MAKE?

When a young child crawls toward the fire in the hearth and a mother snatches up the child before it can get burned, is that cognition? When a team of British mathematicians decodes enemy ciphers during wartime, is that cognition? When ants carrying food back to their nest lay down chemical signals and thereby mark trails to a food source, is that cognition?

Note the commonalities among these situations. First and foremost, in each case the body (individual and social) is in peril. The well-being and continued successful functioning of various organisms are at risk. To survive and flourish, the organism must make adjustments in its way of acting, both within its current environment and in its relations with other creatures. The child must be snatched from the imminent danger of the flames, the mathematicians work desperately to prevent their country from being overrun by the enemy, and the ants must find food and bring it back to the queen in order for the colony to survive. Second, note that in each case the cognition is social, composed of multiple organisms acting cooperatively in response to problems posed by the current environment. And finally, all of these situations have been taken by theorists as emblematic of cognition par excellence (Dewey 1925/1981; Hodges 1983, 160–241; Deneubourg, Pasteels, and Verhaeghe 1983; Brooks and Flynn 1989).

2. I use the term *interactionist* with great trepidation, for it carries the misleading implication that there are two or more independent "things" that are interacting, whereas the view I am presenting treats the "things" within the interaction as just abstracted dimensions of the basic, continuous process of experience. Dewey used the word *transactional* precisely to avoid this dualistic suggestion, and Varela, Thompson, and Rosch (1991) use the nice term *enactionist* to stress the active, dynamic, directed process of experience.



Embodiment theory is now well supported by research in the cognitive sciences, yet there remains considerable debate as to what exactly the term *embodiment* might mean (Rohrer 2001a, forthcoming). Is the "body" merely a physical, causally determined entity? Is it a set of organic processes? Is it a felt experience of sensations and movement? Or is it a socially constructed artifact? I will suggest, in the final chapter of this book, that the body is *all* of these things. In previous chapters, we have seen how parts of the pragmatist view of thought help explain how meaning is grounded in our embodiment. We can now begin to sketch the broader nonrepresentational view of mind that emerges from this naturalistic perspective.

Embodiment theory shares several key tenets of the pragmatist view of cognition. Embodied cognition

- is the result of the evolutionary processes of variation and selection;
- is situated within a dynamic, ongoing organism-environment relationship;
- is problem-centered and operates relative to the needs, interests, and values of organisms;
- is not concerned with finding some allegedly perfect solution to a problem but, rather, one that works well enough relative to the current situation; and
- is often social and carried out cooperatively by more than one individual organism.

Pragmatists advance a view of cognition radically different from the one we are most familiar with from classical (or first-generation) cognitive science. For classical cognitive science, it is assumed that cognition consists of the application of universal logical and formal rules that govern the manipulation of "internal" mental symbols, symbols that are supposedly capable of representing states of affairs in the "external" world. Fodor's representational theory, which treats mind as a computational program, is an exemplary instance of such a first-generation view of cognitive science.

The internal/external split that underlies this view presupposes that "mentalese," construed as part of a functional program, could be detached from the nature and functioning of specific bodily organisms, from the environments they inhabit, and from the problems that provoke cognition. Given this view, it would follow that cognition could take place in any number of suitable media, such as a human brain or a computing machine. This theoretical viewpoint was instrumental in the development of the first electronic calculating machines and general-purpose computers.

In fact, these machines were originally developed by the British military to reduce the tedious workload of military mathematicians (or human "computers"—in the sense of humans who compute). But this thought experiment did not end merely with offloading the tedium of calculation onto electronic machines. From its original conception in the work of Alan Turing (1937), the idea of a universal computing machine became the metaphor of choice for future models of the mind and brain. For example, Allen Newell and Herbert Simon (1976), in their conception of the brain as a physical symbol system, consider the human brain to be just a specific instance of a Turing-style universal machine. In short, *for classical, first-generation cognitive science, cognition is defined narrowly as mathematical and logical computation with intrinsically meaningless internal symbols that can supposedly be placed in relation to aspects of the external world.*

The pragmatist alternative to this classical cognitive science view of mind is to argue that *cognition is action*, rather than mental mirroring of an external reality. Cognition is a particular kind of action: a response strategy that involves both nonconscious processes and occasional conscious processes that apply some measure of forethought in order to solve some practical, real-world problem. In the 1940s, during World War II, the problem of breaking the German codes was of utmost importance to the British war effort, and this led to the development of a series of machines (the Bombes) that could try a vast number of possible cipher keys against intercepted German communications. These decoding machines were among the predecessors of the modern computer. Early computers were designed to model human action—*computing* possible cipher keys—so that machines would replace human labor (Hodges 1983, 160–241).

However, this success in the modeling of a very specific intellectual operation was soon mistakenly regarded as the key to understanding cognition in general. If one thinks that mathematical and logical reasoning are what distinguish human beings from other animals, one might erroneously assume that any computational machine that could model aspects of this peculiarly human trait could also be used to model cognition in general. Hence the MIND AS COMPUTER metaphor swept early (first-generation) cognitive science.[3]

By contrast, on the pragmatist view, we human beings *are* animals. Our rationality is not something apart from our animal bodies, but instead

3. For a conceptual analysis of the MIND AS COMPUTER metaphor and some of its fateful implications for contemporary analytic philosophy, see Lakoff and Johnson 1999, chapters 12 and 21.



emerges from, and is shaped by, our embodied engagement with our environment. Thus, Dewey famously asserted that "to see the organism *in* nature, the nervous system in the organism, the brain in the nervous system, the cortex in the brain is the answer to the problems which haunt philosophy" (Dewey 1925/1981, 198).

## THE CONTINUITY OF EMBODIED EXPERIENCE AND THOUGHT: JAMES AND DEWEY

We have already seen some of the ways that James and Dewey provide us with exemplary nonreductionist and nonrepresentational models of embodied mind. Their models combined the best biological and cognitive science of their day with nuanced phenomenological description and a commitment that philosophy should address the pressing human problems of our lives. I contend that James and Dewey thus provide us today with ideals of a philosophy that maintains a constructive dialogue with the sciences that can guide us in how to live. The fundamental assumption of the pragmatists' naturalistic approach is that everything we attribute to "mind"—perceiving, conceptualizing, imagining, reasoning, desiring, willing, dreaming—has emerged (and continues to develop) as part of a process in which an organism seeks to survive, grow, and flourish within different kinds of environments. As James puts it:

> Mental facts cannot be properly studied apart from the physical environment of which they take cognizance. The great fault of the older rational psychology was to set up the soul as an absolute spiritual being with certain faculties of its own by which the several activities of remembering, imagining, reasoning, and willing, etc. were explained, almost without reference to the peculiarities of the world with which these activities deal. But the richer insight of modern days perceives that our inner faculties are *adapted* in advance to the features of the world in which we dwell, adapted, I mean, so as to secure our safety and prosperity in its midst. (James 1900, 3)

This evolutionary embeddedness of the organism within its changing environments, and the development of thought in response to such changes, ties mind inextricably to body and environment. The changes entailed by such a view are revolutionary, relative to classical dualistic views of mind and thought. From the very beginning of life, the problem of knowledge is *not* how so-called internal ideas can re-present external realities, because the mind was never separate from its environment in the first place. In Dewey's words, "Since both the inanimate and the human environment are involved in the functions of life, it is inevitable, if these functions evolve to the point of thinking and if thinking is naturally serial with biological functions, that it will have as the material of thought, even of its erratic imaginings, the events and connections of this environment" (Dewey 1925/1981, 212–13).

Another way of expressing this rootedness of thinking in bodily experience is to say that there is no ontological rupture in experience between perceiving, feeling, and thinking. More complex levels of organic functioning are just that—levels, and nothing more, although there are emergent properties of "higher" levels of functioning. Dewey describes the connectedness of all levels of cognition via his principle of continuity, according to which "rational operations *grow out of* organic activities, without being identical with that from which they emerge" (Dewey 1938/1991, 26).

The notion of continuity is fundamental for Dewey, because it is the key to avoiding ontological and epistemological dualisms. Dewey employs at least two different senses of the term. *Higher-lower continuity* is the twofold claim that "higher" organisms are not the result of some additional ontological *kind* emerging in the history of the world, and also that our "higher" self (reason, will) is not utterly different in kind from our "lower" self (perception, emotion, imagination). *Inner-outer continuity* is the denial that what is inner (e.g., mental) needs ontological principles for its explanation that are different from those used to explain the outer (e.g., the physical). The continuity thesis is the basis of Dewey's claim that we can provide a naturalistic explanation of events in our world without reference to the alleged activities and powers of supernatural agents or causes. Dewey's views on continuity thus collide head-on with religious and moral traditions that posit disembodied souls and spirits or that regard "mind" as nonnatural.

The principle of continuity entails that any explanation of the nature and workings of mind, even the most abstract conceptualization and reasoning, must have its roots in the embodied capacities of the organism for perception, feeling, object manipulation, and bodily movement. The continuity hypothesis, however, does *not* entail that there are no demarcations, differentiations, or distinctions within experience. Of course there are demarcations, and they are real and important! The continuity hypothesis insists only that wherever and whenever we find actual working distinctions, they are explicable against the background of continuous processes. Furthermore, social and cultural forces are required to develop our cognitive capacities to their full potential, including language and symbolic



reasoning. Infants do not speak or discover mathematical proofs at birth. Dewey's continuity thesis thus requires both evolutionary and developmental explanations. For James and Dewey, this meant that a full-fledged theory of human cognition must have at least three major components:

(1) There must be an account of the connections between humans and other animals as regards the emergence and development of meaningful patterns of organism-environment interactions—patterns of sensorimotor experience shared by all organisms of a certain kind and meaningful for those organisms. Such patterns must be tied to the organism's attempts to function within its environment. The continuity here is between what we call "lower" and "higher" organisms. (This is the subject of the present chapter.)

(2) There must be an account of how we can perform abstract thinking using our capacities for perception and motor response. There would need to be bodily processes for extending sensorimotor concepts and logic for use in abstract reasoning, as well as an account of how the processes embodying such abstract-reasoning capacities are learned during an organism's development. This story has at least two parts: (a) an evolutionary and physiological account explaining how an adult human's abstract reasoning utilizes the brain's perceptual and motor systems; and (b) a developmental and anthropological account of how social and cultural behaviors educate the sensorimotor systems of successive generations of children so that they may communicate and perform abstract reasoning. The continuity here pertains to both higher/lower and inner/outer.[4]

(3) Because judgments of value are essential to an organism's continued functioning, there must be an account of the central role of emotions and feelings in the constitution of an organism's world and its knowledge of it. Again, this will involve a continuity between the "higher" (rational) and the "lower" (emotional) aspects of the self, and also a continuity between the inner (associated with disembodied reason) and the outer (associated with the emotional body).

## ORGANISM-ENVIRONMENT COUPLING

Dewey's principle of continuity states that there are no ontological gaps between the different levels of functioning within an organism. One way to see what this entails is to survey a few representative types of organism-

4. In the present chapter and in chapters 7 and 8, I can only address part (a) of this complex issue.

environment couplings, starting with single-celled organisms and moving up to more complex animals and eventually to humans. In every case, we can observe the same adaptive process of interactive coordination between a specific organism and recurring characteristics of its environment.

But does that mean that we can trace human cognition all the way back to the sensorimotor behavior of single-celled organisms? On the face of it, this seems preposterous; and indeed, from an evolutionary biologist's perspective, there are clear differences in the size, complexity, and structural differentiation of human beings as compared with single-celled organisms such as bacteria. The behavior of single-celled organisms is not ordinarily relevant to the behavior of multicellular organisms, except insofar as there might be structural, morphological analogies between the sensorimotor activity of single-celled organisms and particular sensorimotoric cells within multicellular bodies.

## From Chemotaxis to Cognition

In fact, just this sort of morphological analogy plays a key role in the argument by Humberto Maturana and Francisco Varela that central nervous systems evolved in multicellular organisms to coordinate sensorimotor activity (Maturana and Varela 1998, 142–63). In a single-celled organism, locomotion is achieved by dynamically coupling the sensory and motor surfaces of the cell membrane. When an amoeba engulfs a protozoan, its cell membranes are responding to the presence of the chemical substances that make up the protozoan, causing changes in the consistency of the amoeba's protoplasm. These changes manifest as pseudopods—digitations that the amoeba appears to extend around the protozoan as it prepares to feed upon it. Similarly, certain bacteria have a tail-like membrane structure called a flagellum that is rotated like a propeller to move the bacterium. When the flagellum is rotated in one direction, the bacterium simply tumbles; reversing the direction of rotation causes the bacterium to move. If a grain of sugar is placed into the solution containing this bacterium, chemical receptors on the cell membrane sense the sugar molecules. This causes a membrane change in which the bacterium changes the direction of rotation of its flagellar propeller and gradually moves toward the greatest concentration of the sugar molecules, a process known as chemotaxis. For both the amoeba and the bacterium, changes in the chemical environment cause sensory perturbations in the cellular membrane, which invariably produces movement. The key point here is that *without any awareness of anything like an internal representation,* single-celled organisms engage in sensorimotor coordination in response to environmental changes. Even at this apparently primitive level, there is a finely tuned, ongoing coupling of organism and environment.



Multicellular organisms also accomplish their sensorimotor coordination by means of changes in cellular membranes. However, the cellular specialization afforded by multicellularity means that not every cell needs to perform the same functions. Maturana and Varela (1998) discuss the example of an evolutionarily ancient metazoic organism called the hydra (a coelenterate). The hydra, which lives in ponds, is shaped like a two-layered tube with four or six tentacles emanating from its mouth. On the inside layer of the tube, most cells secrete digestive fluids, while the outside layer is partly composed of radial and longitudinal muscle cells. Locomotion is accomplished by contracting muscle cells along the body of the organism, some of which cause changes in the hydrostatic pressure within the organism, changing its shape and its direction of locomotion.

Between the two layers of cells, however, are specialized cells—neurons—with elongated membranes that can extend over the length of the entire organism before terminating in the muscle cells. These tail-like cellular projections are the axons, and evolutionarily speaking they are the flagella of the multicellular organism.[5] Changes in the electrochemical state of other, smaller cellular projections (known as dendrites) of the neurons cause larger changes in the electrochemical state of the axonal membrane, which in turn induces the muscle cells to contract. These neural signals typically originate in either the tentacles or the "stomach" of the hydra, such that these structures' electrochemical state responds to molecules indicating the presence or absence of food and/or excessive digestive secretions. The neurons consistently terminate in the longitudinal and radial muscles that contract the hydra body for locomotion or for swallowing. The topology of how the nerve cells interconnect is crucial: when a tentacle is touched, a chain of neurons fires sequentially down the tentacle toward the hydra's mouth and causes the muscle cells to curl the tentacle about the prey. The structural coupling between the hydra (with its changes in electrochemical states) and its environment (e.g., prey that touches its tentacle) is what allows the hydra to function adaptively in its environment—this allows the hydra to contract the correct muscles to swallow or to move up or down, right or left.

5. Recent research shows that this may be more than a surface morphological analogy: all microtubular cellular projections stem from a common ancestor (Erickson et al. 1996; Goldberg 2003).

It is clear that the hydra does not experience internal representations of an external world that it could use as a basis for operating within its environment. A protozoan swims into the grasp of its tentacles and the neurons fire as they have been wired to do, stimulating adjacent muscle cells so that the tentacles curl about the food and bring it to a mouth that is simultaneously preparing to swallow. Now, it is obviously quite a jump from the hydra to humans, but Dewey's principle of continuity requires an account of human cognition that is not different in kind—though it would certainly differ massively in complexity—from the story of organism-environment coupling that applies to the simpler multicellular animals. In general, "higher" human cognition is evolutionarily continuous with sensorimotor coordination, because all nervous systems "couple the sensory and motor surfaces through a network of neurons whose pattern can be quite varied" (Maturana and Varela 1998, 159). At the human level, cognition is action—we think in order to act, and we act as *part of* our thinking.

Here the skeptic will interrupt my account of organism-environment coupling with an obvious objection: while these reflexive behaviors are examples of patterns of meaningful sensorimotor activation for coelenterate organisms, what would constitute similar patterns for humans? The whole point, they will say, is that it is precisely the human capacity for abstract representational thinking that distinguishes us from the lower animals. There is no continuum here, but rather a radical ontological gap that separates humans from all other creatures. Surely human cognition is something more complex, and of a different order, than what occurs in the simple reflex of a hydra.

## From Neural Maps to Neural Plasticity

Obviously, human cognition involves orders of complexity that far exceed those of primitive organisms. But even though human cognition is a little more like what happens in frogs, owls, and monkeys, it is nevertheless surprisingly continuous in important ways with that of the hydra functioning within its environment. Frogs, for example, must couple with their environment via primitive reflexes. They have a certain regularly occurring life-maintenance problem: they need to extend their tongues to eat flies. This was the subject of a classic experiment in the history of neurobiology (Sperry 1943). When a frog is still a tadpole, it is possible to rotate its eye 180 degrees (making sure to keep the optic nerve intact). The tadpole is then allowed to develop normally into a frog. When a frog whose eye was rotated goes for a fly, its tongue extends to the point in its visual field that

[redacted]



is exactly opposite the point where the fly is located. No amount of failure will teach the frog to move its tongue differently; the nervous system acts entirely on the basis of the connections between the retinal image and the tongue muscles. Maturana and Varela conclude that for frogs, "there is no such thing as up or down, front and back, in reference to an outside world, as it exists for the observer doing the study" (1998, 125–26). The frog has no access to our notion of the external world and our 180-degree rotation of its eye; it has only an environmentally induced change of state in the neurons constituting its retinal map.

One of the most profound findings in neuroscience is that nervous systems exploit topological and topographic organization. In other words, organisms build neural "maps." In neural maps, adjacent neural cells (or small groups of neural cells) fire sequentially when a stimulus moves across adjacent positions within a sensory field. For example, scientists have manipulated the visual field of the frog and measured the electrical activity of a region of its brain to show that as one stimulates a frog's visual field, the neurons of its optic tectum will fire in coordination with the visual stimulus. Scott Fraser (1985) covered the frog's optic tectum with a grid of twenty-four electrodes, each one recording electrical activity that was the sum of the signals from a receptive field containing many optic nerve fiber terminals. When a point of light was moved in a straight line in the frog's right visual field, from right to left and then from bottom to top, the electrode grid recorded neuronal activity in straight lines with sequential firing, first from the rostral (front) to the caudal (back) areas and then from the lateral to the medial areas. We call this pattern of activation the frog's retinal map, or retinotectal map, because it encodes environmental visual stimuli in a topographically consistent manner. The spatial orientation of this topography is rotated in various ways; thus, visual right-to-left movement becomes front-to-back in the retinotectal map, and so on. But the topographic mapping between movement in the vertical visual plane and the plane of the retinotectal neural map remains consistent. Even though there is considerable spatial distortion in the neural map, the key relational structures are preserved. In some other cases, such as some auditory maps and color maps, the correspondences are less about shape and position, and the organization is more properly called topologic than topographic; but the organizing principle of the neural mapping of sensation still holds. In the auditory regions of human brains, for example, successive rising pitches activate contiguous areas in the auditory cortex, so that "adjacent" pitches (such as B going to A) activate regions of the auditory cortex immediately next to one another.

The degree to which such neural maps might be plastic and subject to reorganization has been the subject of much recent study. It is important to remember that in the case of the frog, Sperry performed a radical and destructive intervention that is outside the realm of "normal" Darwinian deviation. In other words, if this were to occur by natural selection, such a frog would be unable to catch flies and would die quickly, without passing on its genes. Interventions that are more subtle and perhaps more likely to occur in nature, such as cutting the optic nerve of a goldfish and destroying part of the optic tectum, result in a recovery of function; the goldfish's optic-nerve axons will regenerate and make a complete retinal map in the remaining part of the tectum (Gaze and Sharma 1970).

Or consider another, even more subtle intervention: suppose we were to make a barn owl wear glasses that changed its perception of its visual field? Like frogs, owls have developed an extremely accurate method of attacking prey. The owl hears a mouse rustling on the ground and locates the mouse primarily by using the tiny difference in time that it takes for a sound to reach one ear as opposed to the other. This establishes the mouse's approximate position in the owl's retinotectal map, and the diving owl then looks to find the exact location of its prey as it strikes. Eric Knudsen (2002) put prismatic glasses on adult and juvenile owls that distorted the owls' vision by twenty-three degrees. After wearing the glasses for eight weeks, the adult owls never learned to compensate, although juveniles were able to learn to hunt accurately. However, when the glasses were reintroduced to adult owls who had worn them as juveniles, they were able to readjust to the glasses in short order.

This behavioral plasticity has anatomical underpinnings in the organization of neural maps. When the experimental owls were injected with an anatomical tracing dye, comparison of the neural arbors—the patterns of neural connections—from normally reared and prism-reared owls revealed greatly differing patterns of axonal projections between auditory and spatial neural maps, "showing that alternative learned and normal circuits can coexist in this network" (Knudsen 2002, 325). In other words, in order to deal with wearing glasses, the owl brain had grown permanent alternative axonal connections in a cross-modal neural map of space located in the external nucleus of the inferior colliculus (ICX). The ICX neural arbor of prism-reared owls is significantly more dense than in owls that developed normally, with neurons typically having at least two distinct branches of axons (DeBello, Feldman, and Knudsen 2001). By contrast, the retinotectal maps of the visual modality alone do not seem to exhibit quite the same plasticity, neither in owls (whose retinotectum did not change) nor in frogs. Anatomical research on frogs reared and kept alive with surgically rotated eyes has shown that after five weeks, the retinotectal neural arbors exhibited a similar pattern of "two-headed" axons—that is, they had two major axonal branches. However, after ten weeks, the older axonal connections started to decay and disappear, and after sixteen weeks no two-headed axons could be traced (Guo and Udin 2000). Apparently, the frog's single-modal retinotectal maps do not receive enough reentrant neural connections from other sensory modalities to sustain the multiple branching neural arbors found in the cross-modal map of an owl who develops normally at first and subsequently with glasses.



Working on adult squirrel and owl monkeys, Michael Merzenich and colleagues (1987; reviewed in Buonomano and Merzenich 1998) have shown that it is possible to dynamically reorganize the somatosensory cortical maps subject to certain bodily constraints. Similar to the owls and frogs that grew dual arborizations, these monkeys exhibited a plasticity based on their brains' ability to select which parts of their neural arbors to use for what kinds of input. In a series of studies, Merzenich and colleagues altered the monkey's hand sensory activity by such interventions as (1) cutting a peripheral nerve, such as the medial or radial nerve, and (1a) allowing it to regenerate naturally or (1b) tying it off to prevent regeneration; (2) amputating a single digit; and (3) taping together two digits so that they could not be moved independently. The results show that cortical areas now lacking their previous sensory connections (or independent sensory input, in the third condition) were "colonized" in a couple of weeks by adjacent neural maps with active sensory connections. In other words, the degree of existing but somewhat dormant neural-arbor overlap was large enough that the cortex was able to reorganize. And in the case of (1a), where the nerve was allowed to regenerate, the somatosensory map gradually returned to occupy a similar-sized stretch of cortex, albeit with slightly different and filtering boundaries. These experiments illustrate that learning in adults is accomplished in part by neural gating—sequencing of neural firings—between redundant and overlapping neural arbors.

All of these examples of ontogenetic neural change suggest that there is a process of neural-arbor selection akin to natural selection taking place in concert with specific patterns of organism-environment interactions. On precisely these grounds, the neurobiologist Gerald Edelman (1987) has proposed a theory of "neural Darwinism," or "neuronal group selection," to explain how such neural maps are formed in the organism's embryonic development. Different groups of neurons compete to become topobiological neural maps as they migrate and grow during neural development. Successful cortical groups, driven primarily by regularities in the environment passed on from those neurons that are closer to the sensory apparatus, will fire together and wire together in a process of axonal sprouting and synaptogenesis. Some neuronal groups will fail to find useful topological connections, and they eventually die and are crowded out by the successful neuronal groups, while others will hang on in something of an intermediate state of success (Edelman 1987, 127–40). In the adult organism, these latent axonal arbors remaining from only partly successful attempts to wire together lie dormant, ready to reorganize the map as needed by means of further synaptogenesis. Edelman (1987, 43–47) calls these latent reorganizations of the neuronal groups *secondary repertoires,* as distinguished from their normal *primary repertoires.*

Like frogs, owls, and monkeys, humans have sets of visual, auditory, and somatosensory maps. The more obvious of these map perceptual space in fairly direct analogs—preserving topologies of pitch, the retinal field, color, the parts of the body, and so on. But subsequent maps preserve increasingly abstract topological structure (or even combinations of structure), such as object shape, edges, orientation, direction of motion, and even the particular degree of the vertical or horizontal. *Like the frog, we live in the world significantly (but not totally) defined by our maps. Topologically speaking, our bodies are in our minds.* Our "minds" are processes that arise through our ongoing coupling with our environment. Mind is in and of this embodied experiential process, not above it all.

## ARE NEURAL MAPS INTERNAL REPRESENTATIONS? NO, AND YES!

Some people might suppose that talk of neural maps would necessarily engender representational theories of cognition. On this view, the map would be construed as an internal representation of some external reality. But the account given above *does not* entail any of the traditional metaphysical dualisms that underlie representational views—strict dichotomies such as subject/object and mind/body. Such dichotomies might describe aspects of organism-environment interactions, but they do not indicate ontologically different entities or structures. According to our interactionist view, maps and other structures of organism-environment coordination are prime examples of nonrepresentational structures of meaning, understanding, and thought. Maturana and Varela (1998, 125–26) remind us that we must not read our scientific or philosophical perspectives



(i.e., our theoretical stance) on cognition back into the experience itself that we are theorizing about. William James called this error the "Psychologist's Fallacy," namely, "the *confusion of his own standpoint with that of the mental fact* about which he is making his report" (1890/1950, 1:196). In observing something scientifically, one must always consider the standpoint of the scientist in relation to the object of study. When scientists use terms such as *retinal maps, pitch maps, sensorimotor maps,* and *color maps* to describe the operations of various neural arrays in a frog's or a human's nervous system, they do so from their standpoint as observers and theorists who can see mappings and isomorphisms between the neural patterns and their own experience of the "external world." But for the engaged frog, and for the human in the act of perceiving, that map *is* the external world, or at least the topological structure of the world as currently experienced! It is not some internal representation of that world. The reason is that if you took the neural map out of the situation, that situation would not exist as it now does—it would be a different situation. The frog's neural map has its origin not in the immediate mappings that we observers see in the moment, but in a longitudinal evolutionary and developmental process during which those neural connections were "selected for" by Darwinian or neo-Darwinian mechanisms.

In short, what we (as scientists) theoretically recognize and describe as an organism's maps are not—*for that organism*—internal representations that can correspond to external realities. Rather, what we call sensorimotor and somatosensory maps (whether in multicellular organisms, monkeys, or humans) are *for that organism* precisely its structures of experience and reality. Consequently, we must be careful not to be misled by philosophers of mind and language who, in speaking of the intentionality (the "aboutness") of these maps and other neural structures, surreptitiously introduce an inner/outer split that does not exist in reality for the organism. What must be avoided is the illusory setting up of an inner theater of the mind, in which immaterial ideas parade on the stage of consciousness, to be seen by the mind's eye.

I must acknowledge at least one sense in which neural maps—a certain neuronal cluster that is activated as an organism perceives and moves within its environment—may properly be called a representation, which is why that term is widely (though uncritically) used by neuroscientists. Just insofar as a specific neural map is loosely isomorphic with some structure of an organism's environment-as-experienced, neuroscientists are prone to call these maps representations. Something "in the brain" appears to correlate structurally with patterns of the "external environment," even though the environment is not, strictly speaking, independent of the organism. So, from the point of view of the scientist (but *not* from the perspective of the experiencing organism), the neural map represents some structure in the world. It is in this sense that a neuroscientist might say that adjacent neural assemblies in parts of the auditory cortex "represent" adjacent tonal relations. However, we must always be clear that an organism never actually experiences its neural maps as internal mental structures. We do not experience the *maps*, but rather *through them* we experience a structured world full of patterns and qualities.

In challenging the representational theory of mind, I am in no way denying that human beings have images, feelings, and thoughts that they experience as *theirs* and that they understand to be "about" something in their world. But it is quite a leap from such a claim to the very much stronger claim that thought is the mental entertaining and manipulation of internal representations. What I am denying is that we have mental *entities* called "concepts" or "representations" in our "minds" and that thinking is a matter of manipulating these entities by surveying their properties, discerning their relations to each other and to mind-external objects, and arranging them in internal acts of judgment. To say that we have concepts is to say, with James, that within the continuous flow of our perceptual experience, we can attend to aspects of the flow for purposes of understanding our situation, planning what to do, and then acting. Concepts, on this view, are stable patterns of neuronal activation, but they are *not* quasi-entities. Of course, we have the ability to abstract aspects of our experience and then to consider how these so-called abstract concepts relate to one another. There is nothing wrong with this way of talking, unless it leads us to suppose that there must be some center of thought, conceptualization, reasoning, or consciousness that "sees" or "grasps" (i.e., understands) the concepts and compares them to one another, as if we were holding up one juicy red tomato and comparing it to another tomato.

## IN WHAT SENSES ARE THERE MENTAL REPRESENTATIONS?

The classical representationalist picture of thought and how the mind works is so powerful and so deeply rooted in our self-understanding that it is hardly likely to ever be dislodged, either in how we think about thought or in how we talk about our thinking. Nevertheless, it is a mistaken picture that is not supported by what we are learning about how brains work and what cognition and conceptualization are. The empirical evidence from neuroscience suggests that there is no single center of mind and thinking and that scientifically, conceptualization has to depend on activated neural connections. I want to be clear that I am, of course, not denying that we conceptualize; I am only denying that we do so by manipulating mental entities that have the remarkable capacity to be "about" external things. It is also perfectly acceptable to say that we have "mental states," just as long as we don't define these by means of the representational model.



I am not recommending that we launch a massive campaign to expunge the term "mental representation" from our vocabularies. Such a campaign would be an impossible undertaking, doomed to failure. There are some perfectly acceptable uses of the word *representation,* and here is a short list of those uses.

1. *Patterns of sensorimotor neural activation.* As we saw in the previous section, scientists often identify recurring patterns of neural activation that correlate with what the scientist perceives as structural features of the organism's environment. The scientist is likely to call these neural patterns representations. The neural maps just discussed, and also images and image schemas (to be discussed in the next chapter), are examples of representations in this limited sense.

2. *Conceptual structures.* We have described concepts not as abstract, internal mental entities, but as selective discriminations from the ongoing, continuous flow of our experience. In chapter 8, we will see how there is no discontinuity between concrete and abstract concepts. As long as we understand concepts in this nondualistic manner, we can think of them as representations of various aspects of our experience. However, I am inclined always to avoid calling concepts "representations," simply because that term too easily activates a cognitive frame with all of the classical representationalist architecture.

3. *External systems of symbolic interaction.* It is perfectly acceptable to say that linguistic signs and other symbols (which can be regarded as either objects or events in the world) can represent various things and events. The word *dog* can be said to pick out or represent dogs. A photograph of Grace Kelly can be said to represent Princess Grace. A topological map of Oregon can be said to represent the topography of Oregon. An elaborate coronation ritual can be said to represent the grandeur, authority, and royalty of the newly crowned king. Picasso's painting *Guernica* can be said to represent the horror and terror of some incident in 1937 during the Spanish civil war. Even though these are all distinct senses of the term *represent,* and they need to be carefully distinguished and analyzed to identify the key differences, they are appropriate uses.

4. *Theoretical models.* In the sciences, mathematics, and philosophy, we develop theoretical models to help us explain natural and human phenomena. By virtue of structural isomorphism, analogies, and propositional models, we attempt to represent entities, events, states, relations, and processes. Such models are representations of what they purport to be about.

In sum, people are never going to stop using the term *representation*. Nor should they. The term should be used when it is appropriate, which means whenever it does not activate any strong form of the representational theory of mind. The representational theory is incompatible with cognitive neuroscience and out of touch with evolutionary accounts of mind, thought, and language. It is philosophically problematic, because it reinforces a set of ontological and epistemological dualisms that make it impossible to explain meaning, understanding, knowledge, and values without relying on supernatural, or at least transcendent, realities. In our theories of meaning and mind, we must exercise due caution to avoid falling back into any form of the representational view of mind.

CHAPTER 7

# *The Corporeal Roots of Symbolic Meaning*

In the previous chapter, I described patterns of organism-environment coupling mostly for nonhuman animals and argued that a strong representational theory of mind is of no use in explaining how such interactions work. For nonhuman animals, meaning is fully embodied. However, for human animals as well, meanings arise from organism-environment interactions, and we too have neural maps. The structural features and relations that shape our encounters with aspects of our environment are preserved in our neural maps.

In general, every aspect of our spatial experience will be defined by recurring patterns and structures (such as up-down, front-back, near-far, in-out, on-under) that constitute the basic contours of our lived world. It should not be surprising, therefore, that we have evolved to take special notice of these recurring shapes, relations, and patterns, and that these patterns exist as topological features of our neural maps. Such patterns are the structural elements of our ongoing engagement with our environment. They are one of the primary ways we are in touch with our world, understand it, and can act within it.

Since the earliest episodes of ancient Greek philosophy, humans have been inclined to distinguish themselves from "brute" animals and all lower organisms by their supposedly unique capacities for abstract conceptualization and reasoning that are typically associated with the possession of language. According to this view, human reason is a unique capacity having a different source than our capacities for perception, motor activities, feeling, and emotion. Therefore, the problem for an embodied view of