1 The Enactive Approach to Perception: An Introduction

The theory of the body is already a theory of perception.
—M. Merleau-Ponty

1.1 The Basic Idea

The main idea of this book is that perceiving is a way of acting. Perception is not something that happens to us, or in us. It is something we do. Think of a blind person tap-tapping his or her way around a cluttered space, perceiving that space by touch, not all at once, but through time, by skillful probing and movement. This is, or at least ought to be, our paradigm of what perceiving is. The world makes itself available to the perceiver through physical movement and interaction. In this book I argue that all perception is touch-like in this way: Perceptual experience acquires content thanks to our possession of bodily skills. *What we perceive* is determined by *what we do* (or what we know how to do); it is determined by what we are *ready* to do. In ways I try to make precise, we *enact* our perceptual experience; we act it out.

To be a perceiver is to understand, implicitly, the effects of movement on sensory stimulation. Examples are ready to hand. An object looms larger in the visual field as we approach it, and its profile deforms as we move about it. A sound grows louder as we move nearer to its source. Movements of the hand over the surface of an object give rise to shifting sensations. As perceivers we are masters of this sort of pattern of sensorimotor dependence. This mastery shows itself in the thoughtless automaticity with which we move our eyes, head and body in taking in what is around us. We spontaneously crane our necks, peer, squint, reach for our glasses, or draw near to get a better look (or better to handle, sniff, lick or listen to what interests us). The central claim of what I call *the enactive approach* is that our ability to perceive not only depends on, but is constituted by, our possession of this sort of sensorimotor knowledge.[1]

One implication of the enactive approach is that only a creature with certain kinds of bodily skills—for example, a basic familiarity with the sensory effects of eye or hand movements, and so forth—could be a perceiver.[2] This is because, in effect, perceiving is a kind of skillful bodily activity. It may also be that only a creature capable of at least some primitive forms of perception could be capable of self-movement. Specifically, self-movement depends on perceptual modes of self-awareness, for example, proprioception and also 'perspectival self-consciousness' (i.e., the ability to keep track of one's relation to the world around one).[3]

A second implication of the enactive approach is that we ought to reject the idea—widespread in both philosophy and science—that perception is a process *in the brain* whereby the perceptual system constructs an *internal representation* of the world. No doubt perception depends on what takes place in the brain, and very likely there are internal representations in the brain (e.g., content-bearing internal states). What perception is, however, is not a process in the brain, but a kind of skillful activity on the part of the animal as a whole. The enactive view challenges neuroscience to devise new ways of understanding the neural basis of perception and consciousness.[4] I return to this controversial topic in chapter 7.

This idea of perception as a species of skillful bodily activity is deeply counterintuitive. It goes against many of our preconceptions about the nature of perception. We tend, when thinking about perception, to make vision, not touch, our paradigm, and we tend to think of vision on a photographic model. You open your eyes and you are given, at once, a sharply focused impression of the present world in all its detail. On this view, the relation between moving and perceiving is only instrumental. It is like the relation between the lugging around of a camera and the resulting picture. The lugging is preliminary to and disconnected from the photograph itself. And so with perceiving. By moving yourself, you can come to occupy a vantage point from which, say, better to see your goal. And then, having seen your goal, you can better decide what to do. But the seeing, and the moving, have no more to do with each other than the photograph and the schlepping of the camera, or the boxer's left hook, and the training that preceeded it. Which is to say, they have a lot to do with each other, but the relation is nonconstitutive: The effectiveness of the punch is strictly independent of how the boxer learned to do it, and the qualities of the picture are independent of how the camera ended up where it was.

Susan Hurley (1998) has aptly called this simple view of the relation between perception and action the input-output picture: Perception is input from world to mind, action is output from mind to world, thought is the mediating process. If the input-output picture is right, then it must be possible, at least in principle, to disassociate capacities for perception, action, and thought. The main claim of this book is that such a divorce is not possible. I doubt that it is even truly conceivable. All perception, I argue, is intrinsically active. Perceptual experience acquires content thanks to the perceiver's skillful activity. I also argue—but I don't turn to this until late in the book (chapter 6)—that all perception is intrinsically thoughtful. Blind creatures may be capable of thought, but thoughtless creatures could never be capable of sight, or of any genuine content-bearing perceptual experience.[5] Perception and perceptual consciousness are types of thoughtful, knowledgeable activity.

My aim in this initial chapter is to set out the book's central themes.

1.2 A Puzzle about Perception: Experiential Blindness

For those who see, it is difficult to resist the idea that being blind is like being in the dark. When we think of blindness this way, we imagine it as a state of blackness, absence and deprivation. We suppose that there is a gigantic hole in the consciousness of a blind person, a permanent feeling of incompleteness. Where there could be light, there is no light.

This is a false picture of the nature of blindness. The longterm blind do not experience blindness as a disruption or an absence. This is not because, as legend has it, smell, touch and hearing get stronger to compensate for the failure to see (although this may be true to some degree; see Kaufman, Théoret, and Pascual-Leone 2002). It's because there is a way in which the blind do not experience their blindness at all. Consider, you are unable visually to discern what takes place in the room next door, but you do not experience this inability as a gaping hole in your visual awareness. Likewise, you don't encounter the absence of the sort of olfactory information that would be present to a bloodhound as something missing in your sense of smell. Nor do you notice the absence of information about the part of the visual field that falls on the "blind spot" of your retina. In this same way the blind do not encounter their blindness as an absence.

It is easy to demonstrate that there are or could be forms of blindness that were not at all like being in the dark. Imagine that you are out in a fog so dense that no matter where you turn or how you strain you only experience a homogeneous whiteness. This is what psychologists call a *Ganzfeld* (Metzger 1930, described in Gibson 1979, 150–151). You can reproduce the experience of a Ganzfeld by placing half a Ping-Pong ball over each eye (Hochberg, Triebel, and Seaman 1951; Gibson and Wadell 1952; Block 2001). Gibson used this method to argue that stimulation of the retina by light is not sufficient for vision. For even though you enjoy a pattern of visual stimulation—in some sense, you see the Ganzfeld—you are in effect blind. You have visual impressions, but they are bleached of content.

The enactive view of perception predicts that there are, broadly speaking, two different kinds of blindness. First, there is blindness due to damage or disruption of the sensitive apparatus. This is the familiar sort of blindness. It would include blindness caused by cataracts, by retinal disease or injury, or by brain lesion in the visual cortex. Second, there is blindness due not to the absence of sensation or sensitivity, but rather to the person's (or animal's) inability to integrate sensory stimulation with patterns of movement and thought. Let's call this second kind of blindness *experiential blindness* because it is blindness despite the presence of something like normal visual sensation.

Does experiential blindness actually occur? If it does, then we must reject the input-output picture. To see is not just to have visual sensations, it is to have visual sensations that are integrated, in the right sort of way, with bodily skills. Experiential blindness would provide evidence for the enactive approach to perception.

There's good reason to believe that experiential blindness does occur. As an example, consider attempts to restore sight in congenitally blind individuals whose blindness is due to cataracts. Cataracts impair the eye's sensitivity by obstructing light on its passage to the retina. From the standpoint of the input-output picture, it would be natural to suppose that removing the cataract would be like sweeping aside the blinds, letting in the light and thus enabling normal vision. This is not in fact what the medical literature on this teaches us.[6] What we learn from the case studies is that the surgery restores visual *sensation*, at least to a significant degree, but that it does not restore sight. In the period immediately after the operation, patients suffer blindness despite rich visual sensations. That is to say, they suffer experiential blindness.

Consider a few examples. Gregory and Wallace describe a cataract-surgery patient, S.B.:

> S.B.'s first visual experience, when the bandages were removed, was of the surgeon's face. He described the experience as follows: He heard a voice coming from in front of him and to one side: he turned to the source of the sound and saw a "blur." He realized that this must be a face. Upon careful questioning, he seemed to think that he would not have known that this was a face if he had not previously heard the voice and known that voices came from faces. (1963, 366)

Sacks makes a similar observation of his patient Virgil:

> He seemed to be staring blankly, bewildered, without focusing, at the surgeon, who stood before him, still holding the bandages. Only when the surgeon spoke—saying "Well?"—did a look of recognition cross Virgil's face.
>
> Virgil told me later that in this first moment he had no idea what he was seeing. There was light, there was movement, there was color, all mixed up, all meaningless, a blur. Then out of the blur came a voice that said, "Well?" Then, and only then, he said, did he finally realize that this chaos of light and shadow was a face—and, indeed, the face of his surgeon. (1995, 114)

Finally, Valvo's patient made the following entry in his diary:

> after the operation, I saw the light of the doctor's probe, appearing like an atomic explosion on a background of black. Then I saw something which I understood afterwards was the doctor's hand and, clearly, his fingers; they seemed small and red (and to me it resembled the hand of the devil). . . . What I took to be black holes I recognized after about a month as windows in houses facing the hospital. (Valvo 1971, 9)

These patients suffer from experiential blindness, or so I propose. Their visual sensitivity is restored, to be sure. Each of them undergoes dramatic and robust visual impressions or sensations in the immediate aftermath of the surgery. But none of them, in having these sensations, has acquired the ability to see, at least not in anything like the normal sense. The visual impressions they now receive remain confusing and uninformative to them, like utterances in a foreign language. They have sensations, but the sensations don't add up to experiences with representational content.

The existence of experiential blindness is of great importance. It demonstrates that merely to be given visual impressions is not yet to be made to see. To see one must have visual impressions that one *understands*. This is brought out forcibly in connection with Gregory and Wallace's S.B. They write, concerning S.B.'s state about a month after his operation:

> At first impression he seemed like a normally sighted person, though differences soon became obvious. When he sat down he would not look round or scan the room with his eyes; indeed he would generally pay no attention to visual objects unless his attention were called to them, when he would peer at whatever it was with extreme concentration. (Gregory and Wallace 1963, 364)

S.B. has visual impressions, but he lacks, at least in part, a practical understanding of their significance for movement and thought. The point is not only that S.B. lacks the ability to use his impressions to guide movement, although this is true. In normal perceivers, sensation is smoothly integrated with capacities for thought, and for movement; so, for example, we naturally turn our eyes to objects of interest, we modulate our sensations with movement in a way that is responsive to thought and situation. A sharp sound makes us turn in the direction from which the sound emanates. A ball rushes toward us and we reflexively duck. A person speaks to us, we turn to him or her. In this sort of way, and in countless ways like this, sensory impressions are immediately coupled with spontaneous movement. This coupling is missing for S.B. and the other patients. S.B.'s deficit, however, is more far-reaching even than this; S.B.'s inability to use what he sees to guide movements is caused by what is in effect an inability to see (experiential blindness). S.B. lacks understanding of the sensorimotor significance of his impressions; he lacks knowledge of the way the stimulation varies as he moves or would move. As a result, or so I propose, his impressions are without content and he is, to a substantial degree, blind.

Defenders of the input-output picture may be skeptical. Perhaps, they might argue, one can grant that the newly post-operative patients are blind, but without conceding that they are *experientially* blind. After all, there would seem to be evidence that their difficulty stems not so much from abnormal sensorimotor integration, as from abnormal *sensations*. Look at how they describe their experience. Sack's Virgil reports encountering movement, color, "all meaningless, a blur," and Valvo's patient describes impressions of atomic explosions on a background of dark. These aren't normal visual sensations. They are clearly abnormal. This line of objection may be strengthened by considering that inactivity of retina and visual cortex could lead to some degree of stunting of the development of neural connections needed for mature adult vision. Until these possibilities are eliminated, the skeptic can insist that we are not entitled to treat the condition of these patients as *experiential* blindness (i.e., as blindness due to lack of sensorimotor knowledge rather than to lack of perceptual sensitivity). To establish genuine experiential blindness, we need to control for changes in the quality of visual impressions themselves. Until we can do this, we have no argument for the enactive approach and no argument against the input-output picture.

This objection has some force. In section 1.3 I turn to an example of putative experiential blindness that is not vulnerable to this criticism. Taken together the two examples make a strong case for experiential blindness, and so for the enactive approach.

1.3 Being Blinded by What You See

Glasses, or spectacles, belong to the humdrum everyday technology of perception. One of the most common kinds of glasses, or corrective lenses, are for myopia (or nearsightedness). In myopia, light from distant objects, which enters the eye in parallel rays, is brought to a focus before the retina, rather than on it. Light from nearer objects does not consist in parallel rays and is brought to a focus on the retina. What glasses for myopia do is bend light from distant objects so that it enters the eye at the same angle as light from nearer objects, thus allowing it to be brought to a focus on the retina.

What happens if glasses consist of prisms that distort or bias the light entering the eyes in strange or unnatural ways? Suppose you construct lenses so that light from objects on the left enters the eye just as light coming from an object on the right would enter the eye if you were not wearing the lenses. A left-side object would thus stimulate right-side retina, and also right-side brain (that is to say, the parts of the retina and brain normally stimulated by objects on the right). It is reasonable to suppose that in a case such as this you would have an experience as of an object on the right side.

In fact, as experiments by Stratton (1897), Kohler ([1951] 1964), and later Taylor (1962) demonstrate, this is not what happens, or at least not what happens right away. The initial effect of inverting glasses of this sort is not an inversion of the content of experience (an inversion of what is seen) but rather a partial disruption of seeing itself. Inverting lenses give rise to experiential blindness. Consider what one subject, K, wrote of his initial experiences in Kohler's experiment with displacing spherical prism spectacles:

> During visual fixations, every movement of my head gives rise to the most unexpected and peculiar transformations of objects in the visual field. The most familiar forms seem to dissolve and reintegrate in ways never before seen. At times, parts of figures run together, the spaces between disappearing from view: at other times, they run apart, as if intent on deceiving the observer. Countless times I was fooled by these extreme distortions and taken by surprise when a wall, for instance, suddenly appeared to slant down to the road, when a truck I was following with my eyes started to bend, when the road began to arch like a wave, when houses and trees seemed to topple down, and so forth. I felt as if I were living in a topsy-turvy world of houses crashing down on you, of heaving roads, and of jellylike people. (Kohler [1951] 1964)

K is not completely blind, to be sure; he recognizes the trucks, the trees, and so forth. But nor is he completely able to see. His visual world is distorted, made unpredictable and topsy-turvy. To this extent, K suffers blindness. Crucially, the kind of blindness K suffers is not caused by any defect in sensation. K receives normal stimulation. The light reaching his eyes is sharply focused and fully information-bearing. He receives exactly the stimulation he would receive were he looking at an object in a different spatial location without the inverting lenses. The inability to see normally stems not from the character of the stimulation, but rather from the perceiver's understanding (or rather failure of understanding) of the stimulation.

This is exactly what the enactive approach would lead us to expect, as O'Regan and I have argued (O'Regan and Noë 2001a,b; Noë 2002a; see also Hurley and Noë 2003a). The basis of perception, on our enactive, sensorimotor approach, is implicit practical knowledge of the ways movement gives rise to changes in stimulation. When you put on the distorting lenses, the patterns of dependence between movement and stimulation are altered. This alteration has the effect of abrogating sensorimotor knowledge or skill, even though there is no change in the intrinsic character of stimulation. As a consequence, movements of the eye and head give rise to surprising and unanticipated changes in sensory stimulation. The result is not *seeing differently*, but failing to see.

Strictly speaking, the goggles do not produce *total* experiential blindness. This is because the only sensorimotor dependencies that are affected are those pertaining to aspects of spatial content. For example, left-right reversing prisms do not affect one's sense of up and down (although they do affect one's sense, say, of the speed with which the visual world "swings by" as one moves one's eyes). Moreover, left-right reversing goggles do not affect one's sense of light and dark, color, and so on. When you put on left-right reversing goggles, you enjoy *some* perceptual experience. For example, you can tell whether the lights are on. This is not surprising, given that the goggles don't change *all* the patterns of sensorimotor dependence, only those that are related to spatial orientation.

The enactive view would also lead us to expect that vision will be restored once one comes to grips with the new patterns of sensorimotor dependence. The experimental literature supports this. Kohler's reports suggest that adaptation occurs in stages. The first stage of adaptation is the experience of inverted content. Now objects on the left do indeed look just as if they are on the right. Your visual experience has acquired nonveridical content. But this state of partial adaptation is highly unstable. Your left hand may look as if it is on the right, but it continues to *feel* as if it is on the left (Hurley and Noë 2003a). And when you snap your fingers, the sound of your "hand on the right" seems to come from the left. At the next stage of adaptation, visual experience "captures" auditory and proprioceptive experience, resolving conflicts between these sensory modalities in favor of vision. The object on the left not only looks as if it is on the right, but it now sounds and feels as if it is too. If subjects are allowed (indeed required) to actively engage with and explore their environment, a third stage of adaptation comes about in which experience comes to "right itself" and veridicality is restored. Now objects on the left look as though they are on the left, even though they continue, as before, to activate retinal and brain areas associated with right-placed stimuli. This is the final stage of adaptation. (For discussion, see Hurley and Noë 2003a.)

From the standpoint of the enactive view, this is an extraordinarily important phenomenon, a powerful illustration of the fact that perceptual experience acquires content as a result of sensorimotor knowledge. I return to some of these issues in chapter 3. For now the point is this: Once full adaptation has been achieved, the result of *removing* the lenses is comparable to the initial effects of putting them on. Taking the glasses off induces exactly the same kind of experiential blindness, and for exactly the same reasons that putting them on did at first: The glasses (or their absence) cause a sudden abrogation of the patterns of dependence of sensation and movement. Kohler's subject describes the effects of taking the lenses off as follows:

> As I begin to move and walk about, the room begins to move too. What I am experiencing are the apparent movements of the objects around me. As I approach one of them, it seems to move to the right. I reach out for it and touch—air: my arm has completely missed it, passed to the left of it. . . . Even more peculiar are the relative changes inside the room. When I move my head (vertically or horizontally), not a single point remains stationary in relation to another point. If a certain point moves along with me in the visual field, then some other point will infallibly move in the opposite direction, as if indicating to me in no uncertain terms that it is not the least bit bound by what the other points appear to be doing at the time.
>
> The world I am in seems to have become a total chaos of continuously changing distances, directions, movements, and Gestalten. Nothing remains stable and the experience is so confusing that I am unable to detect what laws the transformations abide by . . . everything remains without rhyme or reason. There is no such thing as *a* size or *a* movement; as soon as I move my body or my head, any object is apt to become smaller or larger, stationary or mobile. (Kohler [1951] 1964, 65)

The effect of removing the lenses, then, is to produce nonveridical, distorted, chaotic visual impressions, even though the patterns of visual sensation now produced are exactly as they were before the lenses were first put on. Objects on the left stimulate the parts of the eye and brain that have always supported the sensory experience of leftness. The inability normally to perceive is the result not of changes in the intrinsic character or location of the sensory stimulation, but rather of the induced breakdown in our mastery or control over the ways sensory stimulation changes as a function of movement.

To summarize, experiential blindness exists and is important for two reasons. First, it lends support for the enactive view. Genuine perceptual experience depends not only on the character and quality of stimulation, but on our exercise of sensorimotor knowledge. The disruption of this knowledge does not leave us with experiences we are unable to put to use. It leaves us without experience. For mere sensory stimulation to constitute perceptual experience—that is, for it to have genuine world-presenting content—the perceiver must possess and make use of *sensorimotor knowledge*.

Second, it provides a counter example to the more traditional input-output picture. Kant ([1781–1787] 1929) famously said that intuitions without concepts are blind. The present point is that intuitions—patterns of stimulation—without knowledge of the sensorimotor significance of those intuitions, are blind. Crucially, the knowledge in question is practical knowledge; it is know-how.[7] To perceive you must be in possession of *sensorimotor bodily skill*.

1.4 The Joys of Seeing

A natural line of objection to the enactive approach goes like this: True, our perceptual capacities are bound up with bodily skill and action. We use our eyes to guide our movements and to enable action. But that is not always the case. Sometimes we see not in order to act, but just in order to know, or to enjoy our experiences of seeing. When you lie back and watch the passing clouds, or when you visit an art gallery, or watch TV, you are not using visual skills for purposes of action. Pylyshyn (2001) has made this point; he adds that "much of what we see guides our action only indirectly by changing what we believe and perhaps what we want" (999).

This criticism of the enactive view would seem to gain support from the study of neurological disorders of vision. Patients with *optic ataxia* (resulting from lesions in posterior parietal cortex) are unable to make use of what they see to guide movements. As Milner and Goodale write, "Yet despite the failure of these patients to orient their hands, to scale their grip appropriately, or to reach towards the right location, they have comparatively little difficulty in giving perceptual reports of the orientation and location of the very objects they fail to grasp" ([1998] 2002, 520). Milner and Goodale argue that there are two largely autonomous visual systems. Damage to the dorsal stream (from striate to posterior parietal cortex) impairs visuomotor skill without harming vision or visual awareness as such. Damage to the ventral stream (from striate to inferotemporal cortex), in contrast, can produce striking visual agnosias, impairing object recognition and judgments of size, orientation and location, while leaving visuomotor skill largely intact. Their subject D.F., for example, showed excellent visually guided control of grasp, reaching, and hand posture in general. According to Milner and Goodale, "Yet when she was asked to use her finger and thumb to make a perceptual judgment of the object's width on a separate series of trials, D.F.'s responses were unrelated to the actual stimulus dimensions, and showed high variation from trial to trial" ([1998] 2002, 520–522). This neurological evidence suggests that although some facets of vision are bound up with visuomotor skill, this is not true of vision as a whole. The enactive approach, it would seem, exaggerates the importance of action in perception.

This criticism rests on a misunderstanding of the enactive approach. The basic claim of the enactive approach is that the perceiver's ability to perceive is constituted (in part) by sensorimotor knowledge (i.e., by practical grasp of the way sensory stimulation varies as the perceiver moves). The enactive approach does not claim that perception is *for* acting or for guiding action. The existence of optic ataxia, therefore, does not undercut the enactive view, for from the fact that a patient suffers optic ataxia, it doesn't follow they he or she lacks the relevant sensorimotor knowledge. What would undercut the enactive approach would be the existence of perception in the absence of the bodily skills and sensorimotor knowledge which, on the enactive view, are constitutive of the ability to perceive. Could there be an entirely inactive, an *inert* perceiver?

Before we turn to this question, consider a simpler worry. Paralysis is certainly not a form of blindness. But isn't that precisely what the enactive view requires, that the paralyzed be experientially blind? No. The enactive view requires that perceivers possess a range of pertinent sensorimotor skills. It seems clear that quadriplegics have the pertinent skills. Quadriplegics can move their eyes and head, and to some extent, at least with help from technology, they can move their bodies with respect to the environment (e.g., by using a wheelchair). More important, paralysis does not undermine the paralyzed person's practical understanding of the ways movement and sensory stimulation depend on each other. Even the paralyzed, whose range of movement is restricted, understand, implicitly and practically, the significance of movement for stimulation. They understand, no less than those who are not disabled, that movement of the eyes to the left produces rightward movement across the visual field, and so forth. Paralyzed people can't do as much as people who are not paralyzed, but they can do a great deal; whatever the scope of their limitations, they draw on a wealth of sensorimotor skill that informs and enables them to perceive.

Quadriplegics, who are without sensation as well as movement, live extremely active lives. As the clinical neurophysiologist Jonathan Cole remarks, "Try balancing in a chair without any sensation from the neck down" (personal communication, but see Cole 2004). Quadriplegics are continuously engaged in the task of orienting themselves in relation to the world around them and to gravity (as Cole 2004 discusses).[8]

There is in fact strong empirical evidence that a more thoroughgoing paralysis—for example, of the eyes themselves—would cause blindness. In normal perceivers, the eyes are in nearly constant motion, engaging in saccades (sharp, ballistic movements) and microsaccades several times a second. If the eyes were to cease moving, they'd lose their receptive power. In particular, it has been shown that images stabilized on the retina fade from view (Ditchburn and Ginsborg 1952; Riggs et al. 1953; Krauskopf 1963; Yarbus 1967). This is probably an instance of the more general phenomenon of *sensory fatigue* thanks to which we do not continuously feel our clothing on our skin, the glasses resting on the bridge of our nose, or a ring on our finger. This suggests that some minimal amount of eye and body movement is necessary for perceptual sensation.

There is also developmental evidence that normal vision depends not only on movement of the body relative to the environment, but on *self-actuated* movement. Held and Hein (1963) performed an experiment in which two kittens were harnessed to a carousel. One of the kittens was harnessed in such a way that it stood firmly on the ground. The other kitten was suspended in the air. As the one kitten walked, both kittens moved in a circle. As a result, they received identical visual stimulation, but only one of them received that stimulation as a result of self-movement. Remarkably (but not surprisingly from an enactive viewpoint), only the self-moving kitten developed normal depth perception (not to mention normal paw-eye coordination). From an enactive standpoint, we can venture an explanation for this: Only through *self*-movement can one *test* and so *learn* the relevant patterns of sensorimotor dependence.[9]

There are, however, deeper and more compelling reasons to be skeptical of the very idea that there could be a truly passive, inert perceiver. One of the main aims of this book is to demonstrate this. A few preliminary remarks now can set us on the path.

The extraordinary case of Ian Waterman, documented by Jonathan Cole (1991), serves as an illustration. Waterman, as a young man, took ill with a virus that produced a dramatic and far-reaching neuropathy. Although his motor nerves remained unaffected, he lost all sensation from the neck down, except for the sensation of pain (e.g., pin pricks) and temperature. In particular, he lost what is sometimes called "the sixth sense," namely, the sense of movement and position known as proprioception and kinaesthesis. Waterman was initially, in effect, paralyzed. Despite the fact that he possessed a normally functioning motor system, he was unable to bring his limbs and body under his control. In the absence of proprioceptive feedback, he was unable to move. Eventually, he regained a good measure of motor skill by learning to substitute vision for muscular sense. By intense visual concentration, he was able to control his body movements. However, if he were put into a position in which he could not view his body (say, reclining on a couch), or if the lights were to go out, he would collapse to the ground, unable to move. As Cole says, in the case of Ian Waterman, for his body to be out of sight was, literally, for it to be out of mind.

What would Ian Waterman have done if he had been blind as well? Let's consider a made-up case. Suppose that you suffer from a neuropathy like Waterman's—that is to say, that you have lost all sense of movement, position and posture—but imagine that you are, in addition, deaf and blind. Let's further imagine that you have *normal* sensation from the neck down. Strictly speaking, this last detail is not consistent with the supposition that you lack all proprioception, since proprioception depends in part on cutaneous sensitivity (in addition to the activation of muscle spindles and tendon receptors). But let's put this complication aside and imagine that you have normal tactile sensation, but that you lack a sense of movement, position, and posture, and that you are deaf and blind. To imagine this, then, is to imagine that you are inert, that you are radically unable to act with your body. It is to imagine that your body has been lost to you as an animated part of yourself.

Now let us ask, would you be able to perceive by touch? Could you enjoy tactile experience of the world around you? By hypothesis your cutaneous sensory receptors are intact, so there is no question whether you can feel, that is, have tactile sensations. The question is, in having tactile sensations, would you perceive how things are around you?

In general, there are reasons to doubt that tactile sensation or feeling is sufficient for tactile perception. To perceive by touch, for example, the rectangularity of something you hold in your hands, or the layout of furniture in a room (as a blind person might, by moving around and reaching and touching) is not merely to have certain feelings or sensations. After all, the rectangularity is not captured by specific sensations. There is no unitary sensation or feel of a rectangle. The rectangularity is made available to you, in touch, by your active touching (probing, prodding, stroking, rubbing, squeezing) with your hands. What informs you of the shape of what you feel or hold is not the intrinsic character of your sensations, but rather your implicit understanding of the organization or structure of your sensations. The shape is made available thanks to the way in which your sensations co-vary or would co-vary with actual or possible movements. In perceiving the thing as rectangular, you understand, implicitly, that, for example, if you move your hands like *so*, you'll encounter corners that stand in a certain relation to each other, and so forth. The same sort of point can be made about the tactile perception of the layout of furniture in the room. Your tactile impression that things are arranged thus and such consists not in the sensations in your hands and feet, but in the way those sensations result from attentive movement through the space. What is informative is the fact that you bump your foot here, that you cannot press forward there, and so on. You perceive the furniture layout when you understand the way your sensations are fixed as a function of movement through the space. In this way, sensation and sensorimotor knowledge work together to produce the perception of the spatial layout of the room.

For this reason it seems plausible that feeling alone is not sufficient to enable you to learn about or discover the properties of objects or layouts around you. It is altogether unclear, in the extreme case of inert perception I am imagining, that you would be able to learn how things are around you, for you would be unable to probe in response to sensations, and so would be unable, even in thought, to coordinate them. How could you perceive the object *as rectangular* without moving it across your body surfaces, or without moving your body surfaces across it?

One response is that you could at least perceive *heat*, say, or *texture*. For these simple tactile qualities, it might seem that feeling is sufficient for tactile perception. This is plausible, but we need to be cautious. You will have sensations, to be sure, but will they amount to perception of how things are, even with respect to heat or texture? Because you are completely inert, you may be unable to localize your sensations on your body. Suppose someone presses, say, a warm spoon against your thigh. What will you experience? A feeling of warmth on your thigh? Or merely a feeling of warmth? In either case, your experience will be confined to the character of your own sensations. Your would-be perception of the warmth of something will collapse into the mere sensation of warmth somewhere (perhaps conjoined with the inference that that sensation is likely to have an external source). Such a sensation-plus-guesswork falls short of constituting perceptual experience with content (at least of the normal sort). At best, it would seem, it is a primitive antecedent of the latter.

Remember, what is in question here is the experience of one who is radically inert. Ian Waterman, and others with similar conditions, are not radically inert. They *are* able to locate sensations of heat on, say, their legs; without proprioception (or vision), they are unable to locate the leg, however, in surrounding space. My suggestion is that for one whose sensations bear no familiar dependence on patterns of movement, even this localization on the surfaces of one's own body would be impossible.

One objection might be that sometimes mere touch is enough for perception. A sense of touch, for example, signals the presence of a fly, or some other object. Yes, and no. We *do* experience the presence and location of a fly, say, by the merest sense of touch on the skin, but this is only because we also possess the sensorimotor skills needed to interpret that touch as referring to a type of movement or position in space. We spontaneously withdraw our arm from the touch, for example, and in this way we give expression to the understanding that such a movement of the arm is a movement *away* from the point of contact with the fly. What would it be to experience the touch as an instant of contact with the fly, if one were not also able, thus, to understand the way movements would alter one's relation to that point in space?

The enactive view insists that mere feeling is not sufficient for perceptual experience (i.e., for experience with world-representing content).[10] O'Shaughnessy (2000) has argued that it is not even necessary for perceptual experience.[11] You could perceive the presence of a wall by reaching out and pressing it with your numb hand. Your ability to do that probably depends on your having feelings elsewhere. But, as O'Shaughnessy points out, those feelings are not part of your experience; they do not belong to the scope of your attention in perceiving the wall by touch. This point is nicely illustrated by the case of a blind person perceiving by means of a cane. There is no feeling at the end of the cane, yet it is with the end of the cane that the blind person makes contact with the world. It is probable that the ability thus to perceive depends on one's capacity for sensation (say, in the hand that holds the cane). But crucially, sensations in your hand are *not* constituents of your cane-based perceptual experience of the environment.

On the enactive view, all perception is in these respects like touch. Mere sensation, mere stimulation, falls short of perceptual awareness. As stated earlier, for perceptual sensation to constitute experience—that is, for it to have genuine representational content—the perceiver must possess and make use of *sensorimotor knowledge*. To imagine a truly inert perceiver is to imagine someone without the sensorimotor knowledge needed to enact perceptual content.

1.5 Action in Perception in Cognitive Science

The enactive approach to perception draws on a number of distinct traditions in philosophy, psychology, and cognitive science. The touch-like character of vision plays an important role in Merleau-Ponty's philosophical writing ([1948] 1973, [1945] 1962), and in the writing of other phenomenologists (e.g., Jonas 1966). Berkeley ([1709] 1975), Poincaré ([1902] 1952, [1905] 1958), Husserl ([1907] 1997), and Evans (1982) offer accounts of the spatial content of perceptual experience that anticipate elements of the enactive approach. (I turn to this topic in chapter 3.) In cognitive science, both the motor theory of perception (Berthoz [1997] 2000; Jeannerod 1997) and Gibson's ecological approach to perception (Gibson 1979) lay great emphasis on perception as an activity. Several other influential thinkers have emphasized and developed, in different ways, the sensorimotor basis of perception—for example, MacKay (1967, 1973); Arbib (1989); Koenderink (1984a,b); Varela (Varela, Thompson, and Rosch 1991; Maturana and Varela 1987); and O'Regan (1992). In addition, there has been a great deal of interest in recent cognitive science on the relation between perception and action—for example, Ballard (1991, 1996, 2002); Thompson (Thompson, Palacios, and Varela 1992; Thompson 1995); Humphrey (1992); Churchland, Ramachandran, and Sejnovksy (1994); Kelso (1995); Cotteril (1995, 2001); Clark (1997, 1999); Hurley (1998); Järvilehto (1998a,b, 1999, 2000); O'Regan and Noë (2001a,b,c); Noë (2002a,b).[12] A hallmark of this new work is the idea that the relation between perception and action is more complicated than traditional approaches have supposed.

In this section I give a brief sketch of some lines of thought that converge on the enactive approach. I don't try to give anything like a complete survey.[13] The enactive view gains indirect support from these disparate research lines. Importantly, care is required before the enactive approach is identified with any of these disparate strands. Most recent work on the relation of perception and action stops short of making the constitutive claim that defines the enactive standpoint: It does not treat perception as a kind of action or skillful activity (or as drawing on a kind of sensorimotor knowledge); rather it treats (a good deal of) perception as *for* the guidance of action.

One important source of the idea that perception and action are more tightly connected than the input-output picture tends to suppose is comparative and evolutionary work on perception. It seems probable that vision, for example, evolved as a mechanism of motor control. Certainly it is the case that in simple organisms the absorption of light may have the effect of modulating locomotion thanks to direct biochemical linkages (Bruce and Green [1985] 1990; Humphrey 1992). As an example, consider the phototactic water beetle (*Dytiscidae*). (This example is discussed in Milner and Goodale 1998, 6. See also Schone 1962.) The absorption of light directly produces a modulation of swimming behavior, leading the organism toward the light. In a normal aquatic environment this tends to lead it upward to the air it needs to survive. But the animal will swim to the bottom if that's where a light source is placed, resulting in death. A well-known example is the visual system of the frog, where certain patterns of stimulation are thought to activate "a fly detection response" leading to a darting out of the tongue in the direction of the stimulus (Lettvin et al. 1959). It is probable that our own sophisticated visual capacities develop from these humble sensorimotor beginnings.

A second important source is work in neurology, and psychology, on the existence of two functionally separable visual systems in the brain, one subserving vision and the other subserving the control of visually guided behavior. As mentioned earlier, the neurological evidence is striking: Visual agnosics may have normal visuomotor skills in the absence of normal perception and patients with optic ataxia can make normal perceptual judgments in the absence of normal visuomotor skill (Milner and Goodale 1995; but see Rossetti, Pisella, and Vighetta 2003). There is also psychological evidence that supports this two-systems approach. In particular, evidence exists that vision may guide motor behavior (say, pointing) unconsciously (or implicitly). Subjects may have no access to the visual information, in the sense that they are unable to *say* what they see, even though this information is exploited to guide movement. Bridgeman and his colleagues, for example, gave subjects the task of pointing at a target that was displaced and then extinguished (Bridgeman et al. 1979). They were asked whether the target had been displaced or not. Subjects tended to point correctly, whether or not they noticed a displacement. In a later study, they created an illusion that a target had jumped by moving a background frame in which the target was presented. Pointing accuracy was not affected by the illusion of displacement (Bridgeman, Kirch, and Sperling 1981). Apparent displacement of the target affected only perception, not pointing. In a second condition they allowed subjects to adjust the target's real (as opposed to induced) motion so that it moved in phase with the frame and came to look stationary. Despite this perceived lack of displacement, subjects successfully pointed to the real displacement. In this condition, real displacement went unperceived but affected the motor system. In this way, Bridgeman and his colleagues demonstrate that perceptual and motor functions are successfully dissociated. (For a review of this and other research on the two-systems hypothesis, see Bridgeman 1992 and Bridgeman et al. 2000.)

The significance, for the enactive approach, of this dissociation of perception and perceptual-guidance of action is delicate. The existence of a "how" (dorsal) stream, dedicated to the visual guidance of action, would seem to lend some measure of support to the enactive approach, insofar as it gives additional support to the claim that there are strong constitutive links between perception and action. However, the existence of a "what" (ventral) stream, dedicated to perceptual representation, experience, and identification, would seem to indicate that at least some aspects of perception are independent of links to motor systems.

In fact, Milner and Goodale's two visual systems hypothesis is, at best, orthogonal to the basic claims of the enactive approach. The enactive approach is not committed to the idea that vision is for the guidance of action, so neither the fact that some visual processing *is* for the guidance of action, nor the fact that some visual processing *is not*, has any direct bearing on the enactive approach. From the standpoint of the enactive approach, all perceptual representation, whether the result of dorsal or ventral stream activity, depends on the perceiver's deployment of sensorimotor skills.

One idea that serves to guide investigation into the active character of perception is the recognition that some of the most difficult challenges faced by traditional approaches to perception are, in a sense, debts incurred precisely by a failure to make room, in an account of perception, for the role of action.

As an example, note that traditional approaches to vision suppose that the problem of vision is one of "inverse optics," namely, to produce a description of the three-dimensional environmental layout from a projection of that environment in two dimensions on the retina (Marr 1982). The problem, as is well known, is ill posed. Just as a small object nearby can project the same image as a large object far away, so, in general, one cannot "read off" a description of the scene from the information made available in the retinal image. When the problem is framed this way, the brain's task is frequently supposed to be that of forming a hypothesis (e.g., an inference to the best explanation) as to the distal causes of proximal stimuli (e.g., Fodor 1975).

But why should we suppose that the data for vision is the content of the retinal image? If we think of the perceiver not as the brain-photoreceptor system, but rather as the whole animal, situated in the environment, free to move around and explore, then we can take seriously the possibility that the data for vision (as distinct from data for the photoreceptor) are not the content of a static snapshot-like retinal image. At the very least, the animal or brain has access to the "dynamic flow" of continuously varying retinal information. Optic flow contains information that is not available in single retinal images (Gibson 1979). For example, expanding optic field flow indicates that the observer is approaching a fixed point; contracting optic field flow indicates that he or she is moving away from a fixed point (Gibson 1979, 227).

This suggests that part of what has made the computational problem of vision such a difficult one is that it is framed in an artificially restrictive way. Perceivers aren't confined to their retinal images in the way traditional theorists have supposed.

Gibson took these points further. He argued that the animal has access not only to information contained in optic flow, but also to information about the way optic flow varies as a function of movement. When we move through a cluttered environment, for example, one object may come to occlude another. But occlusion, as Gibson noticed, is reversible (1979, chap. 5). By tracing movements back, you can bring an occluded surface back into view. In perceptual activity the perceiver is thus able to differentiate mere occlusion from obliteration. This is an example of the way it is possible for the animal to explore the structure of the flow of sensory changes and to discern in this structure *invariant* properties of the environment. Gibson also held that his 'ecological' approach can handle the problem of inverse optics mentioned earlier. This problem turns out to be a consequence of the optional assumption that the data for vision is confined to the retinal image. For an active animal, it is easy to disambiguate a large but distant object from a near but large one.

Gibson went further than this, however. He argued that just as there is a fit between an animal and the environmental niche it occupies, thanks to the coevolution of animal and niche, so there is a tight *perceptual attunement* between animal and environment. Because of this attunement, animals (as embodied wholes, not as brain systems attached to photoreceptors) are directly sensitive to the features of the world that afford the animal opportunities for action (what Gibson 1979, chap. 8, called "affordances"). For the active animal, the ground is directly perceived as walk-uponable, and the tree stump as sit-uponable. The theory of affordances is very controversial, as is Gibson's theory of direct perception more generally. He has been roundly criticized by, among others, Ullman (1980) and Fodor and Pylyshyn (1981). I do not endorse Gibson's views across the board. However, many of the criticisms leveled against him can be answered pretty easily. In fact, from the standpoint of the enactive approach, it is possible to reconstruct certain of his most controversial claims (e.g., the theory of affordances and his account of the so-called ambient optic array).

We return to these themes in chapters 3 and 4. For now the crux is this: There is a solidifying consensus in cognitive science that information available to an active animal greatly outstrips information available to a static retina, and that it is a mistake to suppose that the animal's data for visual perception are confined to the contents of the retinal image.

Once we adopt an active approach to perception, treating the active animal as the subject of perception, we are led to question the assumption (made by Marr and most theorists working in the computational school) that vision is a process whereby the brain produces an internal representation of the world (of what is seen). Churchland, Ramachandran, and Sejnowski (1994) call this the theory of pure vision, namely, the doctrine that vision is a matter of generating a detailed internal representation of the visual world on the basis of information available at the retina alone. If vision evolved for the purpose of enabling creatures to get by in a hostile environment (e.g., to facilitate the famous four Fs, etc.), then why assume, by building it into the definition, as it were, that vision requires the construction of a detailed internal representation? Presumably that is an empirical matter (Noë, Pessoa, and Thompson 2000).

An active approach to perception raises a more significant concern. If the animal is present *in* the world, with access to environmental detail by movements—that is, if it is active, embodied, environmentally situated—then why does it need to go to the trouble of producing internal representations good enough to enable it, so to speak, to act as if the world were not immediately present? Surely we sometimes need to think about the world in the world's absence (when it's dark, say, or when we're blind, or not at the location we're interested in), and for such purposes we must (in some sense) represent the world in thought. But what reason is there to think that this is the case in standard perceptual contexts? In many situations, we need only move our eyes, or move our head, or turn around, to get whatever information we need about the environment. How many bookshelves are there in your room? You don't need to have an internal representation to answer; you need only be able to turn around and take a look. Why not let the world serve as an external memory, as O'Regan (1992) has argued, or why not let the world serve, in Brooks's (1991) phrase, as its own model?[14] It makes good evolutionary and engineering sense to *off-load* the representations. We are built in such a way that we can get the information about the world that we need, when we need it.

The claim is not that there are no representations in vision. That is a strong claim that most cognitive scientists would reject. The claim rather is that the role of representations in perceptual theory needs to be reconsidered. (See Noë, Pessoa, and Thompson 2000; Noë 2001; O'Regan and Noë 2001a.) It is a mistake to suppose that vision just is a process whereby an internal world-model is built up, and that the task-level characterization of vision (what Marr [1982, 23–31] called the computational theory of vision) should treat vision as a process whereby a unified internal model of the world is generated. This is compatible with there being all sorts of representations in the brain, and indeed, with the presence of such representations being necessary for perception.[15] Marr famously claimed of

Gibson that he "vastly underrated the sheer difficulty" of the information-processing problem of vision (1982, 30). As the vision scientist Nakayama has responded (1994), there's reason to think that Marr and his followers underestimated the difficulty of correctly framing what vision is at the task or computational level. Vision isn't a process whereby the brain constructs a detailed internal world representation. Once one acknowledges this, then "detailed internal world representations" can be demoted from their theoretical pride of place.

I have argued that the role of representations in perceptual theory needs to be reconsidered. (See Noë, Pessoa, and Thompson 2000; Noë 2001a; O'Regan and Noë 2001a.) This is exactly the path explored by Dana Ballard's animate vision program (Ballard 1991, 1996, 2002). To understand his approach, suppose you are in strange city and your task is to reach the castle on the hill in the center of town. Compare two possible strategies. On a first strategy, you make use of a map. You plot your position on the map, and that of the castle, and you figure out a path connecting the two points. Now you're ready to roll. As you move along, you keep track of your progress on the map. If the map's a good one—if there is a one-to-one correspondence between points in space, and points on the map, and if you don't get confused about what you're doing, you'll get to your goal.

The second strategy is simpler, and somewhat cruder. You look around and notice that you can see the castle on the hill. You can see it rising up on a ridge on the other side of town. So you dispense with a map and head out in the direction of the castle. You just keep the castle locked into view. This second strategy may be crude, but it has distinct advantages. For one thing, to pursue it you don't need a map. Maps are expensive and they are not all that easy to use. It takes time to study the map, to pinpoint yourself and your goal, and so forth. But there is a downside too. The strategy will only work if you can actually see the target (if your eyes are good, if it isn't night), and if heading toward it is likely to reveal a path leading up to it. In a maze-like city, where many ways dead-end, and others lead around the mountain, not up it, the second strategy won't work. That it works depends on the way the environment is, on your skills, and on the way you are embedded in that environment.

Ballard, who works in robotics and artificial intelligence, has proposed that given the nature of our environment, and the way we are embedded in it, vision is in a position to take advantage of something like the second strategy. Traditional approaches to vision have always assumed that we deploy the first strategy. If your aim is to pick up a coffee cup, reasons Ballard, you don't need *first* to build up a detailed internal representation of the cup in space (Ballard 1996). You can just lock your gaze on the cup—your gaze is a way of pointing at the cup, a *deictic* act—and let the cup play a role in guiding your hand to it. Instead of plotting a course through an internal map, you act on what you look at, and you let the fact that what interests you is there in front of you play a guiding function. An important consequence of this proposal is that it lessens the representational burden of the system, and that it does so by making explicit use of our bodily skills. Instead of having to ground ourselves by sheer cognition—constructing a representation of the point in space in our minds—we take advantage of the fact that we have more immediate links to the world because we are in the world from the start, and that we have the sorts of bodily skills to exploit those linkages.[16]

1.6 Persons and Their Bodies

The computational theory of vision stakes itself on the claim that what Marr called the algorithmic level of description of cognitive phenomena is autonomous with respect to the implementational level. Low-level, concrete facts about the brain and nervous system may be constraints on the processes unfolding at the higher level. But crucially, the transactions of the higher level are independent of what goes on at the lower level in both a metaphysical and an epistemological sense. Metaphysically, they are independent in that they are not constituted by what happens at the implementational level. So, for example, one and the same algorithmic system could be implemented by different physical systems. Epistemologically, they are independent in that one can fully understand the algorithmic processes without understanding how they are implemented. These metaphysical and epistemological factors gain support from methodological considerations as well. Marr thought that you couldn't develop a sound theory of vision from the bottom up. He wrote, "Trying to understand vision by studying only neurons is like trying to understand bird flight by studying only feathers: It just cannot be done" (Marr 1982, 27). The guiding metaphor is familiar: Psychology studies cognitive processes at a more abstract level than that of their biological realization just as the programmer studies computational processes at a more abstract level than that of their realization in the hardware of the machine.

A lot is supposed to hang on this autonomy of levels. For one thing, it is supposed to explain how a materialist can insist that psychology has a special domain of inquiry different from that of brain science (Fodor 1975; Dennett [1981] 1987). Psychology is interested in what the brain does, but at higher levels of abstraction than that of neuroscience. It is precisely this autonomy of levels that enabled Chomsky (e.g., 1965, 1980) to claim that linguistic theory seeks to explore language as part of our biological endowment, but in a manner completely divorced from the study of linguistic performance, on the one hand, or biological realization in the brain, on the other.

The enactive view applies pressure to the autonomy thesis. If perception is in part constituted by our possession and exercise of bodily skills—as I argue in this book—then it may also depend on our possession of the sort of bodies that can encompass those skills, for only a creature with such a body could have those skills. To perceive like us, it follows, you must have a body like ours. In general it is a mistake to think that we can sharply distinguish visual processing at the highly abstract algorithmic level, on the one hand, from processing at the concrete implementational level, on the other. The point is not that algorithms are constrained by their implementation, although that is true. The point, rather, is that the algorithms are actually, at least in part, formulated *in terms of* items at the implementational level. You might actually need to mention hands and eyes in the algorithms!

As an illustration, consider that, according to the enactive approach, vision depends on one's knowledge of the sensory effects of, say, eye movements, for example, movements of the eye to the right causes a shift to the left in the retinal image. This knowledge is eye-dependent. Or consider a different kind of case. We noted above that Ballard proposes that the perceptual localization of an object, such as a cup on the table before us, may depend on the gaze-fixing mechanisms of the eye. The algorithm says "reach where I'm looking now" or "put your hand here now" rather than something like "the cup is at such and such a point in space; move your hand there." Space may be represented not absolutely, but rather precisely in terms of movements. In this way, eyes, hands, and the neural systems that enable eye and hand movements are not merely ways of implementing a spatial perception and action algorithm, they are elements in the computations themselves.

A phenomenological example can help illustrate the way our bodies can enter into our experience. Suppose you are in an airplane. At takeoff it will look to you as if the front of the plane, the nose, rises or lifts up in your field of vision. In fact, it does not. Because you move with the plane, the nose of the plane does not lift relative to you. No lifting, strictly speaking, is visible from where you sit. What explains the illusion of the apparent rising of the nose? When the plane rises, your vestibular system detects your movement relative to the direction of gravity. This causes it to look to you as if the nose is rising.[17] The nose is rising, and it looks to you as if it is. But not for visual reasons. This phenomenon illustrates, first, one of the errors implicit in the idea of Pure Vision. How things are experienced *visually* depends on more than merely optical processes. This is a respect in which the content of a visual experience is not like the content of a photograph. Second, the example illustrates the way in which the character of our visual experience depends on our embodiment, that is, on idiosyncratic aspects of our sensory implementation.

I have said that only a creature with a body like ours can have experiences like ours. But now we ask: Must a creature have a body *exactly* like ours to have experience enough like ours to be thought of as *perceptual*, say, or as *visual*? That would be an undesirable consequence, ruling out even a very weak multiple realizability of sensory systems.[18] Clark and Toribio (2001; Clark 2002) have suggested that the enactive approach has this consequence, and that, therefore, the view is guilty of a kind of “sensorimotor chauvinism.”

To respond to this, consider Bach-y-Rita’s prosthetic visual system known as the tactile-vision substitution system (TVSS) (1972, 1983, 1984, 1996). Visual stimulation received by a head-mounted camera is transduced to activate an array of vibrators on the thigh of a blind subject. If the subject is free to move around and thus control tactile-motor dependencies, after a time she reports that she has the experience of objects arrayed in three-dimensional space. She is able to make judgments about the number, relative size, and position of objects in the environment. This is a mode of prosthetic perception. Crucially, it is not a mode of perception by touch, despite the fact that it enables the subject to perceive thanks to the activation of sensory receptors in the skin and neural processes in the somatosensory cortex. For touch is a way of perceiving by bringing things up against you, into contact with your skin. It is reasonable to admit that the resulting experiences are, if not fully visual, then vision-like to some extent. For example, using TVSS subjects describe objects being blocked from view when an opaque object interposes, and subjects are unable to perceive using TVSS if the lights are turned off. So let us say, then, that TVSS enables a kind of tactile vision. This is seeing (or quasi-seeing) without the deployment of the parts of body and brain normally dedicated to seeing, for example, the eyes and visual cortex. This is a striking example of multiple realization and neural plasticity. Somatosensory neural activity realizes visual experiences.[19]

The existence of tactile vision and related forms of sensory substitution provides strong support for the enactive view. As O'Regan and I have argued, they provide evidence for the view because they illustrate that perceptual experience depends constitutively on the exercise of sensorimotor knowledge (O'Regan and Noë 2001a,b; Noë 2002a; see also Hurley and Noë 2003a,b). Tactile vision vision-like because (or to the extent that) there is, as it were, an isomorphism at the sensorimotor level between tactile vision and normal vision. In tactile vision, movements with respect to the environment produce changes in stimulation that are similar in pattern to those encountered during normal vision. The same reservoir of sensorimotor skill is drawn on in both instances.

The enactive view, in turn, exhibits the sort of principles of embodiment that place constraints on what degrees of similarity of body are required to achieve similarity of experience. Tactile vision is vision-like to the extent that there exists a sensorimotor isomorphism between vision and tactile vision. But tactile vision is unlike vision precisely to the extent that this sensorimotor isomorphism fails to obtain. It will fail to obtain, in general, whenever the two candidate realizing systems differ in what we can think of as their sensorimotor multiplicity (i.e., in their ability to subserve patterns of sensorimotor dependence). TVSS and the human visual system are very different in respect to their sensorimotor multiplicity. Compare the crudity and simplicity of the vibrator array in TVSS with the refinement and complexity of the retina. Only a vibrator array with something like the functional multiplicity of the retina could support genuine (full-fledged, normal) vision. To make tactile vision *more* fully visual, then, we need to make the physical system on which it depends more like the human visual system.

In this way, the charge of sensorimotor chauvinism can be answered. Insofar as the enactive approach is willing to count TVSS as quasi-visual, the charge of chauvinism can hardly be made to stick. Nevertheless, differences in body make for differences in sensorimotor skills and in experience. It is not chauvinism to recognize that there will be qualitative differences between TVSS and vision owing to the different ways these systems are embodied.

1.7 A Psychology of the Personal Level?

There is a further enactive challenge to the computer model of mind. Computational theories of vision, for example, model vision as a computation implemented in the brain. Such theories attempt to explain, in the domain of vision, how the brain, which is merely a "syntactic engine," can come to function as a "semantic engine," that is, how it can, for example, produce a detailed representation of the scene on the basis of meaningless patterns of light hitting nerve endings (Dennett [1981] 1987). As Dennett ([1978] 1981) has argued, one of the chief fruits of the computational approach, as a framework for philosophical and empirical investigation of mind, is that it provides, or at least seems to provide, an account of how the brain performs these computational functions, and it does so in a way that satisfies two apparently incompatible desiderata. First, the computational approach explains how the brain gives rise to perception, but it does so not in the idiom of neuroscience (e.g., in terms of action potentials, etc.), but rather in the apparently personal-level idiom of intentional ascription (e.g., in terms of signaling, representing, inferring, guessing, etc.). Second, the computational approach manages to satisfy the first desideratum *without* committing the homunculus fallacy (Kenny 1971 [1984], 1989; Searle 1992; Bennett and Hacker 2001). How is the computational approach supposed to achieve this?

The point of the first desideratum is clear. The alternative to deploying a richly intentional idiom to explain what the brain does, in Dennett's words, "is not really psychology at all, but just at best abstract neurophysiology—pure internal syntax with no hope of semantic interpretation. Psychology 'reduced' to neurophysiology in this fashion would not be psychology, for it would not be able to provide an explanation of the regularities it is psychology's particular job to explain: The reliability with which 'intelligent' organisms can cope with their environments and thus prolong their lives" ([1978] 1987, 64). The point of the second desideratum is equally clear. We won't have succeeded in explaining anything if, in describing the brain in an intentional idiom, we tacitly assume that the subsystems of the brain have the very cognitive powers we are seeking to explain. The solution, according to Dennett, is the insistence that we do not suppose that the internal subsystems have the very powers we seek to explain. Rather, we suppose that they have powers like ours, but simpler. The intuition is that we can decompose the system into homunculi whose powers are so simple as to be, plausibly, powers of the neurons themselves.

Searle has criticized this account of the foundations of the computational theory on the grounds that it confuses the claim that the lowest level of homunculi perform *very simple* functions with the claim that they perform *semantically innocent* functions (Searle 1992). Insofar as we view these maximally simple homunculi as performing functions of symbolic significance, then there's nothing semantically innocent about them.

Whether or not we find Searle's criticism plausible, it seems that from the standpoint of the enactive view at least (which is not Searle's standpoint), Dennett's proposed solution may not avail. Dennett argues that we can explain the brain's semantic powers without attributing non-dischargable semantic powers to the brain's subsystems. But according to the enactive view, perception isn't something that unfolds in the brain *however characterized*, whether in information-processing terms, or those of neurophysiology. It is not the brain, it is the animal (or person), who sees. It's the person, not the brain, that has semantic powers. In a sense, then, the homuncular decomposition never succeeds in discharging the biggest subpersonal homunculus of them all—namely, the brain itself—for the computational approach never allows us to discharge—or better, free ourselves from—the idea that we are analyzing the semantic powers of the brain.

I take it that this is the significance of Nakayama's (1994) remark, mentioned earlier, regarding Marr's oversimplification of the computational problem of vision. Vision shouldn't be thought of as a computation performed by the brain on inputs provided by the retina. What is vision? How should it be characterized computationally? This book suggests the outlines of an answer. *Vision is a mode of exploration of the environment drawing* *on implicit understanding of sensorimotor regularities* (O'Regan and Noë 2001a,b). To model vision correctly, then, we must model it not as something that takes place inside the animal's brain, but as something that directly involves not only the brain but also the animate body and the world.

I have been making use of Dennett's distinction between the personal and the subpersonal (Dennett 1969). But it now appears that we cannot make quite the same use of the distinction that McDowell (1994b) and others have suggested. McDowell sought to reconcile Gibsonian and computational approaches to vision by suggesting that the former provides a theory of vision *at the personal level*, while Marr and the computationalists are concerned with modeling subpersonal processes, that is, the processes that causally underpin and enable the person to see. The flaw in this proposed rapprochement is this: If Gibson is right that the subject of perception is the whole animal, actively exploring its environment, then Marr's characterization of vision at the subpersonal level must be wrongheaded, for he characterizes the subpersonal processes not merely as contributing to the enabling of seeing, but as constituting seeing itself.

The upshot of these reflections, however, is not that we need a theory of perception at the personal level. Dennett insists that this can't be done and he suggests that Ryle, Wittgenstein, and Gibson are *anti-science* in the end because they insist that the only satisfactory account must be at the personal level. Whether or not this is right, I am now inclined to agree with Fodor that the distinction between the person and the subpersonal causal processes enabling mental life may not matter for cognitive science, or may not matter nearly as much as McDowell and others have thought: "Whatever the relevance the distinction between states of the organism and states of its nervous system may have for *some* purposes, there is no particular reason to suppose that it is relevant to the purposes of cognitive psychology" (Fodor 1975, 52). The reason for this is that it turns out that it's not possible to draw a sharp line between what is done by the person, or animal, and what is done by the subpersonal system, or by parts of the animal. This is not to say that there are no straightforward cases. I see. My heart pounds. *I* don't pound my heart. On the other hand, some of the time when my eyes move, it is I who move them, and very often, even if I am not directing their movements, I make use of their movements to keep track of what's going on around me. When my eyes move, whether they move as a result of volition or not, they give rise to changes, some of which (changes in how things look) I may be aware of, and others of which (changes in patterns of retinal activity) I am not. Yet even the subconscious changes (subconscious because subpersonal) may matter to me and impinge on my awareness. As a perceiver I understand, implicitly, how to modulate them. For example, when I cup my ears to hear something better, I modulate receptor-level events to which I have no direct access. But I cup my ears precisely in order to do this, that is, to increase the intensity of stimulation in my ears. Consider the pounding of my heart again. If I am a long-distance runner, then I am used to a certain kind of increased level of pounding. If my heart were to pound that way when I was at rest though, that would be alarming. The point is, as a runner, I have some degree of access to, and control over, subpersonal processes within me. To some extent, my skills as a runner comprise the ability to make my body do this and that. In general, I depend on my subpersonal parts, not merely causally, but constitutively. For I am—we are—beings whose minds are shaped by a complicated hierarchy of practical skills. Our consciousness frequently does not extend to what is going on in our bodies; our consciousness is enacted by what we do with our bodies.

This is not to deny that a distinction can be drawn between the personal and the subpersonal. When I attribute a psychological state to you, it is plausible that I view you as subject to, as it is said, normative constraints of rationality and the holism of the mental. Only a person with a modicum of rationality and a wealth of background knowledge can have, for example, the thought that he or she would like to be rich. And when I attribute to your brain a certain level of activity (say, on the basis of a functional magnetic resonance imagery [fMRI] scan), I do so without regard to such constraints. What you believe or want or expect has no bearing on my attribution of blood-flow activity to your brain on the basis of fMRI.

The understanding of concepts is usually supposed to be a paradigm of personal-level accomplishment. But just as there is no sharp line between the personal and the subpersonal, so there may be no sharp line between the conceptual and the nonconceptual. Indeed, it may be that sensorimotor skills deserve to be thought of as primitive conceptual skills, even if, as is frequently the case, they are subpersonal. I take this up again in chapter 6.

For these reasons, it seems, a theory of perception must straddle the divide between the personal and subpersonal, just as it must straddle the divide between what is conscious and what is unconscious, and what is conceptual and what is nonconceptual. What will such a theory look like? This book is meant to be a step toward an answer.

1.8 Behaviorism Revisited?

I conclude this chapter by considering an objection that may have occurred to the reader. Isn't the kind of identification of perception and action that gets made in this book a form of behaviorism? Experience is *not* something we do; it is something we undergo, something that happens in us! Block (2001), for example, has argued that O'Regan and I are behaviorists because we hold that to have an experience is to partake in certain patterns of input-output relations.

In order to answer this charge, let's consider a different kind of example.

Suppose you hear me say: "Nein!" How do you experience what I say? If you know German, and if the context is right, you may experience me as saying the German word for "No." If you do not know German, but only English, and if the context is different, you may understand me as saying the English word for the number 9. Depending on what you know, and depending on the context, one and the same accoustic phenomenon will lead to very different experiences in you. How you experience my utterance depends not on what you do, but on what knowledge you bring to bear in "making sense" of the stimulus. It is of course true that, given what you know and what knowledge you make use of, your experience of understanding me will dispose you to act in different ways. You will be disposed to reply in some way or other, for example, and the character of your disposition will differ depending on your experience. But it would be a mistake, I think, to say that your experiencing the word one way or the other is simply a matter of your different dispositions. That is the mistake of behaviorism.

According to the enactive approach to perceptual experience, there is all the difference in the world between experiencing the red of a flower, or the shape of a sculpture, and merely having behavioral dispositions. How you experience the flower or the sculpture depends on your perceptual knowledge and on the skill with which you bring this knowledge to bear on what you encounter. As in the linguistic case described earlier, the behaviorist is right that to differences in experiences there correspond differences in behavioral dispositions (other things being equal). But from this it doesn't follow that there is no experience. The enactive view certainly does not embrace the behaviorist's denial of experience. Far from it. As we will see, one of the central aims of this book is to investigate the phenomenology of perceptual experience.

As O'Regan and I stressed in our (2001b) reply to Block, the key to our theory is the idea that perception depends on the possession and exercise of a certain kind of practical knowledge. This is not a behaviorist thesis.[20]

1.9 The Book in Outline

I propose that to perceive is not merely to have sensation, or to receive sensory impressions, it is to have sensations that one understands. The aim of this book is to investigate the forms this understanding can take. There are two main kinds here, although, as I have indicated, there may be no sharp line to be drawn between them. First, there is sensorimotor understanding. Second, there is conceptual understanding. I have said little about the second kind so far. I turn to a discussion of it in chapter 6.

The main argument begins in chapter 2, whose topic is the phenomenology of perception. I argue, on phenomenological grounds, that the content of perception is not like the content of a picture. In particular, the detailed world is not given to consciousness all at once in the way detail is contained in a picture. In vision, as in touch, we gain perceptual content by active inquiry and exploration. When we see, for example, we are not aware of the whole scene in all its detail all at once. We do enjoy a sense of the presence of a whole detailed scene, but it is no part of our phenomenology that the experience represents all the detail all at once in consciousness. The detail is experienced by us as *out there*, not as *in our minds*.

This gives rise to a puzzle. How can we explain our sense, now, of the presence of the whole scene, if we do not actually represent the scene now in full detail the way a picture does? In what does our sense of perceptual contact with the dense and detailed environment consist? I call this the puzzle of perceptual presence. In the course of developing a solution to this proposal, I lay out the enactive (what O'Regan and I have called the sensorimotor) approach to perception. I argue, in particular, that our sense of the presence of detail is to be understood in terms of our *access* to detail thanks to our possession of sensorimotor skill.

The heart of the book is chapters 3 and 4. In these chapters I argue that perceptual experience acquires content thanks to our possession and exercise of practical bodily knowledge. In chapter 3 I focus on the problem of spatial content. The focus of chapter 4 is the experience of color.

In chapter 5, I consider the so-called causal theory of perception. This is a theory of the role of *causation in perception*. I try to show that by emphasizing the role of *action in perception*, the causal theory can overcome important obstacles. But the more far-reaching conclusion of this chapter is that what philosophers call the representational content of experience must be understood to include a *perspectival* aspect. This perspectival aspect marks the place of action in perception. To perceive, we need to keep track of our movements relative to the world. This perspectival aspect belongs to what is experienced.

Perceptual experience is radically ambiguous. The question *What do we experience?* always admits different answers. When we see, we see both how things are, and also how they appear to be. But these are not always the same. For example, we see that the plate is circular, and that it looks elliptical from here. This ambiguity is the source of two important puzzles in the theory of perception, one philosophical and the other psychological. The psychological puzzle is that of perceptual constancy—that is, the phenomenon exemplified by such a fact as that then when you take a book outdoors it does not appear to change color even though the character or the light it reflects changes radically. The philosophical puzzle is that of direct perception, that is, whether the direct objects of perception are mental items such as "sense data." These are puzzles about perceptual content. In chapter 6 I suggest that their solution may turn on an assessment of the place of thought in experience.

Chapter 7 takes up the question of perceptual experience and the brain. In this final chapter I explore the implications of the enactive approach for understanding the brain basis of perceptual consciousness.