The New Science of the Mind

From Extended Mind to Embodied Phenomenology

Mark Rowlands

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

Contents

1 Expanding the Mind

1 The Expanding Mind?

There is a new way of thinking about the mind and things mental that has started to seep out of the ivory tower and set up residence in popular consciousness.[1] Actually, to call it a *new* way of thinking about the mind is not entirely accurate. It is an old way of thinking about the mind that has taken on new form. Previously the preserve of a few scattered, and distinctly renegade, philosophers and psychologists spread out over the centuries, this way of thinking about the mind has started to acquire what many would regard as more persuasive credentials. This is because it is now emerging, in a reasonably consistent and recognizable way, from the confluence of various disciplines in cognitive science, including situated robotics and artificial intelligence (Webb 1994; Brooks 1994; Beer 1995), perceptual psychology (O'Regan and Noë 2001; Noë 2004), dynamical approaches to developmental and cognitive psychology (Thelen and Smith 1994), and cognitive neuroscience (Damasio 1994).

Some people think—and, for what it's worth, I am one of them—that upon this new way of thinking about the mind will be built a new *science* of the mind. The new science in question will employ different methods for studying the mind, and will supply explanations of mental processes that are, at least in some ways, quite different from traditional accounts. But these transformations in methods and forms of explanation are just symptoms of something far deeper and more important. Fundamentally, the new science would be *new* because it is underwritten by a novel conception of what sort of thing the mind is. The subject matter of this book is not the mooted new science of the mind but the conception of the mind that underlies it—a conception that can, I think, be evaluated independently of whether anyone actually gets around to making a science of it.

Traditional attempts to study the mind are based on the idea that whatever else is true of mental processes—perceiving, remembering, thinking, reasoning, and so on—they exist in brains. Mental processes are either *identical* with brain processes or *exclusively realized* by brain processes (see box 1.1). The word "traditional," here, is slightly idiosyncratic. The scientific study of the mind is not much more than a hundred years old, and in that time it has undergone several significant transformations: introspectionism, gestalt psychology, behaviorism, and finally, from the early 1960s onwards, cognitive science. Cognitive science, in its traditional form, is based on the idea that mental processes—specifically cognitive processes, for these are the purview of cognitive science—are abstract "programs" realized in the "hardware" of the brain (an analogy with computers guided much of the early work in cognitive science). The principal tasks of cognitive science are, accordingly, to identify the programs (cognitive psychology) and work out how these programs are implemented in the brain (cognitive neuroscience). For reasons that will be fully explained shortly, I am going to refer to cognitive science, in its traditional form, as *Cartesian cognitive science.*

Cartesian cognitive science is, in many respects, a broad church. There are many important differences in the way cognitive science has developed over the years. For example, in early cognitive science, the emphasis was very much on the "programs" or cognitive "software"; early cognitive science understood itself as engaged in the task of providing abstract formal descriptions of cognitive processes. However, from the mid-1980s on, this emphasis gradually gave way to a renewed emphasis on "hardware" in the form of *connectionist* or *neural network* approaches: approaches to understanding cognition based on neurally realistic models of its underlying architecture.[2] Neural network models base their accounts of cognition on a hardware that is explicitly (if roughly) modeled on the brain (Rumelhart, McClelland, and the PDP Research Group 1986). It is not clear that these two approaches are incompatible. It may be that neural network models are merely accounts of how the more abstract formal descriptions of cognitive processes come to be implemented in the brain. However, nor it is clear that these two approaches are compatible: it may be that neural network models have properties that preclude their being described at a more abstract level by formal descriptions of the sort typically employed.[3]

We do not need to worry about the details of this dispute. What unites these differing faces of Cartesian cognitive science is an unquestioned—indeed seemingly banal—assumption: whatever else is true of mental

processes, whether they are abstract formal processes or patterns of activity in a neural network (or both)—they are processes that occur inside the head of the thinking organism. Cognitive processes—the category of mental processes with which cognitive science is concerned—occur inside cognizing organisms, and they do so because cognitive processes are, ultimately, brain processes (or more abstract functional roles realized exclusively by brain processes). It is this unquestioned assumption that makes Cartesian cognitive science Cartesian. And it is this assumption that I shall try to undermine.

The new way of thinking about the mind is inspired by, and organized around, not the brain but some combination of the ideas that mental processes are (1) *embodied*, (2) *embedded*, (3) *enacted*, and (4) *extended*. Shaun Gallagher has referred to this, in conversation, as the *4e* conception of the mind.[4] The idea that mental processes are *embodied* is, very roughly, the idea that they are partly constituted by, partly made up of, wider (i.e., extraneural) bodily structures and processes. The idea that mental processes are *embedded* is, again roughly, the idea that mental processes have been designed to function only in tandem with a certain environment that lies outside the brain of the subject. In the absence of the right environmental *scaffolding*, mental processes cannot do what they are supposed to do, or can only do what they are supposed to so less than optimally. The idea that mental processes are *enacted* is the idea that they are made up not just of neural processes but also of things that the organism *does* more generally—that they are constituted in part by the ways in which an organism acts on the world and the ways in which world, as a result, acts back on that organism. The idea that mental processes are *extended* is the idea that they are not located exclusively inside an organism's head but extend out, in various ways, into the organism's environment. We shall examine each of these ideas in much more detail in the following chapters. These characterizations are very rough, and in several ways inadequate; but they will probably give us enough to work with for present purposes.

Each of these ideas—embodiment, embeddedness, enactedness, and extendedness—has been understood as denying, or at least questioning, the central assumption of Cartesian cognitive science: mental processes are identical with, or exclusively realized by, brain processes. It is not clear, however, whether it is *correct* to understand all of the ideas in this way. I shall argue, later in this chapter and in more depth in chapter 3, that not all strands of 4e are equally anti-Cartesian. Moreover, even if it were true that all of these ideas deny the central assumption of Cartesian cognitive science, this denial would take a quite different form in each case. And

Box 1.1
Identity and Exclusive Realization

For the purposes of this book, the difference between identity and exclusive realization is of no real importance—that is why I have hitherto used them in the same breath. But it is probably time to explain why the difference is of no real importance. First of all, identity itself can actually be understood in two ways. Broadly speaking, to say that mental processes are identical with brain processes is to say that they are one and the same thing as brain processes. It is not that there are two things there—mental processes and brain processes—that are correlated; there is only one thing there. However, there are two different ways of thinking about this. According to one, this is a claim about mental and neural processes understood as *kinds*—or as philosophers like to call it, *types*—of process (Smart 1959). So, the claim is that kinds or types of mental process are identical with kinds or types of brain process. This is known as a *type identity theory*. According to this, mental processes are one and the same thing as brain processes in much the same way that water is H_2O, and as lightning is an electrical discharge to Earth from a cloud of ionized water particles.

There is another—currently more popular—way of understanding the identity theory. According to this, the identity between mental and physical holds between individual *instances*—or, as philosophers like to call them, *tokens*—of each kind (Davidson 1970). It is the individual episode of pain that I, a particular person, feel at a particular time (e.g., 4:19 PM on March 21, 2008) that is identical with a particular firing of a brain process (that takes place in me at this time). This is an identity between individual episodes, instances, or tokens, rather than an identity of general kinds or types. This view is known as the *token identity theory*.

The token identity theory has proved more popular for reasons that connect up with the notion of *realization* mentioned above. The idea of realization is drawn from the computer metaphor that dominated early cognitive science (Putnam 1960). One and the same program can be run on different sorts of computer—and these computers can, within limits, be built in different ways. Therefore, we can not identify the program with any particular configuration of hardware. But, nonetheless, the program cannot be run without some hardware or other. Therefore, the idea is that though the program cannot be regarded as the same thing as configurations in the underlying hardware—because it can be run on different hardware—the program is, in any particular case, realized by given type of hardware. To say that A realizes B, therefore, is to say, very roughly, that A makes B possible by providing a physical basis for it.

Box 1.1
(continued)

The token identity theory became popular because it is compatible with the idea that mental kinds are *functional* (see box 2.1) kinds that are, in any given instance, realized by some or other physical hardware—but not necessarily the same hardware in all cases. The idea that mental kinds are functional kinds is the claim that they are best understood in terms of what they *do*. To see what this means, consider another example of something defined by what it does. A carburetor is a physical object located somewhere in the innards of a car's engine (or older cars anyway—fuel injection systems have replaced them in more recent models). What is a carburetor? Roughly, it is (or was) something that takes in fuel from the fuel inlet manifold, takes in air from the air inlet manifold, mixes the two in an appropriate ratio, and sends the resulting mixture on to the combustion chamber. It is fulfilling this role that makes something a carburetor, and anything that fulfils this role in a car thereby counts as a carburetor. Most carburetors tend to look pretty similar. But this is at best a contingent fact, because it doesn't matter what a carburetor looks like as long as it fills this role. The details of its physical structure and implementation are of secondary importance compared to the role it fills, for it is filling this role that makes something a carburetor, and not the details of its physical structure or implementation. Of course, not everything physical thing is *capable* of playing the role of a carburetor. A lump of Jell-O inserted into your car engine would have a hard time mixing fuel and air—or doing anything except melting, for that matter. A lump of Jell-O is simply not the right sort of thing for fulfilling the functional role of a carburetor. So, the details of the how the functional role is physically implemented are not irrelevant. But as long as you have a suitable physical structure—one that is capable of fulfilling the role of a carburetor, then it doesn't matter what it is as long as it, in fact, fulfills this role. If it does, then it is a carburetor.

At the level of tokens, each individual carburetor is a physical object. There is a token identity between a carburetor and a physical thing. However, since carburetors can, in principle, be realized by different kinds of object, there is no type identity between carburetors and specific types of physical object. That is, it is not plausible to say that a carburetor as a kind is the same thing as this kind of physical object, because other kinds of physical objects are capable of playing the role of carburetors—and thus of being carburetors. The kind or type carburetor is *realized* by physical kinds or types, but it is not the same thing as them.

By far the most popular version of materialism over the past few decades has combined an identity of mental and physical tokens with a functionalist

Box 1.1
(continued)

account of mental kinds or types. So, the general idea is that mental process-tokens are *identical* with brain process-tokens, but mental process-types are *realized* by brain process-types. Since, it has almost universally been assumed, there is nothing outside the brain that realizes mental types, this realization is *exclusive* to the brain. Together these claims—token identity combined with exclusive neural realization—form the default view of the nature of mental phenomena that 4e attacks.

once we venture beyond the simple denial, we shall see that the ideas of embodiment, embeddedness, enactedness, and extendedness are far from equivalent. Indeed, not only are these ideas different, some of them may actually be incompatible with the others. Indeed, at least one of the strands of 4e can, and indeed has, been employed as a sort of Cartesian fifth column—one designed to acknowledge the force of anti-Cartesian arguments but also strictly limit their scope. In chapter 3, we shall begin the process of tidying these ideas up: working out exactly what each one claims, the ways in which it denies (if it does) the central assumption of Cartesian cognitive science, and thereby working out which ideas are and which are not mutually compatible.

To avoid these difficulties, and to avoid anticipating the results of arguments yet to be given, I shall, for now, simply talk of *non-Cartesian* cognitive science (or, depending on context, of the non-Cartesian conception of the mind that underlies this science), and understand this as made up of at least some, but not necessarily all, of the strands that make up 4e: embodiment, embeddedness, enactedness, extendedness. The diversity and potential incompatibility of the ideas of embodiment, embeddedness, enactedness, and extendedness do present us with a problem. If a new, non-Cartesian, cognitive science were to be built on these ideas, and if some of these ideas didn't really cohere with the others, then the prospects for our new science would be at best uncertain and at worst bleak. What we need, it seems, is a way of identifying with precision the content of each of these ideas, and, on this basis, putting them together in the best way we can and discarding the ones that cannot be thus assimilated. This, in effect, is where philosophy comes in. It would not be entirely true to say that philosophy has *driven* the development of the ideas of embodiment, embeddedness, enactedness, and extendedness. It is true that the

history of philosophy has contained figures that have developed views of the mind that are entirely amenable to, and may even be (philosophical) versions of these ideas. We shall encounter several of these figures later in the book. However, most of the recent impetus for 4e comes from developments in fields of cognitive science such as situated robotics, developmental cognitive psychology, and theories of visual perception. The primary role of philosophy is not to provide new empirical evidence for non-Cartesian cognitive science, but to place this science on a solid conceptual footing.

To this end, this book aims to identify, clarify, render consistent, and, where appropriate, defend the central concepts on which a non-Cartesian cognitive science is to be based. This process begins in chapter 3. There, we shall examine the ideas of embodiment, embeddedness, enactedness, and extendedness with a view to identifying the content of each idea—what it actually says or claims. Then we shall examine the ways in which each idea fits together with the others and the ways in which it does not. That is, we are going to look at the extent to which each idea either entails the others, or is compatible with the others, or is incompatible with the others. In the process, choices will have to be made. Perhaps one idea will simply turn out to be a version of another. In this case, our non-Cartesian cognitive science may turn out to be 3e (or even less), rather than 4e. More worryingly, perhaps one idea will prove incompatible with one or more of the others. Then, it seems, one of the ideas will have to go, and our task is to work out which one, and how to do justice to the overall framework of the non-Cartesian conception of the mind given these constraints.[5]

Once the process of identification, clarification, and rendering consistent has been completed, we get to the most important part of the project: *defending* the new, non-Cartesian, conception of the mind.[6] In a work of philosophy, this means defending the concepts out of which this new conception is going to be built—the concepts we have succeeded in identifying, clarifying, and rendering consistent. This is the principal task of the book, and most of the book will be devoted to it. None of these tasks is particularly easy. Even if we are witnessing the birth of a new science of the mind, some births are protracted and painful. Nonetheless, this book contends that all these tasks can be successfully completed.

Before we get to them, however, there are a few more preliminary matters that need our attention. The idea that the mind is not "in the head"—that the mind can extend out into the body and even into the world—will strike many as a truly crazy idea that no one who is even remotely sane could ever accept. In the rest of this chapter, I want to

sketch, in a preliminary way, the general motivation for thinking that this idea is not as crazy as it seems.

2 Minds and Mental Phenomena

The new way of thinking about the mind is sometimes characterized as the claim that the mind, or even the self, is *outside the head*. Now that, at least on one way of thinking about the mind or self, *would* be a truly crazy claim. Happily, no version of non-Cartesian cognitive science commits us to this. That is, none of the various strands that make up 4e—the theses of embodiment, embeddedness, enactedness, and extendedness—should be interpreted as saying anything at all about the mind (and a fortiori, the self)—unless you want to think of this as a construction out of mental states and processes. What a non-Cartesian cognitive science is actually concerned with is mental states and processes, and not whatever it is that has them.[7]

Where does your mind begin, and where does it end? This is an unusual question. The usual question, at least if the history of philosophy and psychology is anything to go by, is: what is the mind? And the usual answer is: the brain. If this is right, then your mind begins and ends where your brain begins and ends—for your mind is simply your brain and nothing else. But where does your brain begin and end? People who think that the mind is the brain typically draw a firm distinction between the central and peripheral nervous systems. The brain is the lump of gray, gooey matter located in your head, made up of brain stem, hippocampus, and cerebral cortex. And if that is what the brain is, then that is what the mind is. Or more accurately, your mind is part of this triune structure, the part that is responsible for your being able to think and feel, and so on: your mind is (parts of) your cerebral cortex and hippocampus.

This idea is probably still unacceptably vague. But the most serious problem is with it is this: between you and me, I'm not sure I even have a mind—if this is understood as something different from my mental states and processes. This point, or at least something recognizably similar to it, was made a long ago by the philosopher David Hume:

> Whenever I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never catch *myself* at any time without a perception and never can observe anything but the perception. . . . If anyone upon serious and unprejudiced reflection thinks he has a different notion of himself, I must confess I can reason

no longer with him. All I can allow him is that he may well be in the right as well as I, and that we are essentially different in this particular. He may, perhaps, perceive something simple and continued, which he calls *himself*, though I am certain there is no such principle in me. (1739/1975: 252)

Hume's remarks, here, are addressed to the existence of what he calls the *self*. It is not clear that the self and the mind are the same thing for Hume.[8] For our purposes, however, this does not matter. We can make a precisely parallel point about the mind. When I enter, as Hume put it, "most intimately into what I call myself"—that is, when I turn my attention inward or introspect—I never encounter my mind: all I ever come across are mental states and processes. That is, when I introspect, I might become aware of what it is I am thinking, what it is I am feeling, and so on. I might become aware of my beliefs, my desires, my sensations, my emotions; my hopes, fears, aspirations, and anticipations. But what I never come across is the *subject* of these states and processes, at least if this is understood as something different from these states and processes. I never come across my *mind* thought of as something lying behind these states and processes.

We must be careful here, because we can easily transform this Humean-inspired observation into a very bad argument. Suppose I am looking at something—say, the can of diet Coke that invariably sits next to my laptop when I work, bearer of the psychotropic substance that I need to give my aged and rather reticent brain the morning kick start it needs. I can see the shape of the can, and its shiny silver surface. I can see the writing on that surface in black and red. But do I ever see the can itself? Well, what I don't see is the can as something separate from, or over and above, these properties. I don't see the shape, the surface, the adornment, and then, in addition, see the can. But does this mean I don't see the can? That would be the bad argument advertised at the beginning of this paragraph. I see the can in virtue of seeing the shape, surface, and adornment. In general, we see objects by way of, or in virtue of, seeing their properties.

That is the best way of understanding this Humean-inspired observation. I am not claiming that when you introspect you don't ever encounter your mind. Rather, I am making a point about what it is to introspectively encounter your own mind: to encounter your mind is to encounter your mental states and processes. You are aware of your mind in virtue of being aware of your mental states and processes, just as you are aware of the can of diet Coke in virtue of being aware of the shape, surface, and adornment of the can.

Nevertheless, lack of subtlety often has a habit of creeping back into our thinking. Often, when we think of the mind we have an idea of something that underlies all our mental states and processes: something to which those states and processes attach—something that holds them all together. Similarly, when we think of the can, we typically think of it as something that underlies its various properties—as something to which those properties attach. To avoid this questionable way of thinking about the mind—the mind as a *bare substratum*—I am, in the rest of the book, going to do my best not to talk about the mind at all (unless I am talking about someone else who is talking about the mind in this sense). I am only going to be concerned with mental states and processes: mental *phenomena* broadly construed. The non-Cartesian conception of the mind defended in this book is not, in fact, a conception of the mind at all—not if we understand this as something underlying mental states and processes. It is a conception of mental phenomena: it claims that at least some mental phenomena are either embodied, embedded, enacted, or extended. In this, it rejects the conception of mental phenomena embodied in Cartesian cognitive science: mental states and processes begin and end with the brain, because mental states and processes are *identical* with, or *exclusively realized by*, brain states and processes.

The view that mental states and processes are identical with, or exclusively realized by, brain states and processes is not a peculiar eruption of contemporary thought. On the contrary, it derives from a view of the mind that emerged in seventeenth-century France.

3 Descartes's Ghosts

It is easy to understand why the ideas of identity of mental and neural tokens combined with exclusive realization of mental types by neural types assumed such position of dominance in scientific thinking about the mind. To deny it seemed, at least at first glance, to commit you to an untenable position: to a form of *dualism* about the mind. Dualism received its most famous development at the hands of the philosopher, mathematician, and sometime mercenary René Descartes.

According to Descartes, the mind is a nonphysical *substance*. Today, we tend to use the word "substance" to mean something like "the stuff from which a thing is made." However, Descartes inherited his use of this term from medieval philosophers, and, for them, "substance" meant "thing" or "object." So the mind, according to Descartes, is a nonphysical object. In some respects, Descartes thought of the mind as similar to other bodily

organs. The heart, liver, kidneys, and so on, are all objects found in the body and, more importantly, they are objects of a certain sort: ones defined by their *function* or *what they are supposed to do*. The function of the heart is to pump blood around the body; the function of the liver is to regulate metabolism; the function of the kidneys is to process waste products; and so on.

According to Descartes, the mind, like other things found inside the body, is defined by its function: and its function is to *think*. However, he argued that there is a crucial difference between the mind and all other organs found inside the body: the mind is a *nonphysical* substance. By this, he meant, fundamentally, that the mind is *nonspatial*. Physical things, Descartes claimed, have a single defining feature: *extension*. By this he meant that physical things *occupy space* or *take up room*. This, he thought, is precisely what makes them physical things. Therefore, the mind, being nonphysical, must also be nonspatial.

However, there are actually two different aspects of the idea of space, and Descartes never clearly distinguished between them. On the one hand, there is spatial *extension*—taking up room. On the other hand, there is spatial *location*—existing *at* a particular place. If an object has spatial extension then it must also have spatial location. It is not possible for an object to occupy space without the space that it occupies being somewhere or other (even if this "somewhere or other" is rather vague). But just because an object has spatial location, it does not necessarily follow that it has spatial extension. For example, scientists are willing to countenance (and some even insist on) the existence of *point particles*: particles that exist at some particular place but do not take up any room. Even if they ultimately turn out to be wrong, their willingness to countenance the possibility shows, at least prima facie, that the *ideas* or *concepts* of spatial extension and spatial location are not the same. The idea of spatial extension seems to entail the idea of spatial location. But the idea of spatial location does not entail the idea of spatial extension.

Although Descartes never clearly distinguished between the ideas of spatial extension and spatial location, this distinction can be used to make sense of Descartes's position.[9] In effect, Descartes's view was that minds do not have spatial extension—this is what makes them nonphysical—but they do have spatial location. Every mind is located inside a (functioning) brain, and every (functioning) brain is located inside a body. Descartes was never exactly clear on the precise location of the mind; but somewhere in the vicinity of the brain's pineal gland seemed to be his favored hypothesis.[10]

Descartes's view is known as *dualism*—since it asserts that each one of us is composed of two different kinds of thing: physical bodies and nonphysical minds. Despite making something of a recent, limited comeback (e.g., Chalmers 1996), dualism is still almost certainly one of the most reviled philosophical views ever invented. Generations of professional philosophers have spent much time and energy (a) showing that Descartes's arguments for dualism don't work, (b) arguing that dualism itself has empirical and conceptual difficulties so serious as to render it an effectively untenable position, and (c) inventing catchy slurs with which to disparage his view—most famous of which is, perhaps, Gilbert Ryle's (1949) dismissal of the view as the dogma of the *ghost in the machine*. Today, even the word "Cartesian" is often used as a term of abuse.

The popularity of the mind–brain identity/exclusive neural realization combination stemmed, in large part, from the belief that to deny it was to be committed to dualism of a broadly Cartesian sort. This, we are now beginning to understand, is not true. Indeed, what has been overlooked until very recently is just how much the mind–brain identity theory/exclusive neural realization combination has inherited from Descartes: they are, in effect, fashioned in the image of the Cartesian conception of the mind.

The Cartesian conception of the mind in fact has two distinguishable aspects. First, there is the claim that the mind is a nonphysical thing. Second, however, there is the idea that the mind is something that exists inside the head. So, when Ryle dismissed Descartes's view as the myth of the ghost in the machine, this dismissal actually has two distinguishable aspects. First, Ryle was rejecting the idea that the mind is a *ghost*—that is, a nonphysical thing. But second, and for our purposes more significantly, he was rejecting the idea that the mind is the sort of thing that can be found *inside* the bodily machine. If you reject only the first idea, then you have not fully rejected the Cartesian conception of the mind, but only part of it. And that, in effect, is precisely what the mind–brain identity/exclusive neural realization combination did. It rejected Descartes's idea that the mind is ghostly or nonphysical, but it left intact the second defining idea of the Cartesian conception: the idea that the mind is something that exists *inside* the head. In other words, the mind–brain identity/exclusive neural realization model is a view fashioned partly, but as things turned out, decisively in the image of the Cartesian view of the mind.

Non-Cartesian cognitive science is based on a more complete rejection of the Cartesian view of the mind. This science is, of course, materialistic: there will be no reversion to nonphysical substances—that particular

Cartesian ghost remains well and truly exorcised. However, non-Cartesian cognitive science also rejects, or is at least thought of as rejecting, Descartes's second idea, the idea inherited by the mind–brain identity/exclusive neural realization model. That is, it rejects the claim that mental states and processes occur purely inside the brains. Some of them do; but not all of them do. Mental states and processes are not just things that happen inside our brains; they are also things that happen, *partly*, in our bodies and even, *partly*, in the world outside of our bodies. The qualification "partly" is (i) obvious, (ii) crucial, and (iii) ignored with surprising frequency. With (iii) in mind, let me risk being tediously overemphatic. No one is going to claim that there can be free-floating mental processes hovering around in the world outside the head. That *would* be an example of a truly crazy idea that no remotely sane person would want to hold. Almost as insane would be the idea that there can be mental processes that are entirely made up of processes occurring inside the body but outside the brain. No one wants to maintain that either. The idea, rather, is that in the case of some mental processes but not all, *part* of that mental process—but *never* all—is made up of factors that occur outside the brain of the subject.

Our next obvious question is: why think this?

4 Best Friends and Barking Dogs

The starting point for non-Cartesian cognitive science is the extent to which we make use of things around us in order to solve problems and get things done. Our tendency to do so has become more and more obvious in recent years. Consider, for example, a relatively new acquaintance of mine: my car's GPS (global positioning system). Actually remembering how to get to a destination: that was *so* 2007 (or, for more those of a less technophobic persuasion than I, so 2005). In 2009, my GPS will tell me how to get anywhere I want to go. "Make a safe and legal U-turn," the reassuring, and apparently vaguely stoned, voice tells me. OK, will do.[11]

The knowledge afforded me by my GPS is essentially *situated*. That is, it provides me with practical and easily digestible instructions because it uses my physical location to encode at least some of the information I require to follow those instructions. The information with which my GPS provides me is, in this sense, *indexical*: it has a meaning that is partly made up of is made up of the meaning of words such as "here," "there," "this," "that," and associated locutions. Because of this the instructions, we might put it, *show* more than they *say*. If the instruction is "make a safe and legal

U-turn": then that is precisely what the instruction *says*. But what it shows is that you should make the U-turn at the next intersection—which is "there," in front of you.[12] The instructions show more than they say because part of their informational content—part of their *meaning*, broadly construed—is encoded in my physical location at the time the instructions are given.

Think about the difference between my GPS and its precursor: MapQuest. For our purposes, both the similarities and differences between the GPS and MapQuest are important. Consider, first, the similarities. Both the GPS and MapQuest are external forms of information storage—external to *my body*. I don't need to remember how to get somewhere, because the information about how to get there is contained in both the GPS and in MapQuest. These external forms of information storage, therefore, reduce the burden on my biological memory. Memory tasks that I would find difficult are offloaded onto the environment.

The differences between the GPS and MapQuest concern how this external information is stored. MapQuest works by providing you with an *algorithm*: a set of instructions such that if you follow them faithfully you will arrive at your destination—at least in theory. Leaving your residence, you turn west on 156th. Then at 77th you turn north. At the second light, you turn east on 144th, and so on. When MapQuest wants you to turn left, it needs to provide you with information that specifies precisely where you are to do this (e.g., at the intersection of 77th and 144th). There is no need for your GPS to do this. With your GPS, the time and place of the utterance is inextricably linked to the meaning of the utterance. The information contained in the GPS instructions is situated in a way that the information contained in the MapQuest algorithm (typically) is not. The GPS information is both external and situated. The MapQuest information is external but not situated (or, at least, not *as* situated).

We all know the problem with MapQuest directions. The chances are there are too many instructions for you to remember, and so you have to take them with you in printed form. And then, consulting your sheet to find out what you should do next, you slam into the back of the Camry that's stopped in front of you. Your GPS, on the other hand, gives you easily digestible snippets of information as and when you need them—which it can do precisely because the information it provides is situated. Nevertheless, although they accomplish it in different ways, both the GPS and MapQuest have essentially the same function. You need to accomplish a task—getting from A to B. If you couldn't avail yourself of MapQuest or a GPS or some other form of external information storage system, then

you would have to remember how to get from A to B "all by yourself." That is, the task would have to be entrusted to your naked biological—that is, neural—memory. However, to the extent that you are able to avail yourself of an external information source, the complexity, and therefore difficulty, of the task you have to accomplish in your head is correspondingly reduced. The task that you would have had to accomplish in your head is, in part, *off-loaded* onto the environment. Equivalently, we might say that the task is *distributed* onto the environment.

It is the possibility of this sort of off-loading that provides the starting point for the new science of the mind. However, recent technological developments of the sort that result in MapQuest or GPS systems are only the most recent manifestations of a process that began almost as soon as humans became humans. Human cultural development is, in part, a process of creating external information-bearing structures: structures that could be used to enhance our ability to accomplish important tasks. One of the most significant of these structures is the development of written language. In a classic early study, the Soviet psychologists Alexander Luria and Lev Vygotsky (1930/1992) identified the implications of this development for human biological memory.

Consider an early form of written language—a simple system of visuographic representation. Imagine two people (the example is Luria and Vygotsky's). One is an African envoy, entrusted with the task of remembering word for word the message of his tribal chief. The other is a Peruvian *kvinu* officer—an "officer of knots"—who uses a conventional system of knots tied in string to achieve the same purpose. As Luria and Vygotsky point out, for each new message, the African envoy must employ anew his (biological) memory. The *kvinu* officer, on the other hand, need employ his memory resources only once—in learning the "code" that allows him to access the information contained in the knots. Once he has done this, a potentially unlimited amount of information becomes available to him since a potentially unlimited number of knots can be presented to him. The African envoy faces a comparatively difficult task. But, for the *kvinu* officer, part of the difficulty of this task has been off-loaded onto the environment. By making use of an appropriate environmental structure, the *kvinu* officer significantly reduces the complexity and difficulty of the task that he must accomplish in his head (see also Rowlands 1999, 134–137).

Suppose you have to do something; it doesn't matter what. And suppose successfully doing this thing requires you to put in a certain amount of work. Then, this is a truism: if you can get something (or someone) else

to do some of this work for you, you will have correspondingly less work to do yourself—as long, and here is a crucial caveat, as the work you put into to getting that something or someone to do the work for you is less than the work they thereby do. There is an old adage that captures this idea quite nicely. It says: why keep a dog if you are going to bark yourself? If there is some barking to be done, and you have a dog that will do at least some of it for you, then you have correspondingly less barking to do yourself (Rowlands 1999, 79–80). At the conceptual heart of non-Cartesian cognitive science is this *barking dog principle*.[13]

Any cognitive science, whether Cartesian or non-Cartesian, will be concerned with, at the very least, the following sorts of tasks: (1) *perceiving* (visually, aurally, etc.) the world, (2) *remembering* perceived information, (3) reasoning on the basis of information perceived or remembered, and (4) expressing this information, or listening to information expressed by others, in the form of *language*. Perceiving, remembering, reasoning, and the processes involved in the production and comprehension of language may not exhaust the legitimate business of cognitive science; but they certainly lie at the core of this science.

The central idea of non-Cartesian cognitive science is that cognitive tasks are not, in general, the sort of thing that need be accomplished only in the head or by a brain. If we had to accomplish a cognitive task using only our brains, then complex and difficult neural maneuvers might be required. However, if we are able to use relevant structures in our environment, then some of the complexity and difficulty of this task might be reduced: we off-load at least some of the task onto the environment around us. Roughly: we get the environment to do some of the work for us; and this reduces the work that we need do. What makes an environmental structure relevant? Again roughly: it carries information relevant to the task that we need to accomplish, and by using this structure, or acting on it in the right sort of way, we are thereby able to appropriate—make available—and employ this information in the accomplishing of the cognitive task in question.

If this sounds a little abstract, just think of our MapQuest instructions that we have printed out prior to our journey. This printout contains information relevant to the task we need to accomplish—getting from A to B. I am able to make this information available and use it by manipulating the page in the right way—for example, by picking it up, turning it the right way up, and holding it in front of my eyes. In the absence of this sort of manipulation, the information I require is present on the page but unavailable to me. Because I am able to act on the external structure—the

page—in such a way as to make the information it contains available to me, the number and complexity of neural operations that I must perform to accomplish the task of getting from A to B is accordingly reduced (given that we are encultured in linguistic communities, we generally find reading novel information easier than remembering it). Acting on external structures in such a way that the information they contain is transformed from the merely *present* to the *available* lies at the heart of non-Cartesian cognitive science. This sort of action is, according to the non-Cartesian conception of cognition, *part* of what cognition is. That is, action of this sort—action that transforms information that is merely present in external structures to information that is available—forms a properly cognitive part of an overall cognitive process.

Consider another example that makes much the same point. Imagine how fiendishly difficult it would be to do jigsaw puzzles if we weren't allowed to pick up and manipulate the pieces (Kirsh and Maglio 1994). If we couldn't do this, then we would have to form detailed mental images of each individual piece, and by a process of mental rotation try to work out with which other pieces it fits. In such circumstances, jigsaws would be even less fun than they already are. But this, of course, is not how we do jigsaws. We don't do them this way because the pieces themselves contain information about which of the other pieces they are going to mesh with and which they are not. By picking up individual pieces and "trying them out"—bringing them into close proximity with other pieces and adjudicating the likelihood of their fitting together—we make this information available to us. Therefore, we don't need to construct this information in the form of mental images and processes of mental rotation. We get the world to do some of the work for us, and we do this by acting on the world—manipulating and exploiting structures that contain useful information and thereby making this information available to us. Each piece of the jigsaw contains information that uniquely specifies the pieces with which it will fit. By manipulating the pieces, we transform this from information that is merely present to information that is available. And this, according to non-Cartesian cognitive science, is, in part, what cognition is.

5 Non-Cartesian Cognitive Science: Framework

We are beginning to glimpse at least the general contours of the non-Cartesian conception of mental phenomena that underlies the prospective new science:

1. External structures carry information relevant to the accomplishing of a given cognitive task (or task that has a cognitive component). This information is *present* in these structures.
2. By using such structures in appropriate ways, I can transform the information they contain from information that is merely present to information that is *available*—available for detection by my sensory apparatus, and for deployment by my subsequent cognitive operations.
3. The information thus made available is, therefore, information that I need only *detect* rather than information I need *construct* or *store*.
4. Detection of information is *cheaper* than construction or storage.

And, finally, the *money ball*:

5. Action on external structures that transforms the information they contain from the merely present to the available (and so amenable to detection rather than requiring construction or storage) is part of what cognition is.

To see these principles in action, let us once more revisit MapQuest. The external structure in question is the printout of the directions. This printout contains information relevant to the accomplishing of a given task: getting from A to B. This information is *present* in the printout. Moreover, this is a task that is part locomotive, but also part cognitive (that is, it has a clear cognitive component). Therefore, condition 1 is satisfied. By using the structure in appropriate ways—for example, by picking it up, turning it the right way up, and holding it up in front of my eyes—I transform the information contained on this page from information that is merely present to information that is *available* for detection by me and deployment in my subsequent cognitive operations. This is condition 2. The information that is made available is information that I need only detect, for example, through perception. I do not need to store this information in the form of memory. Nor do I need to construct this information by way of a process of reasoning. This is condition 3. We find perceiving of information easier to accomplish than storing or constructing it. Typically—not always, not necessarily, but typically—perceiving makes fewer demands on our neural apparatus than either reasoning or storage. This is condition 4, and we shall look at reasons for thinking it is true later on. It is condition 5 that is the controversial one. We will look at this principle, and why it is controversial, in the next section.

The first condition of this general framework appeals to external information-bearing structures. In this context, to call something "external" is to say that it is outside the brain (or central nervous system). The structure

in question is not a neural state or process. The structures I have looked at so far are also ones that are outside the body as a whole—GPS devices, printed pages, knots, and so on. However, like its Cartesian counterpart, non-Cartesian cognitive science is a broad church—as we shall see later, perhaps too broad. In some cases, the external structures in question might be bodily ones: things located inside the body of the organism but not part of the brain or central nervous system. This means that the notion of "use" referred to in the second condition must also be understood in a very broad sense. The way in which we use bodily structures in accomplishing our cognitive tasks is typically very different from the way in which we use structures outside our body. We shall return to this point in a later chapter.

At present, however, we need to look at a more fundamental question, one that pertains to condition 5.

6 The Big Question (Or, How Many Es Are Enough?)

It is easy to understand the attraction of utilizing external forms of information storage in the ways we have encountered above. The attraction, bluntly put is: *thinking is hard*. The perils and pitfalls of irresponsible brain-use were, as far as I am aware, first pointed out by the philosopher Alfred North Whitehead, early in the twentieth century:

> It is a profoundly erroneous truism, repeated by all copy-books and by eminent people making speeches, that we should cultivate the habit of thinking of what we are doing. The precise opposite is the case. Civilization advances by extending the number of operations we can perform without thinking about them. Operations of thought are like cavalry charges in a battle—they are strictly limited in number, they require fresh horses, and they must only be made at decisive moments. (Whitehead 1911, 55)

As if we needed reminding, this idea has been the subject of extensive empirical confirmation in recent years (e.g., Baumeister et al. 1998). Cognizing is hard, especially when we use only the naked brain, and is something one should do only at decisive moments. The rest of the time, it is a good idea to *cognitively delegate*. As Andy Clark (1997) once put it, we make the world around us smart so we don't have to be. Underlying this delegation, of course, is the barking dog principle. Delegating works, as long as the work you put into the delegation is less than the work your delegates then actually do for you.

However, this idea begs one crucial question: what reason, if any, do we have for thinking that the sorts of extraneural processes invoked by

non-Cartesian cognitive science are cognitive processes?[14] Take, for example, manipulating the MapQuest printout—picking it up, turning it the right way around, holding it up in front of my eyes, and so on. This is an action that transforms the information merely present in the printout into information that is available. But what reason, if any, do we have for regarding this as a part of my cognitive processing? Intuitively, this seems far-fetched. It seems far more natural to divide up what is going on in the following way. First of all, there are the processes of perceiving, recognizing, and understanding the words contained on the page. These are genuinely cognitive processes, and they occur in the brain. That is, they are processes occurring somewhere in the brain's visual and language modules. The processes are *facilitated* by actions that I perform. It is, of course, much easier to read the page if I turn it the right way up. Everyone knows that. But that doesn't mean that this manipulation forms part of my processes of cognition. The right way to think about things, surely, is that my genuinely cognitive processes—ones occurring in my brain's visual and language modules—are supplemented and supported by actions that I perform on the world. To be supported by actions that I perform on the world is not the same thing as to be partially constituted by those actions (Rupert 2004; Adams and Aizawa 2001, 2010).

We can make the same point about the jigsaw puzzle. The real cognitive processes involved here are ones of visual perception and mental image formation and rotation. Of course, it is true that if I were not able to pick up the pieces and manipulate them, then I would have to engage in far more complex processes of image formation and rotation. The need to do this is considerably reduced by the fact that I can pick up pieces and manipulate them—bring them into close proximity and see if they fit together. Nonetheless, the need for image formation and rotation is not eliminated. At any given time, I select which piece to pick up on the basis of an antecedent sense of whether it is a likely candidate for fit. And I get this sense through prior processes of at least partial image formation and rotation. So, it is natural to think that the process of completing a jigsaw puzzle can be divided into two sorts. First, there are the *real* cognitive processes: perception of the pieces, and subsequent image formation and rotation. These occur exclusively in the brain. Second, there are bodily actions I can perform on the individual pieces: picking them up and "trying them out." These are not real cognitive processes at all; but they do provide a useful supplement to those processes—a supplement that reduces the extent, and perhaps alters the nature, of the real cognitive processes that we must perform.

These questions are obvious ones. The inference from the claim that external operations I perform on the world *supplement* or *complement* my cognitive processing to the claim that these operations therefore form *part of* that processing is unsound. We cannot afford to run together the question of whether things going on in the environment *drive* (i.e., causally contribute to) a cognitive process with the question of whether things going on the environment partially *constitute* a cognitive process. The answer to the former question is, of course, yes. Things going on in the environment do causally drive cognitive processes. Even René Descartes could have agreed to that—assuming he could work out his problems with understanding how mental–physical interaction took place. The idea that things going on in the environment causally drive cognitive processes is an utterly mundane claim that anyone should accept. Cartesian cognitive science would not only accept this idea, it would insist on it.

Indeed, not only is this question an obvious one that must be answered in any development of a non-Cartesian alternative, it also makes it clear why we shouldn't be in any rush to blithely identify non-Cartesian cognitive science with 4e. In effect, the question begins the process of whittling down 4e into something that is leaner and, I hope to show, meaner. To see this, consider a way of rephrasing the objection developed above. In the case of MapQuest and the jigsaw puzzle, we might say that the external operations I perform provide a useful *scaffolding* or *framework* within which the real cognitive processes can operate. That is, the actions I perform on the world provide a scaffolding or framework within which my cognition can be usefully *embedded*. And this is the thesis of the embedded mind. This scaffolding in which my cognition is embedded considerably reduces the amount of (real) cognitive processes that I must perform in working out how to get from A to B, or in completing the jigsaw puzzle. But this doesn't mean that the framework is literally part of the cognitive processes in which I engage when I am navigating or jigsaw-completing.

The traditional—Cartesian—way of thinking about cognitive processes draws a firm distinction between cognitive processes proper and the framework or scaffolding within which these processes occur. This scaffolding can causally affect, and so facilitate, the cognitive processes proper. But, according to Cartesian cognitive science, we must be careful not to confuse real cognition with extraneous causal accompaniments (Adams and Aizawa 2001). Notice that this point has, in effect, been developed in terms of the language of embedding. Thus, endorsing the thesis of the embedded mind can actually provide a way of attacking non-Cartesian cognitive science—at least in the way I have developed this idea.

The correct response to this is not to deny the distinction between cognition and scaffolding. Such a move might be tempting, but only, I suspect, because it has, as Bertrand Russell once put, all the advantages of theft over honest toil. We need to reiterate the distinction between causation and constitution and cannot fall into the trap of supposing that whatever causally contributes to a cognitive process is part of that cognitive process (cf. Wheeler 2008). To say that the distinction between cognition and scaffolding has been drawn in the wrong place (or has, in general, not been drawn in the right place) is not to deny that there is a distinction between cognition and scaffolding. Without this distinction, non-Cartesian cognitive science becomes true by stipulation.

If non-Cartesian cognitive science is to provide a genuine (i.e., interesting and informative) alternative to its Cartesian counterpart, then it must be based on a stronger claim than the mere environmental embedding of cognitive processes. The required claim, I am going to argue, is that processes occurring in the environment—that is, outside the brain—can, in part, literally *constitute* cognitive processes. That is, things existing or occurring outside of the brain can be partial constituents of cognitive processes. So, the question that should be used to demarcate non-Cartesian cognitive science from the Cartesian alternative is this: can states, processes, and structures that exist outside the brain form part—a genuinely cognitive part—of cognitive processes? Cartesian cognitive science answers *no*: cognitive processes are exclusively realized by brain processes. To the extent that cognitive processes rely on processes occurring outside the brain, these form part of the scaffolding or framework within which these processes are situated, but do not literally form part of cognition.

Non-Cartesian cognitive science, at least as I shall develop this idea, answers *yes*: some cognitive processes—not all by any means, but some—can have, as literal constituents, processes occurring outside the brain. "Outside the brain," of course, does not necessarily mean "outside the body." Here, the new science bifurcates. The *embodied* strand of the science emphasizes the way on which cognitive processes can, in part, be constituted by processes and structures occurring outside the brain but inside the body of the cognitive organism. The *extended* strand, however, argues that cognitive processes can, in part, be made up of processes and structures that occur outside the body in the wider environment of the cognitive organism.

At stake, of course, is the Cartesian vision of mental phenomena as ones located inside the heads of thinking subjects. This vision has under-

written both our commonsense conception of the mind, and the scientific attempts to study the mind that have been erected on the foundation of common sense. The Cartesian vision stands or falls on this question: can processes occurring outside of the brain form genuinely cognitive parts of cognitive processes? If they can, and if cognitive processes are mental processes, then the Cartesian vision of the mental must be abandoned. The aim of this book is to show that the Cartesian vision must in fact be abandoned.