Varieties of Meaning

The 2002 Jean Nicod Lectures

Ruth Garrett Millikan

Varieties of Meaning

The Jean Nicod Lectures
François Recanati, editor

The Elm and the Expert: Mentalese and Its Semantics, Jerry A. Fodor (1994)

Naturalizing the Mind, Fred Dretske (1995)

Strong Feelings: Emotion, Addiction, and Human Behavior, Jon Elster (1999)

Knowledge, Possibility, and Consciousness, John Perry (2001)

Rationality in Action, John Searle (2002)

Varieties of Meaning: The 2002 Jean Nicod Lectures, Ruth Garrett Millikan (2004)

Varieties of Meaning

The 2002 Jean Nicod Lectures

Ruth Garrett Millikan

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

This book was set in Palatino by SNP Best-set Typesetter Ltd., Hong Kong
Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Millikan, Ruth Garrett.
Varieties of meaning: the 2002 Jean Nicod lectures / Ruth Garrett Millikan.
p. cm.—(The Jean Nicod lectures)
"A Bradford book."
Includes bibliographical references and index.
ISBN 0-262-13444-6 (hc: alk. paper)
1. Meaning (Philosophy). I. Title. II. Series.

B105.M4M55 2004
121'.68—dc22

2003063301

10 9 8 7 6 5 4 3 2 1

Contents

Series Foreword

The Jean Nicod Lectures are delivered annually in Paris by a leading philosopher of mind or philosophically oriented cognitive scientist. The 1993 inaugural lectures marked the centenary of the birth of the French philosopher and logician Jean Nicod (1893–1931). The lectures are sponsored by the Centre National de la Recherche Scientifique (CNRS) and are organized in cooperation with the Fondation Maison des Sciences de l'Homme (MSH Foundation). The series hosts the texts of the lectures or the monographs they inspire.

Jacques Bouveresse, President of the Jean Nicod Committee
François Recanati, Secretary of the Jean Nicod Committee and Editor of the Series

Preface

What are the varieties of meaning? And what do they have in common, so as to be treated together under one cover?

People mean to do various things. Body organs are meant to do various things. Tools and other artifacts are meant for various things. Conventional signs such as words and sentences mean various things. People mean various things by using words and sentences, and they don't always mean the same things that the words or sentences mean. Natural signs mean things. Contemporary psychologists and neurologists claim that there are representations in people's brains. Presumably these also mean things. What is in common among the various things that are said to mean things? Nothing, I think. Yet the story of how these various phenomena are related to one another, how they cross and overlap such that the term "meaning" moves freely among them, is deeply interesting. Indeed, none of these kinds of meaning can be thoroughly understood, I believe, without grasping its relation to the others.

In one basic sense, what something means, or is meant for, or is meant to do has to do with its purpose. What a sign means, however, is not usually described as its purpose, but rather as what it represents or signifies. What does meaning in the sense of purposing have to do with meaning in the sense of representing or signifying? According to the *Oxford English Dictionary*, all three of the following senses of the verb *to mean* go back as far as can be traced: *to have in mind*, *to intend*, *to signify*. The original connection seems to be that intention or purpose is what one has in mind as one acts, whereas signification is what one has in mind as one speaks. What one has in mind as one speaks is what it is one's purpose to represent or signify. But there are other places where purpose and signification intersect as well.

The paradigms of purpose are explicit human intentions. I will argue that these intentions are represented purposes. This does not mean

merely that they represent purposes. They possess the purposes that they represent. They are self-representing purposes. The paradigms of things that signify are intentional signs such as sentences. These are distinguished from natural signs by their capacity to misrepresent or be false. (For example, black clouds can be a natural sign of rain, but they don't mean rain unless it actually rains. They cannot be false.) Intentional signs, I will argue, have purposes essentially. The reason they can be false is that they have purposes, and purposes can always fail to be fulfilled. This is true of all conventional signs, of all signs by which animals communicate, and of all inner representations such as perceptions and thoughts.[1]

I begin in Part I by discussing what purposes are, the purposes of people, of their behaviors, of their body parts, of their artifacts, and of the signs they use. Then I begin again at right angles in Part II, describing a variety of natural signs that I call "locally recurrent" natural signs.[2] These are more user-friendly than those that convey "natural information" in the sense of Dretske (1981). I discuss similarities between locally recurrent signs and natural language signs, for local signs are intrinsically "productive" and admit of embedding. Moreover, they can represent individuals, which natural signs have not previously been thought to do. I then introduce several varieties of intentional signs and describe their relations both to natural signs and to purposes. Part II ends with a discussion of how representations themselves sometimes get represented, and hence how the phenomenon of intensionality (with an "s") emerges.

Part III concerns important continuities between the ways local natural signs and public language signs are read and understood. Perception involves the interpretation of local natural signs, and interpreting conventional language signs is surprisingly like perception. The result is quite a different interpretation of the relation between semantics and pragmatics than has been traditional.

Part IV speculates about stages in the development of inner representations, from the most primitive kinds in the simplest organisms up to the perceptions and intentional attitudes of humans. The most primitive representations are what I term "pushmi-pullyus." Their functions

1. Notice that I do not say that the purpose of an intentional sign is to represent. That formulation muddles together a number of issues that need to be carefully separated. See chapter 5 below.

2. In (Millikan 2000), appendix B, I called these signs, or an earlier version of them, "soft natural signs."

are undifferentiated between description and direction. The problem is how and why any more specialized signs should ever have evolved, and ultimately, what in humans is the value of subject-predicate judgment and the peculiar ability to think negative thoughts. Sophisticated biological mechanisms are often built on top of more primitive ones, riding piggyback, as it were; nor is their higher authority always secure. For this reason, speculation about the evolution of increasingly differentiated forms of behavior control in animals is directly relevant to understanding the complexities of human behavior.

The book is written so that those more interested in thought than language can safely skip Part III, and those more interested in language than perception and thought can safely skip part IV.

I am extremely grateful to the Centre National de la Recherche Scientifique (CNRS) for sponsoring the lecture series that initiated my writing this book and to my audiences in Paris who challenged me with excellent constructive questions every step of the way. I profited greatly from those meetings, during which much of the first two parts of the book were initially presented, and I much enjoyed both the good company and the cultural setting in Paris. Gunnar Björnsson and Nicholas Shea offered comments on the first two parts of the book that were both sympathetic and helpfully critical. Carol Fowler, Bruno Galantucci, and I discussed the first three parts, at Bruno's request, in a wonderfully cheerful tiny seminar in the fall of 2002. I learned from them where terminology and ideas that are home ground for philosophers are foreign to psychologists, and I have tried to adjust various passages accordingly. Crawford (Tim) Elder has read all of the chapters in all of the parts. His interest in the project has been unflagging and an indispensable source of support. He is the chairman of my department, and I am deeply grateful to him for this and for many other kind gestures as well. Karl Stocker did the index—the second time he has done an index for me. Thanks, Karl.

I

Purposes and Cross-purposes

1 Purposes and Cross-purposes of Humans

Imagine that the eye doctor is trying to put drops in your eye but you keep blinking. You insist you don't *mean* to blink but that no matter how hard you try, when the eyedropper comes too close, your eye just closes. Perhaps unconsciously you don't want that medicine in your eye? What could your underlying motive be?

The Freudian move is a joke, of course. But there does seem to be a sense in which that medicine is not wanted in your eye. We say that the eye "is *meant* to close automatically" when a foreign object comes too near. The point is to prevent foreign objects from entering it. That is *the purpose of* the eye-blink reflex. The difficulty is that you and your eye, or you and your eye-blink reflex, are at cross-purposes. You are trying to let the drops in but the reflex's purpose is to keep them out.

Maybe you will object that only one of these crossing purposes is a *real* purpose. The other is a "purpose" not literally but only by analogy or metaphorically. The real purpose is the conscious human intention not to blink. Only the intention not to blink is a purpose of the whole person, rather than merely a "subpersonal" purpose. The purpose of the eye-blink reflex is only a "subpersonal" or a "biological" purpose, and these are purposes only metaphorically.

I am going to try to dissuade you of that. I will try to persuade you that no interesting theoretical line can be drawn between these two kinds of purposes. Purposes of the whole person are made up out of intertwined purposes at "lower" or more "biological" levels.

In elementary psychology classes students are sometimes given this as a homework assignment: For the next few days, every time you and your roommate are talking together, smile whenever your roommate blinks. When performed successfully, the experiment brings home to the student the power of operant (instrumental) conditioning, even in humans. For the roommate soon begins to blink more frequently, yet

will be unaware of this, and certainly completely unaware of the *reason*. The purpose of the blinking is to collect smiles. This is known to the student trainer, of course, but not to the blinker. Now ask yourself, where is this purpose resident? The idea that this purpose was, for some reason, "repressed" would surely be ridiculous. But if this purpose isn't the blinker's purpose, then whose purpose is it? Is it only a "biological" purpose, like that of the protective eye-blink reflex? Suppose that the student trainer casually mentions the frequent blinking to the roommate. The roommate will find it difficult to stop. Are these crossing purposes—the purpose of the blinker to stop blinking and the purpose of the blink to bring in smiles—both *real* purposes, or is one of them a purpose only metaphorically?

Some facial gestures are invariant in meaning across cultures. Smiling is one, frowning in anger is another, and apparently raising one's eyebrows is a third (Ekman 1980). Suppose that raising one's eyebrows, like various animal signals such as the cat's arched back or the frog's mating call, is an adaptation, a product of natural selection, selected for serving a particular sort of communicative function. Comprehension of its meaning by observers will have been selected for too. That this is the case with smiling and angry frowning, at least, is not in doubt. Indeed, half of the muscles used in a real, heartfelt smile are involuntary (Damasio 1994). But raising one's eyebrows, like breathing, is also under conscious control. Probably the biological purpose of raising the eyebrows and the eyebrow-raiser's purpose when purposefully raising the eyebrows usually coincide or overlap. Probably they don't, anyway, typically cross. But why is it so hard to *say* exactly what you communicate with a raise of the eyebrows, or with a smile? Do you really *know*, articulately, what you intend to convey when you raise your eyebrows? Does the child know, or the man on the street? If you can't say, offhand, *exactly* what function the raising your eyebrows has (try it!)—as you *can* say offhand exactly why you utter "please pass the salt" when you do, and as you *can* also say why you were trying *not* to blink at the eye doctor's—is it possible that the raised eyebrows' purpose, even when the eyebrows are raised purposefully, really is not a fully *conscious* purpose? But again, knowledge of the raised eyebrows' purpose surely has not been repressed. (It is not just like an explicit conscious purpose except that it's gone underground.) So if you don't fully understand the function of raising your eyebrows, is its purpose entirely real? Or is it purposeful only in a metaphorical way?

Raising one's eyebrows may not be the clearest example here. So consider instead greeting rituals. It may be that all mammals perform greeting rituals; certainly, for example, dogs, cats, mice, sea otters, and at least some whole communities of killer whales do. The phenomenon of greeting is so common across diverse animal species that, despite debate about just what the function of greetings is, no student of animal behavior doubts that greeting rituals have a function, a survival value. We humans, on the other hand, are not born with ready-made natural ways of greeting. Greetings are done different ways in different cultures. There are various verbal formulas and various hand gestures, head gestures, and so forth. There is curtseying and bowing, handshaking, kissing on one or both cheeks, breath sniffing, and so forth.

These rituals do not have biological functions in the sense of particular *genes* having been selected for producing them. Presumably, however, they *have* been selected for culturally. Their continued reproduction, paired with reproduction of the psychological responses of those receiving the greetings, is accounted for by some function they are serving. There is some effect that loops back to encourage continued reproduction of both the particular forms of greeting used in a particular culture and the standard responses to them. There is some cultural purpose that they serve. They are also under conscious control. But can you describe the purpose of greeting, offhand? Are you aware whenever you greet someone what that purpose is, as you are aware of the purpose when you ask for the salt? You know that it's impolite not to give a greeting, and you want to be polite. But *why* is it polite in every society to give some sort greeting? What's the purpose? Clearly there is some purpose that our greetings have over and above just what you and I consciously intend them to have. Again we can ask, is the sense in which greetings themselves have a "purpose" a *real* purpose, or is it a purpose only metaphorically speaking?

It is clear what the biological purpose of a taste for sweet things is. A sweet taste is a natural attractor or reinforcer designed to increase behavior that leads to the intake of high-calorie foods. Although the genes are responsible for the fact that sweets (and also smiles) will reinforce human behaviors, the purpose of the conditioned behaviors themselves is not the same, of course, as that of the genes. Conditioned behaviors have been selected for on their own level, selected for bringing in their own designated rewards—in this case, for bringing sweet tastes into the mouth. On this level, for example, the purpose of

behaviors conditioned by sweets is served by foods containing saccharin. Moreover, many foods high in calories don't serve these purposes.[1] Further, granted that people don't have inborn concepts, in humans a taste for sweets often brings in sweets not as a result of automatic operant conditioning, but by a person's having learned from experience that they like sweets, and having consequently formed a conscious psychological goal—another level of purpose—of obtaining and eating sweet foods. Still, a taste for sweets can also work prior to, and perhaps also after, "sweet" is conceptualized, by reinforcing behaviors directly, as smiles can reinforce eye-blinks. Psychological purposes induced by acquaintance with sweet things are entirely different from their underlying biological purpose. A desire for sweets is not a desire for calories any more than reinforcement by sweets is reinforcement by calories. It is merely a desire for or reinforcement by sweet tastes, which happen to have been strongly correlated with calories in the historic environment of humans. Similarly, although pain avoidance has tissue-damage avoidance as a biological purpose, the desire to relieve pain is not a desire to relieve tissue damage. In the contemporary context, of course, obtaining more calories is not always a good means to the further biological purposes of increased energy and health. Thus both the conscious and the unconscious (the conditioned) psychological purposes of obtaining sweets may cross purposes with their own deeper biological purposes.

Of course, most conscious psychological purposes of humans are not immediately rooted in original attractions and aversions such as the attraction to sweets or to smiles or, say, the aversion to pain. Most conscious psychological purposes are highly derived, distantly rooted in a variety of more original attractions and aversions, mediated by a vast number of beliefs, true or false, about causes and effects and about other aspects of the environment. Thus it may happen that a modern person realizes what the usual source of sweetness is (sugars), realizes what the effects may be of indulging a taste for sweets (getting fat), and thus acquires the avoidance of sweets as a rational conscious purpose. It can happen that the last thing you rationally want is more calories, and yet you still crave sweets. This sort of inner crossing of purposes is all too familiar, of course. It has been discussed at length by both ancient and modern philosophers under the name of *weakness of will* or

1. On the remarkably strict analogy between genetic selection and the selection of behaviors by operant conditioning, see Hull, Langman, and Glenn (2001).

akrasia. Purposes produced circuitously by pale thought do not necessarily win over purposes derived through more biologically ancient and direct routes. Indeed, perhaps there is a good reason for this. "Doing what comes naturally" is not necessarily good for you, but then neither is doing what reason and/or culture currently happen to dictate. (Consider, in biological context, the highly derived goal of celibacy. Consider the history of medicine: Bishop Berkeley's recommendation of tar water as a cure for all ills; forcing the boys at Eaton to smoke tobacco in order to prevent plague.)

Could sweets operating as an unconscious reinforcer directly conflict with an explicit desire to avoid sweets? Certainly one can absently eat "too many cookies" while engrossed in conversation, whereas one would be unlikely to eat too many of something one didn't like. Returning again to our original question, if this sort of conflict can occur between explicit desire and conditioned response, is the conflict here between two genuine purposes, or is the more subversive purpose, the procurement of cookies and hence sweet tastes, a purpose only analogically? Does the answer to this question depend, perhaps, on whether the Skinnerian or the Freudian is right about the status of motives such as the nonconscious cookie-eating motive—or, taking a different example, about why Hans is attracted to a woman who is so much like his mother? If Skinner is right, these motives are biological and hence only analogical purposes; if Freud is right they are psychological hence real purposes. Is that how it goes?

Do you still think there is a clean divide somewhere between real and merely metaphorical or biological purposes? Then which of these are governed by real purposes and which are not? Which are purposes of the whole person and which only "subpersonal" purposes?

- Pushing the snooze button on the alarm without waking up.
- Fearfully retreating from a harmless snake, or from the fence at the edge of a precipice, despite knowing there is no real danger.
- Gently applying the brake to negotiate a curve while completely absorbed in a conversation.
- Swinging your left foot forward on your way to answering the telephone.
- Swinging your left foot forward while walking when your attention has just been drawn to this motion.

Suppose that you are having trouble sleeping and suddenly become super aware of your breathing. Does the purpose of your breathing suddenly shift from being a biological purpose to being a psychological purpose? Does it suddenly shift from being a "subpersonal" purpose to being a purpose of your whole self? Suppose that the awareness of your breathing contributes to keeping you awake, and you make every effort to breathe naturally. Now can you say, about each breath, whether it is taken purposefully in the literal sense or purposefully only metaphorically? It is not possible, I believe, to draw a principled line between real purposes and merely metaphorical purposes, or between people's purposes and merely biological or "natural" purposes, or between purposes of the whole person and subpersonal purposes, purposes of a person's parts.

The purposes that at first seem farthest removed from mere biological purposes are explicit human goals, desires, and intentions. These are different in part because they are mentally represented purposes. They represent the conditions of their own fulfillment. But, of course, mental representations couldn't represent their own purposes unless they *had* purposes to represent, and these purposes are derived from various levels of selection. Explicit desires and intentions are mental representations whose purposes are to help to produce what they represent. They were selected for helping to bring about the conditions they represent.[2]

---

2. Claims have been made that the mechanisms currently accounting for human thought may not have resulted from natural selection. Perhaps the mechanisms appeared on the scene quite accidentally. Most recently, Fodor (2001), following Chomsky, who seems to be following such figures as Lewontin (1978), Gould and Lewontin (1979), Gould and Vrba (1982), and Gould (1991), argues against the idea that human cognition is an adaptation as follows:

> What matters to the plausibility that the architecture of our minds is an adaptation is how much genotypic alteration would have been required for it to evolve from the mind of the nearest ancestral ape whose cognitive architecture was different from ours. . . . If changing the physiology a little makes a very large change in fitness, the difference between a selection theory and a saltation theory disappears. . . . Nothing we know about how cognitive structure supervenes on neural structure impugns the possibility that quite small variations in the latter may produce very large reorganizations in the former. Well, likewise, nothing impugns the idea that quite small changes in a creature's cognitive *structure* may produce very large changes in its cognitive capacity. (pp. 89–90)

And how does Fodor suppose that that very small genetic change in one of our lucky ancestors just happened to get handed down to all the rest of us? The job of natural selection is never, strictly speaking, design. What natural selection does is only to weed out the creatures that are less fit. If there is a good thing going, like, for example, the turtles,

They were not selected for one by one, of course, certainly not on the level of genetic evolution. Only the cognitive and conative mechanisms responsible for forming desires and intentions were designed or chosen by natural selection. They were selected for their capacity, on the basis of experience, to form representations of goals, of possible future states of affairs, which, when brought about, sometimes furthered our biological interests. The job that these representations did, in turn, was to serve as blueprints, guiding the causal processes by which behaviors that brought about the represented states of affairs were constructed. Compare the design of a camera or of a calculator. The camera is not designed, specifically, to take any particular picture that it takes, nor is the calculator designed, specifically, to make one particular calculation rather than another. Still, when the camera is working right, it is designed to turn out each picture that it turns out, given its input. And the calculator that is working right gives each individual result in accordance with design, again, depending on its input. An explicit intention does what intentions were designed to do when it initiates its own fulfillment. Exactly similarly, a desire does what desires were selected for doing when it is eventually transformed into a fulfilled intention. Of course the majority of our desires may never be fulfilled. Desires very often conflict with one another as well as with purposes on more primitive levels. And often we can come up with no means of fulfilling our desires. Similarly, most pounces of the cat may miss the mouse. Lots of things fail to serve their biological or natural functions more often than they succeed. The point is that the capacity to develop and to act on desires would seem to have been selected for only because desires are sometimes fulfilled and, of course, sometimes do represent means to fulfillment of our biological interests.

The purpose that is an intention is clearly a purpose of the whole person who has that intention. When you have an intention to do something it is *your* purpose to do it. So, you may suppose, it is not merely

---

who have not changed in nearly three hundred million years, natural selection has remained very busy all that time maintaining the status quo. Its job is to throw away the trash that would otherwise accumulate, not to make changes. The notion that the current human brain was not selected for is patently absurd. That it was selected for is what happened to the various earlier now extinct species of the genus *Homo*. (It is also what happened to the Neanderthals.) Similarly, the common notion that making reference to natural selection is making reference to ancient history is badly mistaken. Possibly our brains arrived in Paleolithic times or before. But they have managed to proliferate considerably in very recent years, to the detriment, of course, of many other species. They have been selected for.

a purpose of some part or aspect of your self. Yet looked at another way, to talk about your explicit purposes, your intentions, is merely another way of talking about the natural purposes of your represented intentions themselves. They have natural purposes just as do many other parts of your self such as your heart and your eyes and the nerve connections responsible for your reflexes, your conditioned responses and your taste for sweets. And turning the coin over again, the purposes of these other parts and aspects of you are as much purposes of yours as are the purposes of your explicit intentions. After all, what you are is the sum and interaction of all the various parts and aspects of your self, and vice versa. So of course all these purposes are your purposes, and vice versa. Looked at critically, the distinction between whole-person purposes and subpersonal purposes collapses.

Explicit goals, desires, and intentions may seem very different from other purposes because we are "aware" of them. We are aware of what we are doing when we act on our desires and intentions. We can introspect, and say what our desires and intentions are, whereas introspection will not reveal the purpose of our eye-blink reflex or of our hearts' beating, nor will it reveal to the roommate the purpose of his blinks.[3] But lack of awareness of an action does not affect its purposiveness. This has long been recognized by psychiatrists and hypnotists. One can certainly have purposes that are one's own purposes but that are hidden from one's introspective view. Taking a very different kind of example, consider what it is that happens if you suddenly become aware of each breath you are taking, or of the steps you are taking on the way to answering the telephone. Surely these do not suddenly become purposive activities or purposes of your whole person whereas before they were not. If you suddenly awake from deep thought while you are driving and notice where you are and what you are doing for the first time in ten minutes or so, surely your driving does not suddenly turn into a purposive activity for the first time or into an activity of your whole person when it was not before. According to the social psychologist John Bargh, "our psychological reactions from moment to moment" are "99 and 44/100% automatic" (1997, pp. 243–244). Surely this does not mean that only 56/100 percent of our reactions are purposive! That adult humans typically have the capacity to discern and to think about their own intentions is certainly interesting and certainly requires explanation, but very young children don't have this capacity.

3. Thanks to Sarah Buss for pointing out to me the need for this paragraph.

They do not even acquire concepts of beliefs and intentions until they are three or four. Surely they perform many intentional acts long before that.

Though no more our purposes than are the purposes of any of our other aspects, consciously represented purposes result from a higher level of selection than do the purposes of our genes and the purposes of conditioned behaviors. As Popper, Dennett, and many others have noted, on this level, experimental thought attempts to reach consciously projected goals by trial and error. Based on our past experience, we think through the consequences of various courses of action, one after another, until we find a plan that seems to work. Then we form an explicit intention, which, when the behavioral systems function properly in the right supporting environment, initiates its own fulfillment. The most basic capacities that allow for this, by supporting our capacities to develop concepts, to collect information and so forth are, of course, built in by our genes. But the resulting behaviors have not been selected for by the genes. The behaviors selected in this way have as their purpose to achieve the projected goals for which they were selected. But suppose that we now ask: What determines which conscious goals are projected? Whose goals are these? What mechanism has selected these goals?

My plan to buy donuts and milk on the way home from work may at first seem, clearly, to express a human purpose, whereas the reflex eye-blink that prevents a sand grain from entering my eye seems to express merely a biological purpose, a purpose derived only from a history of natural selection. But explicit goals and intentions emerge out of a sea of more primitive behavior controls, and the details of the execution of explicit goals are again submerged. The explicit intention to buy donuts emerges from a primitive attraction to sweet tastes, designed to motivate my indulgence in high-calorie foods. The explicit intention to buy milk may have emerged from a history of reinforcement by smiles when I drank my milk as a child. And as I retrieve my donut from the package, convey it to my mouth and chew, each minute adjustment of the fingers, hand, tongue, and jaw has a definite purpose, though I am not conscious of most of these motions. Indeed, I am quite incapable of becoming explicitly aware of most of them, let alone of their individual purposes.

Clearly we don't aim just for what our genes aim for. Our original conscious goals are not merely to survive and to have lots of children. An easy assumption would be that what we originally aim for, or away

from, with our calculated behaviors is just whatever reinforces us, positively or negatively, and that we then aim for whatever we have calculated to be means to those original ends. This is the answer suggested by much classical twentieth-century motivational psychology. But I think this is almost certainly mistaken. There is a large gap between possessing a mechanism that reinforces behaviors that lead to certain reinforcing internal states, sensations, perceptions, and so forth, and understanding what those reinforcers are. Surely we are not born with the ability to think thoughts of sweet tastes or of food, of water or of other classical "primary reinforcers." Often we may know *when* we are attracted or repulsed, but knowing exactly *why* we are attracted or repulsed is another matter entirely. Knowing exactly what it is that we want, or why, is not automatic. Likely we only find this out by experience, indeed, by something analogous to hypothesis-formation and testing. "There is something about him that I don't like," we say, or "For some reason, huge supermarkets make me very anxious," or "I don't know what it is about her that attracts me so"—these are common kinds of reactions. But if it is true that we don't always know what will satisfy us, what will repel us and so forth, then here is yet another opportunity for mismatches among our purposes. Moreover, knowing what it is that attracts us or repels us need not lead to a reasoned desire for or against that thing. Conflicting interests often appear on the same level as well as on different levels.

In chapter 17, I will argue that there may also be another level of selection behind human purposes, which falls between the selection that is conditioning and the selection that is trial and error in thought. If you watch a squirrel trying to get to a well-armored bird feeder, for example, you will witness another level at work. The squirrel is not on the conceptual level I have just been discussing. It does not conceptualize its purpose, make inferences; of that I am pretty sure. But it studies the perceptual situation at length, from first one angle, then another, walking from one side to the other, climbing up a little here and there, experimenting until it "sees" a way it might try to get up. It experiments until it "sees" a Gibsonian affordance. We often do the same kind of thing. While on a rather steep hike, while still completely engrossed in conversation, you may hesitate for a second or two as you examine the best way to get a leg up to the next level, or mentally experiment with which stones in which order to step on to get safely across a stream. This is not conceptual thought, but a sort of trial and error on

the perceptual level, yielding purposive behaviors on yet another level of selection.

In sum, if we look at the whole human person in the light of our history of evolution by natural selection, minding the continuities between humans and other animals, it appears that all levels of purpose have their origin in adaptation by some form of selection. In this sense all purposes are "natural purposes." Even though there are, of course, many important differences among these kinds of purposes, there is a univocal sense of "purposes" in which they are all exactly the same.

One more kind of purpose that is derived from natural selection though not from a separate level of selection is that of artifacts. The purposes of many artifacts are derived directly from the purposes their producers had in producing them. Spiderwebs and beaver dams, for example, result from the operation of inner mechanisms in the spider or beaver that produce behavioral dispositions that result in the construction of the webs or dams. These producing mechanisms were designed pretty directly by natural selection. The purposes of these artifacts are derived from the purposes of the genes that were selected for producing them. The webs and dams themselves are what Dawkins (1983) termed "extended phenotypes," and the purposes of these webs and dams are likewise extended purposes of the inner mechanisms, and prior to that, purposes of the genes that built the mechanisms. Artifacts made by people are "meant" to do things by the people who make them. If people's purposes are a form of natural purposes, derived from the selection of genes and the selection processes producing conditioned behaviors and rational selection of means to ends, then the purposes of these artifacts are also derived from natural selection. So the purposes of genes, of unlearned behaviors (smiling), of learned behaviors, of conscious intentional actions, of at least some cultural products (greeting rituals), and of artifacts are all purposes in exactly the same sense of "purpose." In all cases the thing's purpose is, in one way or another, what it was selected for doing. Moreover, the purposes we attribute to whole persons, rather than just to various of their aspects or parts, are composed of no more than the purposes of these parts and aspects, and of the ways these have been designed to work together.

Notice that this result has not been gained by the method of conceptual analysis. Putting ordinary language and ordinary ways of thinking aside, I have tried to find what is actually in common,

according to modern empirical theory, among various kinds of purposes. I have tried to describe the common underlying pattern beneath the surface features that we recognize as marks of purposiveness across a variety of domains. If you feel that I am using the term "purpose" not in its usual but in some technical sense, I have no interest in arguing the point. My interest lies only in revealing what I take to be an important sort of commonality among the various phenomena that we think of as purposive. And it lies in trying to dissolve the common view that there is some sort of great cleavage between *real* purposes, purposes of the whole person—*my* purposes—and the purposes of less sophisticated things, including parts and aspects of the whole me.

2 Purposes and Cross-purposes of Memes

One of the many things that have evolved by natural selection is evolvability itself. One example of this is the evolution of sexual reproduction, which mixes genes in such a way as to introduce wide variation for selection in organisms while still ensuring that most remain viable. Another example is the evolution of homeo box genes. These are genes lying close together on the same chromosome that control the expression of other genes whose phenotypic effects lie within strictly limited bodily areas. Thus one part of the animal can be changed in various ways, effectively experimented with, without at the same time affecting random unrelated parts of the organism. The evolution of completely new levels of natural selection that ride piggyback on lower levels is another way in which evolution evolves. The evolution of behavioral systems controlled partly by mechanisms that learn by operant or instrumental conditioning, discussed in the previous chapter, is an example of the evolution of a new level of selection, as are the development of trial and error in perception of paths to a goal and the development of Popperian trial and error by which representations are experimented with in thought. In chapter 1, I argued that the purposes that emerge from these various levels of selection are not always compatible with one another but are sometimes at cross-purposes.

Another aid to evolution results from systematic segmentation and then recombination of elements that have already served usefully in prior combinations. There is evidence that segmentation and recombination are at work in trial and error learning of motor skills. Segments of behavior lineages of various durations are retained or substituted for during the learning process, much as segments of chromosomes of various lengths are recombined during sexual reproduction and selection (Hull et al. 2001). Recombination is also exemplified by the way

the immune system works (Cziko 1995; Hull et al. 2001). One of the requirements for the emergence of adaptive products from a process of competitive selection for reproduction is that the reproduction should be extremely accurate for the most part, reproducing always the same aspects of each new model. Dawkins (1976) calls this requirement "fidelity." Of course some variation is needed or no evolution would occur, but it must be carefully controlled variation. One way to make reproduction faithful is to digitize what is to be copied, and then design the copying machinery to recognize only the presence or absence of each digit, ignoring minor variations in the original. This is how DNA is copied, for example. The digits are the "letters" ACGT, which in combinations of three make up the "codons," each of which stands for one of the twenty amino acids. The copying within the immune system is similar.

In this chapter I will discuss another level of replication and selection from which natural purposes emerge, again, sometimes crossing with lower levels. This is a level on which reproduced cultural items of the sort Dawkins (1976) called "memes" are selected.[1] It is also a level on which segmentation and recombination of definite digital units often occurs, the most striking example being the phonological, morphological, and syntactic structures of natural languages, which account for the accurate reproduction and rapid evolution of linguistic elements and for the possibility of compositional semantics.[2]

Richard Dawkins invented the term "memes" to stand for items that are reproduced by imitation rather than reproduced genetically (Dawkins 1976, chap. 11). According to Dawkins, memes are such things as tunes, fashions in dress, other kinds of fads, handed-down ideas, values, forms of expression, words, and so forth. They get copied, one person copying from another. And given that people have limited memories, and limited energies and time to devote to copying things, these things can be thought of, Dawkins claims, as competing with one another for reproduction. Following the logic of chapter 1 above, what a meme does that accounts for its continuing to be reproduced while other memes die out corresponds to the *purpose* of the meme. It is what the meme has been selected for. An example already

1. These various levels of evolvability should not be confused with the levels constituted by gene selection, individual selection, and group selection, sometimes referred to as "levels of selection."

2. Other examples of the digitization of memes are discussed under the heading "normalization" in Dennett (forthcoming).

mentioned is human greeting rituals. There must be some reason why all human cultures have greeting rituals, though these are passed on through imitation rather than through the genes.

Dawkins originally compared memes to viruses, which use the cellular mechanisms designed to reproduce genes to reproduce themselves. The purposes of viruses need not, of course, coincide at all with the purposes of the mechanisms that reproduce them. Similarly, what makes a meme such as an annoying tune or a superstitious belief stick in one's mind and be reproduced may be some quirk or side effect of people's brains that has no biological or psychological function. Dawkins suspected that very many of the memes people harbor cross people's own purposes or the purposes of their genes. Think, for example, of the meme that is high-heeled shoes and the meme that is pious celibacy.

Susan Blackmore (1999) goes so far as to assert that memes are an independent force making "things happen that serve to spread memes whether or not they spread genes" in a way that mere "Skinnerian learning and Popperian problem-solving" do not: "I suggest that the human brain is an example of memes forcing genes to build ever better and better meme-spreading devices. The brain is forced to grow bigger far faster and at much greater cost than would be predicted on the grounds of biological advantage alone" (p. 119). But Blackmore does not describe any mechanism by which memes could force genes in this way, merely boldly asserting that memes (somehow) have this power because they are a "second replicator" that is "set loose on the world." On the contrary, given that the capacity to replicate memes (cultural artifacts), found only feebly when at all in other species, is clearly genetically determined, memes are dependent for their proliferation on the cooperation of the genes in a way that is deeply asymmetric. If the genes that replicate memes are not helped but hindered by these memes, they will disappear from the gene pool. Similarly, if genes for increasing meme production do not consequently increase their own production, they will not increase in the gene pool. (Our cells don't evolve so as better to accommodate viruses just because viruses use our cells' copy machines for their own purposes.)

On the other hand, surely Dawkins is right that memes can cross purposes with the genes that enabled them, exactly as conditioned behaviors and rationally selected purposes can. They can also cross over people's psychological purposes, conflicting with their tastes, aversions, or preferences. Thus nervous English speakers say ". . . uh

. . . uh . . .” at intervals, nervous German speakers say “. . . also . . . also . . . ,” nervous Russian speakers produce a series of just audible grunts under their breath, and nervous Hungarian speakers (I am told) say “. . . ö . . . ö . . . ö . . .” Exactly why these memes proliferate is a good question, but that many speakers want very much to be rid of this habit is quite certain. It could be that some memes have no purpose at all beyond that of triggering our replicating machinery. That our replicating machinery is triggered by them could be merely a side effect of design for better uses of this machinery. But this leaves open what the purpose of our original capacity for reproducing memes is. How is the capacity to replicate cultural artifacts useful to us?

Susan Blackmore proposes that the original purpose of this capacity was to allow imitation of useful behaviors and the passing on of small bits of technology directly from one person to another. Bits of practical knowledge laboriously learned by one individual through trial and error are copied by others, with an enormous community saving. Hence the advent of culture. It does seem obvious enough that a genetically designed ability to pass on technology in this way might proliferate itself. Notice, however, that if passing on technology were all that memetic replication was used for, no new kinds of purposes would emerge from it. Suppose that one man makes an arrow for shooting deer, others copy his method and design, and still others copy the copies, either using or trading their products. If the design proliferates because the arrow shoots well, the purpose of the original, derived from the original maker’s intentions, and the purpose of the copies, reproduced because their models shoot well, are the same. Artifacts that are copied from one another in this manner have a second *source* of purpose or function from that of the original from which they are copied, namely, the function that accounts for their reproduction. But except in odd cases, this function will be the same as the function of the original. Of course people do sometimes find uses for artifacts for which they were not originally designed, and it is possible that some artifacts might even be proliferated for such purposes. But still, the purpose will be a practical human purpose, coinciding with some ordinary psychological purpose of humans who use it. It will not be some new kind of private purpose invented by the memes. Technological memes get themselves replicated by serving people’s prior interests. They are not like viruses. Some technological innovations may be more easily understood and remembered than others, and for this reason they may proliferate more readily. Part of what they have been selected for is their ability to be

reproduced accurately through the medium of human minds. But this does not subvert their essentially human purposes.

Imitation plays a very obvious role in the arts as broadly conceived. Visual art objects and styles, musical compositions and musical styles, tunes, dances, stories, poems, jokes, games, and amusements of every kind are both purposefully reproduced and often unconsciously reproduced. Here again, however, no new element of purpose is added by the memes. It is possible, of course, that our interest in the arts is an accidental by-product of cognitive features originally selected during our evolutionary history for quite other reasons. But if so, still the memes connected with the arts have not invented these interests, producing the psychological aims and purposes that match them. The memes have merely fed these interests a much richer diet than if each person had to invent all of his own amusements, or invent all of the entertainments he uses to invoke the gratitude and appreciation of others. The functions of artistic memes that survive will be to serve prior human interests.

Similarly, beliefs, and ideas and concepts can be considered memes, handed down by mechanisms that reproduce them fairly accurately. Many of these beliefs and ideas may be false, of course, and may circulate widely for reasons having nothing to do with their practical value or their truth-value. But again, the basic mechanisms involved in their reproduction have either been genetically selected for or selected for by learning and reasoning. The probability is that these mechanisms have served purposes of the genes or of the psyche more often than conflicting with these purposes. Side effects and mishaps resulting from use of these mechanisms will surely occur, but there is no reason to suppose that they systematically produce memes with purposes of a different kind from those either of the genes or of the psyche.

There is a domain of memes that do possess a whole new kind of purpose, however, not found on lower levels. These purposes are cooperative purposes. The mechanisms that produce these memes have apparently been designed to facilitate social coordination. To serve a coordinating function is to facilitate the purposes of two or more individuals at once, bringing the separate behaviors of these individuals into a designed coordination that benefits both but that has not been designed by either individually. I will discuss two examples of coordinating memetic functions, one of which is quite speculative, namely, the purpose of certain kinds of social conventions, and the second of which I take to be quite solid, namely, the memetic functions of various

conventional elements of human languages. The memetic functions of conventional language elements are one of the kinds of "meanings" that they have, so one of the connections between meaning as purposing and meaning as a property of signs will be revealed by this discussion.

A puzzling phenomenon involving memes is the uniformity of human behaviors within a given culture that do not seem, in themselves, to be any more practically, artistically, or epistemically rewarding than any of numerous alternatives would be. Styles of dress, kinds of food eaten, behaviors at meals, in the market, in places of entertainment, when socializing, courting behaviors, ways of greeting, ways of celebrating or grieving, ways of accepting or turning down offers and so forth, and also the times and places at which these various activities occur seem to be fairly arbitrary yet highly uniform within each culture. Surely this uniformity is not accounted for by lack of imaginativeness. Moreover, in all cultures people seem to find comfort in conformity, a strong fear of being considered different is common, indeed, it is typical for conformity to be strongly sanctioned. Consider, for example, the standard way of reproving children: "Sally, we don't run outside in our pajamas!" "No, Johnny, we don't eat peas with our fingers!" What *we* do or what *one* does (Heidegger's *das Mann*) is conceived both as just what people do do, and at the same time as what people must do or ought to do. I suspect that a disposition to conformity may actually be built into the human psyche. If so, the purpose may concern social coordination.[3] Let me explain.

Recent speculations about the history of our minds have emphasized the special demands of living in a complex social community in which many conspecifics must be recognized, learned about, and dealt with individually. It is proposed that the ability to live in such a community requires predicting other individuals' behaviors and that this requires a deep understanding of other individuals' minds. Certainly one characteristic that helps to distinguish us sharply from other species is our inventiveness as individuals, and this trait might naturally lead us to be far less predictable in our behaviors than are the members of

3. One hypothesis is that conformity helps to maintain group identity. Outsiders are easily recognized by their foreign behavior patterns. But the groups to which humans belonged during the critical period of their evolution were small enough that every individual in these groups would have known every other individually by face, by name, and by reputation. Surely no further means of telling insiders from outlanders would have been needed.

other species. A second trait that distinguishes us is our huge dependence on others in nearly all aspects of our lives. We use and rely on others, and they on us, for a multitude of everyday purposes. Just as we could not learn to use natural objects to our advantage if nature were completely unpredictable, we could not rely on or even walk safely among other humans if we could not predict many aspects of their behaviors. It is claimed, then, that the social animal needs to develop a theory of mind, a grasp of intentional psychology, so as to predict the behavior of its fellows. And this need quickly escalates. As others begin to understand my mind better, hence to predict my behaviors better, and to adjust their own behaviors accordingly, I need to develop a more and more sophisticated kind of Machiavellian intelligence so as to continue to compete in this context. Trying to keep one thought ahead of the people around me when they are trying to keep one thought ahead of me puts quite a strain on the intellect. This is how, it is proposed, we humans were driven to become so smart.

This way of thinking about the demands of social living seems to be premised on the idea that social cooperation typically requires altruism on the part of the cooperating parties, who will try to duck their cooperative responsibilities if they possibly can. The ability to anticipate, detect, and punish "cheaters" or "free riders," for example, is taken to be central to the development of our kind of intelligence and to the maintenance of a cooperative society. The problem of maintaining an honest communication system in the face of the advantages that accrue to the liar is claimed to be another central issue.

I do not think this is how it works at all. First, for the most part, thinking of social interaction on the model of a competitive game is quite wrong. Most aspects of social living involve cooperation in ways that benefit everyone. Typical patterns of cooperation do not require altruism on anyone's part, and hence do not need to be maintained by sophisticated methods of cheater detection, lie detection, keeping one thought ahead of the next fellow, and so forth. If you and I were trying to move a couch together, for example, and I failed to pick up my end, or we failed to walk in the same direction, it is not just your purpose that would be frustrated, but mine. Similarly, consider what will happen to the driver who fails to conform to the cooperative right-hand rule of the road in America. Consider recording the next hundred fact-stating sentences you utter and then asking yourself for each of them how you would have benefited by lying in this instance rather than telling the truth! Trying to understand how cooperative living developed is not like trying to

understand how altruism developed. People are not, typically, playing against each other. Social cooperation very seldom resembles a game of prisoner's dilemma. That doesn't mean, of course, that it *never* does. It doesn't mean that there never are occasions on which one needs to be aware of the possibility of someone's cheating. But for the most part, social cooperation benefits both or all parties. There is nothing mysterious about its evolution in this respect.

Second, the idea that we usually predict one another's individual behaviors by speculating about each other's personal motives and beliefs seems to me quite wrong. Partly, we expect people to exhibit behavioral patterns similar to those they have shown in the past. Some people usually come to work on foot and on time, others drive or take the metro and often arrive late. Some people always eat lunch at noon, others at other times or irregularly. Some people will talk on and on if you start conversing with them, others are very reticent. Some always stick to their word, others change their minds frequently. Some always eat eggs for breakfast, others always eat yogurt. We take these patterns into account, betting on their continuation when it is useful or necessary to do so. When we use belief-desire psychology, it is almost always for explanation after the fact, not for prediction. We may explain why John always has yogurt for breakfast by saying he must like it, but if he actually eats yogurt only for his health, it won't matter to our predictions.

Less obvious, but perhaps more important, we are often able to predict one another's behavior owing to patterns of social conformity or social convention. We become acutely aware of this when we find ourselves in another culture where things are done differently. Then we are inconvenienced or embarrassed to discover that people come to market or close markets at a different time, sleep and eat meals at different times, use different eating utensils, prefer to sit in different postures on different sorts of mats, cushions, chairs, or stools, accept or turn down invitations in different ways, give or bring different kinds of gifts and for different kinds of occasions, count on assistance of kinds we don't expect from people connected to them in ways we didn't expect, recreate at different times, in different ways, and in different kinds of places, are especially respectful of persons in different kinds of offices, and so forth. They may also drive on a different side of the road!

Focusing on the example of driving, clearly it is sometimes essential to be able to count on uniform behavior in other people. Other kinds

of conformity may also support everyone's well-being, proving efficient and convenient for everyone or for most people much of the time. In many domains, a built-in disposition to do as others do would have a strong tendency to benefit all. Moreover, the more models there are to copy, the more definite it becomes just what it is that should be copied. Quite good fidelity to some pattern or another would slowly but naturally emerge even from initial chaos by this simple principle of self-organization. On the other hand, there will usually be times when conforming is not particularly convenient for an individual. The disposition to conform will sometimes cross over more primitive purposes. Since it is usually best for others that one conform, however, pressures and sanctions will quite naturally be applied. Indeed, a general disposition to discourage nonconforming behaviors in others probably benefits all in the end.

Behavioral dispositions that are genetically selected for are triggered not by the intentional contents of sophisticated cognitive systems, but in much less discriminating ways. Moreover, whether or not a certain kind of practice serves a useful coordination function typically depends on what other social practices are being followed. It is pretty hard to imagine how you could construct a primitive perceptual or cognitive trigger that would discriminate between behaviors that are socially useful to copy and to sanction and those that are not. It is easy to understand why not only socially useful behaviors, but many behaviors originally proliferated for accidental reasons of taste, attraction, salience, and so forth can easily become part of the standard social repertoire.

Thus the disposition to social conformity and the conventional behaviors produced by it may have a purpose, even though many examples of social conformity fail to serve this purpose. The purpose of these behaviors is a general one, derived from the purpose of the mechanisms that create social conformity. A different question concerns whether these behaviors have memetic purposes, derived from selection of individual behavioral memes from among competitors. A guess might be that memes not serving a coordinating function are less likely to be highly sanctioned and less likely to be faithfully copied since no one is actually damaged or inconvenienced when they are ignored or when they drift into new forms. It may be hard not to chafe at the inconvenience when one's daughter turns vegetarian and you are asked to cook two kinds of meals every evening, but not hard to put up with her new tastes in clothes if she buys her own. And indeed, fads and fashions in dress do tend to change more rapidly than, say, family diets.

There may be mild to strong selection pressures stabilizing some but not all individual social conformities. And of course there are statutes and laws stabilizing some others. Then the behaviors exemplifying these conformities do have memetic purposes of their own, but again, not purposes that cross the purposes of most people most of the time.

Perhaps the clearest analogues to genes among memes are the reproduced elements out of which language is built, such as phonemes, words, syntactic structures, elements of prosody, and so forth. These are memes that definitely do have coordinating memetic functions. They are combined and recombined to produce the functions of phrases and full sentences, that is, roughly, the literal meanings of these. There is evidence, moreover, that we have special mechanisms genetically designed to make possible the rapid evolution and discriminating selection of language forms.

The fidelity with which selected items are reproduced is a crucial factor enabling natural selection to produce functional products. Two mechanisms seem to have been built in to ensure fidelity in copying language forms. First is a capacity to grasp, during the first few months of life, the phonological structure of one's language. Phonological structure determines what will count as correct reproduction of an element such as a word or a sentence, enabling the learner to discriminate those aspects of speech signals that matter to meaning from those that can vary freely. Second, Chomskyan linguists posit that universal grammar serves as a filter determining which aspects of the structure of the language it hears a child will reproduce.[4]

---

4. There are rumors that natural selection was not involved in the process that produced Chomskyan universal grammar. Again I quote Fodor, since he has been more explicit with his reasons for believing this than has Chomsky:

> [T]he facts that make a speaker/hearer's innate beliefs about the universals of language true (or false) *aren't* facts about the world; they're facts about the minds of the creature's conspecifics. Roughly speaking, all that is needed to ensure that my innate beliefs about linguistic structure will allow me to learn the language that you speak is that you and I *are* conspecifics; and (hence) that your linguistic behavior is shaped by the same "innate linguistic theory" as my beliefs about your linguistic behavior. And, presumably, what guarantees all these correspondences is that, qua conspecifics, we have the genotypic determinants of our innate beliefs in common. (2001, p. 95)

Similarly, I suppose, the fact that we are conspecifics should also guarantee that we can guess each other's hair color without looking because our hair colors are all the same? Clearly, that is a silly argument. The crucial question should be why we all are born with the *same*, supposedly arbitrary, universal grammar inside. How did it happen that the first *Homo* who accidentally acquired human universal grammar by a genetic fluke (according to Fodor) handed it down to *all* the rest of us? Why did all the *Homos* lacking this arbitrary trait get selected out?

In the case of human language, prior agreement on the kind of materials that are to be used in communication and the aspects of these materials that are to be significant produces a genuinely new kind of faithful replicator, ready for selection. Language forms are then subject to a new and characteristic kind of selection pressure, guiding the evolution of a cleanly different level of natural purpose or function. Unlike the case of most technical skills passed down by imitation, but more like the case of other conventional social forms, those effects that encourage continued replication of a language form are not determined by the purposes only of the agent producing them. The functions of language devices are fulfilled through cooperation between speakers and hearers, and hence are determined by the interests of both. Language devices will produce effects that interest speakers often enough to encourage continued replication only if hearers replicate hoped-for cooperative responses often enough. And hearers will continue to replicate intended cooperative responses often enough only if the results are, in turn, of interest to hearers.

Consider, for example, a speaker whose purpose in using the word "dog" is to communicate about or to call attention to facts that concern dogs. (This probably won't be an explicitly represented purpose, of course. It can be a purpose of the speaking without that—see chap. 1.) Such a speaker will eventually stop trying to use the word "dog" for this purpose if there is no evidence that it ever has this effect on hearers. Similarly, a hearer whose language-understanding faculties turn his mind to dogs with the purpose of collecting information about dogs whenever speakers use the word "dog" will soon unlearn this response if speakers never use the word "dog" such that it carries information about dogs.

Consider those syntactic forms that get labeled as "indicative" forms in various languages. These forms sometimes have a number of alternative functions, just as one's tongue has alternative functions, being designed, for example, to help both with mastication and with speech production. But no form will be labeled "indicative" unless one

---

Chomsky himself seems to have supposed that arguments for the arbitrary nature of human universal grammar are arguments for its not having been selected for. But every form of animal communication is arbitrary in form if one ignores its concrete history in the specific species—every bird song, every mating display, every pheromone, every danger signal is arbitrary. Each got there to be selected for by an accidental historical genetic fluke or set of flukes. No biologist doubts, however, that these signals have all been coevolved, along with the dispositions to react to them appropriately, through the mechanism of natural selection.

of its central functions is this. It effects production of *true beliefs* that have whatever propositional content the various other aspects of the sentences exhibiting it embody. This effect is often of interest both to speakers and to hearers. Production in hearers of false beliefs may occasionally interest speakers, but it rarely serves the interests of hearers. A hearer unable to interpret the indicative sentences he hears so as sometimes to extract genuine information from them would soon cease to form beliefs on their basis. He might first try out other interpretations of the form, and of other linguistic elements used with it, but eventually he would have to give up on it altogether. And if hearers ceased ever using indicative sentences as guides in forming beliefs, speakers would stop trying to use them to impart beliefs. Production of true beliefs, then, is a linguistic function of the indicative form itself, whether or not a particular speaker and/or hearer have as their purpose to use it that way on a given occasion.

Similarly, a linguistic function of imperative mood sentences is to instigate actions that accord with their propositional contents. If it were not sometimes in the interest of hearers to comply with imperatives—advice, instructions, directions, friendly requests, sanctioned imperatives, and so forth—hearers would soon cease to comply with them. And if hearers never complied with imperatives, speakers would soon cease to issue them. Imperative syntactic forms would either become obsolete or change their functions.

Thus it is that the function of a public language device itself is not on the same level as either speaker purposes or hearer purposes taken alone. Conventional language devices are selected for performing services satisfactory at once to both partners in communication. These language forms are arbitrary, of course, within broad limits. There is nothing magical about the form itself that enables it (sometimes) to serve its memetic function. It can perform that function only because speakers and hearers are trained to respond with and to it in ways that have some stability, each given the expected performance of the other. (These themes are much expanded in Millikan 1984, chap. 4, and in Millikan 1998, 2001a, 2003.)

The functions of conventional language devices considered as such are memetic purposes. But when language parts are used in figures of speech or used as bases for Gricean implicatures, the underlying memetic purposes of these expressions are crossed by the speaker's purposes. Then what the speaker means may not be what the words mean, or it may be more than what the words mean. The very same

expression token then has two purposes derived from two different sources that cross, a literal meaning deriving from its function in the public language and a pragmatic meaning deriving from the speaker's purposes. Public language meaning and speaker meaning often diverge in this way.

I said in the preface to this book that the meanings of signs are not usually considered to be their purposes but rather to be what they represent or signify. But if we look really closely, the memetic purposes of conventional language forms do seem to be what these forms mean. Compare, for example, the following sentences: (1) "Jeanette will stop." (2) "Jeanette, stop!"(3) "Will Jeanette stop?" (4) "Would that Jeanette stop!" These four sentences seem all to *represent* the same thing, for they all represent or refer to Jeanette's stopping. In the first, Jeanette's stopping is asserted; in the second it is directed; in the third it is questioned; in the fourth it is wished for. In philosophical tradition, we say that these sentences all have the same "satisfaction conditions."[5] But though they represent the same thing, they don't *mean* the same. They are not translations of one another. "Jeanette, stop!" translated into French, is "Jeanette, arrete!", not "Que Jeanette arrete" or "Jeanette, arretera-t-elle?"; and this is because only "Jeanette, arrete!" has the same purpose or function as "Jeanette, stop!" Memetic purpose seems to be what is preserved in literal translation of conventional language forms. I will return to this theme in chapter 7.

Returning to a theme from chapter 1, are the purposes of public language forms real purposes, or are they purposes only metaphorically or analogically? They have exactly the same kind of foundation as does any other kind of purpose. The purposes of conventional language forms are as real as any purposes can be.

5. The satisfaction conditions of assertions are usually called "truth conditions." Truth conditions are one kind or example of satisfaction conditions. Similarly, the satisfaction conditions of directives are sometimes called "compliance conditions."

II

Natural Signs and Intentional Signs

3 Local Natural Signs and Information

Intentional signs are signs that can be false or unsatisfied. Natural signs cannot. Black clouds do not mean rain unless it actually rains, and an elevated temperature does not mean illness unless one is actually ill. True intentional signs are sometimes thought to be just natural signs that have been purposefully produced. The information they carry is thought to be just natural information that they happen to have been designed to carry. It is their purpose or function to carry this information. Certainly that is Fred Dretske's claim (Dretske 1986, 1988, 1995). I will agree that true intentional signs often do carry natural information. But to support this claim it will be necessary to understand the terms "natural sign" and "natural information" quite differently from the way Dretske defines them. Also, not all true intentional signs do carry natural information. Some of them are true by accident. Further, I will argue that it is not the *purpose* of an intentional sign to carry natural information. Carrying natural information is merely the usual means by which an intentional sign gets to be true.

In *Knowledge and the Flow of Information* (1981) Dretske defined the notion "natural information" quite strictly. I will argue that his definition is not adequate even to the simplest tasks he wished natural information to perform, and I will propose a description of "local natural information" that I believe is more adequate. The most useful kind of local natural information is carried by "locally recurrent" natural signs. In this and the next chapter, I will introduce locally recurrent signs and local natural information and describe some of their many virtues. In chapters 5 and 6, I will explore their relations to intentional signs. The perceptual and cognitive systems of every animal are deeply dependent on the local natural information found both in the environment and within the organism itself. Without such information there could not be any intentional signs or intentional information.

Dretske's use of the term "natural information" vacillates in *Knowledge and the Flow of Information* (1981) and also in *Explaining Behavior* (1988).[1] Sometimes his usage conforms to his original strict definition in (1981) but other times it gestures more broadly. According to Dretske's original definition and the discussion immediately following it, a signal carries "information" about a source only if the signal's occurrence yields a probability of one, determined in accordance with strict natural and logical necessity, that the source is a certain way. On the other hand, many of Dretske's examples of natural information seem to rest not on strict natural necessity, but merely on statistical frequencies at the source. For example, he tells us that although in some woods a certain kind of tracks made by quail might carry the information that quail are present there, on the other hand, "[i]f pheasants, also in the woods, leave the very same kind of tracks, then the tracks, though made by a quail, do not indicate [carry the information] that it was a quail that made them" (1991, p. 56). Here, not natural law but statistical frequencies at the source end of the information channel appear to be determining whether the tracks carry natural information. Yet Dretske explicitly claims that statistical frequencies are not enough to determine the presence of natural information:

> Even if the properties F and G are perfectly correlated . . . this does not mean that there is information in *s*'s being F about *s*'s being G. . . . For the correlation . . . may be the sheerest coincidence, a correlation whose persistence is not assured by any law of nature or principle of logic. . . . All Fs can be Gs without the probability of *s*'s being G, given that it is F, being 1. (1981, pp. 73–74)

But surely, whether or not there are pheasants as well as quail in these particular woods is a matter of statistical frequency, not natural law. If we look only at natural law, given that pheasants as well as quail could lawfully produce such tracks, there could never be a probability of one that such tracks are produced by a quail, no matter how far these tracks happened to be from any actual pheasants. Logic and natural law do not change over space and time with variations in the distribution of pheasants and quail.

This is an insurmountable problem, as I see it, for Dretske's explicitly stated theory of intentional representation. Nearly all of the kinds

1. See Millikan (2000), appendix B.

of information needed by us, and by all other organisms as well, for securing what we need in an inclement world, is information that cannot possibly be acquired without leaning on certain merely statistical frequencies. Consider a rabbit that needs to know when a predator is near. However she detects a predator, no natural law can require it to *be* a predator that causes her predator detectors to fire. Whatever information channel she uses, it is always nomically possible that non-predators should exist who would activate it. Suppose for the sake of the argument (though very implausibly) that there are unbreakable natural laws that concern the effects of foxes on rabbit sense organs. Still, there surely are no laws that nothing *else* could possibly produce these same effects on rabbit sense organs. For example, no natural laws prevent the introduction of new species.

Moreover, Dretske is often quite explicit that it is only relative to certain channel conditions that the sign carrying natural information must correspond with certainty to what it signifies. Suppose that we contrast, for example, (1) the probability that it is the presence of helium under excitation that produces such and such a spectrum with (2) the probability that it is the gas tank's being half full that produces such and such a reading on the gas gauge. Perhaps the spectrum corresponds to the presence of helium with a probability of one, period. But the gas gauge reading corresponds to half full with a probability of one only on the assumption that there are connections of a rather exact sort between the tank and the gauge. Similarly, when the sounds of the recorded music get first louder and then softer, this may be a sign that the orchestra executed a crescendo and then a decrescendo. But this depends on stable placement of the microphone during recording, and on whether the recording engineer, or someone at the loudspeaker end, twiddled the knobs. Channel conditions of this sort are not considered by Dretske to be part of the signs they transmit. The signal carrying natural information indicates neither what channel it comes through nor, supposing the channel known, what semantic rule to apply in order to read information coming through that channel. Both would have to be independently known.

The question arises, then, what use signals carrying bits of natural information could be to an organism. The mere fact that a signal carries certain natural information seems not to bear on whether a creature could learn anything from encountering that signal or know anything by virtue of harboring that signal in its brain. To learn anything from

such a signal would seem to require that the organism have a means of detecting not merely the information-bearing signal, but also the presence of a particular kind of information channel, as well as a means of knowing what semantic mapping function, what translation rule, to apply. Returning to our rabbit, the difficulty is that no matter what means she uses for detecting the presence of whatever particular kind of external information channel she knows how to use, it is always nomically possible that the presence of some other external circumstances should cause the very same signals to arrive as through this channel, but with different meaning. Just as no natural laws prevent the introduction of new species that would affect the rabbit's senses in the same way given these channel conditions, no natural laws prevent the introduction of new external circumstances not detectable by the rabbit that substitute for these channel conditions. If the capacity of an organism to represent something mentally were to depend on its ability to discriminate that thing from all others in accordance merely with natural law and logical necessity, it is clear that no organism could possibly represent anything distal.[2]

Nor, of course, does the rabbit care that there are no such laws. As long as she is good at detecting foxes in her actual environment, given its actual statistics, all is well. Similarly for the things we humans detect, for all the various things we acquire empirical information about. Our abilities to represent these things couldn't possibly depend on natural information that fits Dretske's strict characterization, nor would the insertion of this kind of information make much difference to our practical lives.

I conclude that a theory of natural information that will help to explain how real animals manage to obtain useful information will need to introduce statistical considerations about the environment in some controlled way. The relevant statistics would have to bear both on conditions at the source of information and on channel conditions. Moreover, a reference to nonaccidental stability of these statistics over trackable regions would need to be built in, for organisms must

2. Jerry Fodor likes to explain the supposed laws that produce covariance between representation and represented in terms of "ceteris paribus laws," "counterfactuals," and what is true in "near possible worlds," but nowhere has he explained with care what ceteris paribus laws are, exactly how the counterfactuals are to be stated, or what the principles are that determine relevant possible-world nearness and why. Following his lead, irresponsible incantations of "ceteris paribus" everywhere pollute the literature on this subject.

somehow keep track of the regions in which the relevant statistics continue to hold. I will try to describe such a new notion of natural information, information of a more user-friendly kind than Dretske described, calling it "local information." Information defined by Dretske's strict rule above I will call "context-free" natural information.[3]

At the same time I will try to speak to another deep problem intrinsic to any informational semantics that defines information by reference to natural law only, ignoring local statistics. The difficulty is that there are no natural laws about individuals. No natural laws are laws just about George Bush or just about the White House or just for the moon. On a theory that recognizes only context-free information, no natural sign can carry the information that it is George Bush who is speaking from the White House or that the moon is out again. Context-free informational semantics is debarred in principle from explaining how there can be representations of individuals. Consonant with this, informational semanticists have invariably confined themselves to a discussion of predicative representations. For example, when Dretske explains what it is for a signal to carry the information that *s is F*, he concentrates entirely on the representation of *F*, telling us nothing about what it would be, say, for *t* rather than *s* to be represented as being *F*. The subject of the intentional representation *that s is F* is not represented. Compare: If Billy and Johnny are identical twins, that a photograph is a photograph of Johnny rather than Billy is not represented *in* the photograph. Dretske describes a signal carrying the information that *s is F* exactly as if it were, in this way, like a photograph of Johnny. The description of local natural information that I will give explains how there can exist natural information concerning individuals, and how the subject term of a natural sign that *s is F* can be naturally represented.

3. In Millikan (2000), appendix B, I use the terms "soft natural sign" and "soft natural information" rather than "local (recurrent) sign" and "local information." I think the term "local" captures the idea we need a bit better. The central idea is that there is a historically positioned domain to which the sign is bound. One might also call these signs "bounded signs," but the correlative term "bounded information" somehow doesn't work. Those familiar with the description of "historical kinds" and "individual substances" from my (2000) will notice that the domains of local signs are bound together in much the same way that historical kinds and individuals are. In both cases, the underlying claim is that to understand the possibility of any kind of cognition we must recognize that contingently existing historical regions of continuity as well as universal natural laws are tapped by cognizers.

We can begin by reflecting on what count as natural signs and their meanings in everyday life.[4] What makes black clouds a sign of rain and birds flying south mean that winter is approaching? Consider Dretske's bird tracks in the woods. If both quail and pheasant leave tracks of exactly the same kind in the woods, are these tracks signs of quail, or signs of pheasant, or are they signs of neither (as Dretske claimed)? If there's no way to tell the difference between quail tracks and pheasant tracks, it seems to me that a natural remark would be "those tracks might mean quail or they might mean pheasant, we'll just have to wait and see." That is, we'll have to wait and see which they really *do* mean, for they do mean whichever was their *actual* cause. No probability of one is in view here, but only a real causal connection. Similarly, Dretske told us that "[t]he red spots all over Tommy's face mean that he has the measles, not simply because he *has* the measles, but because people without the measles don't have spots of that kind" (1991, p. 56). But suppose that scarlet fever can cause spots like that too. What a cautious physician will say, I believe, is something like this: "Those spots probably mean measles, but they could also mean scarlet fever. I think we had better take a culture." Indeed, a doctor might say this even if measles is ten times more frequent in the schools this fall than scarlet fever. Not even a *high* probability is always required for attribution of natural meaning.

And there are examples that are more extreme. Suppose that I take my daughter's mitten lying in the path to our front door to be a sign that she is home from school. For me to be right, the connection must not be accidental. If she dropped her mitten on the way to school rather than the way home, then even if she is home, her mitten is not a sign of that. As Dretske said, merely correlating (in this case, merely coinciding) is not enough to make one thing mean another. But what is the general probability of a mitten in the path meaning that a daughter is home from school? How would one calculate that? Also, what would count here as channel conditions? Again, suppose you are conspiring

4. The passages immediately following may look like conceptual analysis of the notion "natural sign." I have often complained about the tradition of conceptual analysis in philosophy. Philosophy should be concerned primarily with theory construction. Ordinary usage should be described by lexicographers, who do a much better job of it anyway. But it is true that a careful look at the subtleties of usage can sometimes help us to "assemble reminders" about the complexities of the phenomena that we are trying to address. Ordinary usage can be merely idiomatic, superficial, indecisive, erratic, vague, gerrymandered, inconstant, revealing nothing of theoretical interest. Other times it can be surprisingly deep and subtle, following underlying natural contours of great interest.

with Joe in an assassination attempt. He hands you a briefcase with a homemade bomb in it and says, "Here is the briefcase for you to leave in his office. Be careful though. When you hear it beep twice it will explode in two minutes. Leave the room quickly." The beeps will be a sign that the briefcase is about to explode, and they will be causally connected with that. But surely beeping in briefcases doesn't generally precede their exploding.

What seems to be happening in these cases is that there is a real causal connection between two things such that in the circumstances one does depend on the other, and that given what one already knows, one is able to track that connection and hence come to know or suspect one thing on the basis of knowing the other. In the mitten case, the inference is to the best explanation one thinks of, and this explanation, as it happens, is correct. In the bomb case, the inference is based on knowledge of causal principles plus quite exact knowledge of the channel conditions. The central thing common to all these examples of natural signs, I suggest, is that in each case it is possible for a true belief to be reached about one thing from knowledge of the other, the transition from one belief to the other being based on prior knowledge or experience, where the truth of the belief reached will not be accidental because the connection in thought correctly and nonaccidentally tracks a dependency in nature. A natural sign of a thing is something else from which you can learn of that thing by tracking in thought a connection that exists in nature. The notion of a natural sign is at root an epistemic notion.

The use of isolated natural signs, such as the mitten in the path and the beeping briefcase, clearly depends on having a good bit of prior knowledge with which to combine one's observation of the sign. The use of isolated signs could not then serve as the original foundation on which perception and knowledge is built. Such a foundation would require signs that can be used alone or one by one. Further, for a system to acquire the capacity to use a natural sign either by means of natural selection or by means of learning, the sign would need to be one that recurs, and recurs with the same natural signification or meaning. For these basic natural signs, it must be true with some generality that the same kinds of signs are connected to the same kinds of signifieds. There must be a correlation between similar signs and similar signifieds, and it must be a correlation, as Dretske has said, that is not accidental.

Correlations are defined relative to reference classes. Within what reference class must *A*s be correlated with *B*s for a particular *A* to be

an instance of a recurrent natural sign of *B*s? Dretske says that if *in his woods* only quail leave a certain kind of tracks—call them "ε-tracks" for the quail's three-toed foot—then ε-tracks in his woods are signs of quail. He implies that if *in the woods next door* only pheasants were to leave ε-tracks, then the ε-tracks next door would be signs of pheasant. But why take Dretske's wood and the wood next door as separate reference classes? Why not take both together as our reference class?

Suppose that looking worldwide, we find ten more species that leave ε-tracks, and there is no strong correlation between ε-tracks and any one of these species. Why not pool all these together and assert that ε-tracks are never recurrent signs of any particular species of bird? Or suppose that in my American woods only poisonous mushrooms have a certain reddish color underneath whereas in the French woods only edible mushrooms do. Is that certain reddish color underneath a mushroom ever a recurrent natural sign of anything? What reference class should we use?

We should be clear that this problem cannot be solved by relativizing natural signs to arbitrary reference classes. We cannot just say that this particular ε-track in Dretske's woods exemplifies a recurrent natural sign of quail with reference to the class of ε-tracks in Dretske's woods though not with reference to the class of ε-tracks worldwide. We cannot just say that the reddish color under my American mushrooms is a recurrent natural sign of poisonousness with reference to the class of mushrooms in America but not with reference to mushrooms worldwide. For what is to prevent arbitrary divisions of the reference classes? Suppose I combine a highly discontinuous set of sectors from Jill's woods in Massachusetts with a highly discontinuous set from Jack's woods in Minnesota, such that although each of these small sectors has pheasants living nearby, mainly quail live within my gerrymandered class. Then I name the whole of this scattered terrain "Q-woods," claiming that each quail track in Q-woods exemplifies a recurrent natural sign of quail with reference to Q-woods. I also admit, of course, that the infrequent pheasant tracks found in Q-woods nicely exemplify recurrent natural signs of pheasant relative to some other reference classes, just not with reference to Q-woods. Relativized in this manner, the notion of a recurrent natural sign obviously spins its wheels. It does no work.

Well, what kind of work *should* it do? What we would like, I suggest, is for the notion of a recurrent natural sign to explain *why* a person might be able to use the recurrent sign as an indicator of its signified

with some success. To be genuine, this explanation must have its footing in nature. Consider, for comparison, the fact that John, who happens to be forty years old, five feet ten inches tall, and to like sports, has a mortgage. This fact would not be *explained* by citing the fact that the average man who is forty years old, five feet ten inches tall, and likes sports has a mortgage. Similarly, that the inference from "This is an ε-track in Q-woods" to "This is the track of a quail" is likely to yield a true conclusion is not *explained* by citing the statistics on quail in Q-woods. It is a clear example of the fact, now well known, that a logical derivation from true premises is not always an explanation. What is needed is some way to delineate *relevant natural* classes. We need to require that *A* and *B* are correlated within a relevant natural reference class for one to be a recurrent natural sign of the other.

Another problem needs to be solved here as well. Dretske told us that even a perfect correlation between *A* and *B* is not enough to ensure that *A* carries natural information about *B*, because a perfect correlation could be perfectly accidental. Of course, a particular *A* does not mean *B* unless it actually coincides with a *B*, but Dretske demands more. He wants *A*s that mean *B*s to correlate with *B*s not by accident but for a reason. But what is it, exactly, for a correlation not to be accidental? The standard thought here is that a correlation between *A*s and *B*s is not accidental just when there is a causal connection between *A*s and *B*s, the *A*s causing the *B*s or vice versa, or perhaps something else causes both the *A*s and the *B*s. But suppose that this causing itself happens accidentally. The mechanisms by which the *A*s cause the *B*s are each time entirely different. Each *A* is in some circumstance such that it causes a *B*, but these circumstances are all different. From chaos a correlation accidentally emerges. Are we to say that *A*s are recurrent natural signs of *B*s?

Here is my suggestion. What we are trying to do is to construct a notion of recurrent natural sign that will help us to understand how it is possible for an animal nonaccidentally to recognize the recurrence of such a sign so as to use it effectively. The project is not conceptual analysis but theory construction. If *A*s in some reference class are to be recurrent natural signs of *B*s, we want it to be possible for an animal to come to *learn* of *B*s from encounters with those *A*s, where what counts as learning is acquiring true beliefs nonaccidentally. From the standpoint of natural epistemology, the cause of one's moving from encounters with *A*s to representations of *B*s (or to accommodations to *B*s) should be connected with the reason these moves are sometimes

correct. Now one's moves of this sort will be based, in central cases, on one's experience—or someone else's experience, or the experience of the species—of a correlation of *A*s with *B*s within some sample. The inference then predicts that the correlation will continue to various new samples encountered. The inference will succeed "for a reason" rather than "by accident" only if there is a reason why the correlation *persists* from the old sample into the new. A natural reference class for a sign—the natural domain within which certain *A*s are "locally recurrent signs" of certain *B*s—is a domain within which the correlation of *A*s with *B*s extends from one part of the domain to other parts for a reason, and it must be a domain that it is possible for an organism to track.[5]

Here are some simple examples. Suppose that every ball in the urn is black. Then every ball I pull out today will be black. That a ball comes from this urn is perfectly correlated with its being black. Now consider tomorrow. Granted nothing disturbs the urn overnight it will be no accident that the correlation I discovered today holds tomorrow. Balls left undisturbed do not change their colors overnight, nor do urns left undisturbed change their contents. The correlation continues for a reason. Being a ball in this urn is a local sign of being black for as long as the urn remains undisturbed.

What about the balls in the next urn? Depending on the histories of these two urns, there may or may not be a reason why the contents of one urn would reflect the contents of the other. If there is a reason for the contents to match—say, they came from the same factory in response to the same middleman's order—then the relevant recurrent sign domain includes the contents of the second urn as well. Otherwise not—not even if all the balls in the second urn do happen to be black. What this example brings out, however, is that in order to make use of a local sign, one needs a method of tracking, or recognizing, or managing to remain within, its local domain. In this case, one would need a way of recognizing other urns that came from the same factory in response to the same middleman's order. But then, in the simpler case above involving only one urn, a way of tracking the relevant domain was also necessary. One would need to keep track of which urn it is that contained only black balls yesterday.

Suppose that this morning my gas gauge reads half-full when it is three-quarters full, and, later today when the tank is half-full, the gauge

5. This description of local natural signs is designed, in part, as a way to clarify what it is for Dretske's "channel conditions" to remain constant and how it is possible for an organism to *keep track* of the domains in which channel conditions remain constant.

reads one-quarter full. What about tomorrow? The channel conditions that obtain today between the height of gas in my tank and the gauge readings concern bits of hardware and wiring that, should nothing outside disturb them, strongly tend to remain in place in accordance with natural conservation laws. It will be no accident if my gas gauge continues to underestimate the amount of gas in my tank by about the same amount tomorrow. The readings on my gauge are locally recurrent signs of the amount of gas in my tank.[6]

Suppose that most ε-tracks now being made in these woods are by quail. Quail in these woods this week causally result in quail being in these woods next week and next year, and also in quail soon being in the woods next door. For good reason ε-tracks are repeating themselves, over and over, now and then, here and there, in these woods and in contiguous woods, each time connected to quail. The correlation persists and spreads over a local time and space for a reason. There is a causal stream in which the correlation holds.

In general, it is because many conditions persist and/or spread or replicate themselves over time and space that there are locally recurrent natural signs belonging to recurrent sign domains. Identical signs that come from other domains, for example, from different regions of space and time, may be local natural signs of something quite different. Clouds of a certain sort on this side of the mountain may mean rain but on the other side not. That is because certain contours on each side of the mountain persist, and because the causes of air moving in and the places the air comes from on the two sides tend to persist. The local statistics for each place repeat themselves for a reason.

Suppose that all the spots that appear on the children in our school this week are from measles. And suppose that the children pass measles on to their friends, brothers, and sisters. Thus it happens that spots due to measles appear on lots of other children in our school next week. In this school, spots will be a recurrent natural sign of measles until the epidemic is over. If an epidemic of scarlet fever should now intrude as well, then there will be two separate grounded persistent

6. Dretske discusses exactly this kind of case in his (1981). I am hoping to articulate somewhat better what I believe he was groping for there in his discussion of channel conditions. I believe that his attempt to show how the idea of natural information could explain what knowledge is, coupled with his assumption that knowledge requires certainty, got in his way. He was also concerned about what he called the "Xerox principle." Information that there exists information *that p* ought itself to be information *that p*. For this reason too he wanted probabilities of one for his information. Validity of the Xerox principle for locally recurrent natural signs is discussed in chapter 4.

correlations in our schools, one between this kind of spots and measles, the other between this kind of spots and scarlet fever. Within the natural reference class that is our schools, some of these spots are recurrent natural signs of measles, others of scarlet fever. ("These spots might mean measles. But they might also mean scarlet fever. I think we had better take a culture.") On the other hand, each of these natural signs may be viewed as having its own separate gerrymandered geographic domain, determined by the contingent contacts of individual children carrying one disease or the other with children to whom they have passed these germs. Normally one would not attempt to track these two gerrymandered domains separately. But if one of the diseases is serious enough, exactly this sort of tracking will be attempted by physicians and health authorities, in order to recognize natural signs of the disease more quickly and accurately.

Consider now the look, the appearance, of Johnny's face. Your encounters with that look are likely to be highly, indeed perfectly, correlated with your encounters with Johnny. This is because natural conservation laws plus principles of homeostasis built into Johnny tend strongly to preserve Johnny's appearance from one day to the next, indeed, over many years. The look of Johnny's face is a locally recurrent natural sign of Johnny. It carries local natural information as to the presence of Johnny. There are no laws that concern individuals as such, but there may be very numerous local well-grounded correlations that do.

To interpret a locally recurrent sign successfully you must keep within its natural domain. You must stay within the boundaries. On the other hand, it may not be necessary to discriminate the boundaries of the sign domain in order to stay within them. The rabbit is likely to be born and to die without ever leaving the domain containing various of the locally recurrent signs of fox that it recognizes, and you are not likely to leave the domain in which that face (its look) reliably signifies Johnny. In many fields of knowledge, becoming an expert involves learning to recognize subtle perfectly diagnostic signs for a variety of different kinds, that is, local signs outside of whose boundaries one doesn't traverse. Or it may involve learning how to track various causal streams of signs. The modern doctor diagnosing signs of infectious diseases will be helped by understanding as much as possible about what causes the spread and the boundaries of each disease.

Knowing how to track is most obviously important if one needs to be able to identify each of various similar individuals, individuals whose signs are hard to distinguish. To be able to recognize my glass from among other identical glasses at a party so as to drink out of it and no other, I have to keep it in my hand, or remember where I put it down, or who took it away offering to refill it, or, should I accidentally drop it, I must know the way glasses are likely to fall and roll, and so forth. This means I must understand quite a lot about the possible trajectories of individual objects of that general kind. Squirrels behave differently from drinking glasses. If I want to recognize the same squirrel again as it jumps among the branches, I will have to use quite a different method of tracking. Squirrel shape, color, movement style, behavior (dropping hickory nuts from trees, chattering, squeaking), and other characteristic squirrel effects (ways the branches quiver and so forth) are each recurrent signs of squirrel. As such they may also be signs of *Chipper* specifically, but only as manifested within a certain very much more local domain that may often shift in location.

Failure to account for our capacity to represent individuals in language and thought has been, perhaps, the most serious failing common to contemporary naturalist theories of content. This failing is not often discussed, but it should certainly be considered embarrassing for a tradition that has strongly rejected the classical analysis of individual concepts as composed of definite descriptions. If a cognitive mechanism is to have producing representations of individuals as one of its functions, there had better be some practical way of producing these representations. That there exist local natural signs of individuals will help us to explain this possibility in chapter 4.

But local sign domains may sometimes be recognized for practical purposes without tracking, and without understanding much about why they flow where they do. In northern climes, mushrooms that look like that are always edible, whereas in southern climes they are not, except in certain parts of South America. You can know that without knowing why. You can recognize a sign of a local domain. If the domain of a local sign is signed by another sign or signs, however, this does not make the whole complex involving both signs together into a context-free sign, for the sign of the domain, as a sign of a portion of the actual world, will not itself be context-free. There is no way of adding to a local sign so as to completely free it from context.

The notion of a recurrent local sign is designed to explain how it is possible for an animal to use natural signs to collect information about its world. The kind of knowledge that earthly creatures have is knowledge applicable in the domains they inhabit, not knowledge for arbitrary nomically possible worlds, nor for other domains, regions or eras within the actual world. If the question arises how strong the nonaccidental correlation between *A*s and *B*s must be within a domain for the *A*s that do correspond to *B*s to count as locally recurrent signs, the following is what is important. A strong enough correlation to count in determining a local sign to be such is one that is strong enough to have actually influenced sign use, either through genetic selection or through learning. For the point of introducing locally recurrent signs is as a tool for understanding how perception and cognition are possible. Our concern is only with natural signs that are actually used by organisms. Suppose that an animal has only crude and inaccurate ways of tracking a local sign's domain. Suppose that the animal tracks only a rather vague domain that includes other similar signs as well, and it is unable to distinguish accurately among these signs. (Compare: Both measles spots and scarlet fever spots are found in our schools, and no one knows how to take a culture.) Still, I will soon argue, the animal may be able to use these signs effectively in the process of producing intentional signs. Intentional signs, I will argue, do not have verificationist meaning. No probability of one, nor even a particularly high probability, needs to be involved anywhere in the origin or use of an intentional sign. But here I have gotten ahead of my story.

One last comment about locally recurrent signs. As I have described them, it is not necessary that there be any causal connection between a locally recurrent sign and the affair that it signifies. This is because the relation between tokens of the sign and affairs signified may be reiterated over a domain merely because the signs and the affairs each persist, maintaining the same relation to one another. A nice example of this comes from Dretske (1986), though it was not his intention to illustrate this point. The magnetosomes of certain anaerobic bacteria are tiny magnets that serve as sense organs. In northern-hemisphere bacteria, they steer the bacteria toward magnetic north, which means steering them toward geomagnetic north, which means steering them into deeper water which contains less oxygen. (Oxygen poisons these bacteria.) The direction in which the magnetosomes point are locally recurrent signs of the direction of lesser oxygen. But there is no causal connection between the direction of magnetic north and the direction

of lesser oxygen. The information is locally recurrent information, the reference domain of which is extended through time and space because magnetic north continues to equal geomagnetic north so long as no local bar magnets (and so forth) intervene, geomagnetic north continues to remain in the same relation to the earth and hence to deeper water, and lesser oxygen continues, for entirely independent reasons, to remain where the deeper water is.

4 Productivity and Embedding in Natural Signs

Before turning to the relation between recurrent natural signs and intentional representations, I want to discuss two features well known to characterize the parade cases of intentional signs, namely, conventional linguistic signs, but that are not generally recognized to characterize natural signs. First, locally recurrent natural signs are productive. For in their own way, they are something like compositional. Second, locally recurrent natural signs allow embedding.

From now on by "natural signs" I will mean locally recurrent natural signs, unless I say otherwise. Natural signs are not just objects or properties of objects. They are occurrences or states of affairs—I will say "world affairs" or just "affairs." World affairs, as I understand these, are structured or articulated aspects of the real world roughly of the sort that are sometimes called "truthmakers." I will not say "truthmakers," however, because it so strongly suggests correspondence to just one particular sort of sign, namely, a sign having clear subject–predicate structure that is also sensitive to a negation transformation. World affairs are mind-independent, unless minds happen to be part of their subject matter. Natural signs are structured world affairs and the things of which they are signs are also structured world affairs, analogous to the correlates of complete sentences rather than open sentences or sentence parts. Strictly speaking it is not the black cloud that is a sign of rain. Rather, the structure that is a black cloud in the sky at a certain time, *t*, moving toward a certain place, *p*, may be a sign of the structure that is rain occurring shortly after *t* at *p*. Similarly, the structure that is an ε-track at a place, *p*, and a time, *t*, may be a sign of the structure that is a quail passing *p* shortly prior to *t*. Structures are often referred to using "that" clauses. Another way to put this same point, then, is to say, "that there is a black cloud at a certain time, *t*, moving toward a place, *p*, is a sign that it will rain

at *p* shortly after *t*."[1] Signs that tell by their own time and/or place something about the time and/or place of something else are very simple signs, but it is crucial to see that, as signs, they are in this way *structured*.

Their structure determines their meaning architecturally. The meaning of the sign is determined as a function of values of significant variables or determinables exhibited by the sign. Put another way, the meaning varies systematically to parallel significant (mathematical) transformations of the sign. It is often said that the meaning of public language sentences is determined "compositionally." Exhibiting compositionality, in this sense, is just one among other ways of being architecturally structured. Representations that exhibit compositionality are transformed into other representations in the same system by rearranging the same parts or by substitution of parts. It is not compositionality in particular, but architectural structuring more generally, that yields the productivity of sign systems, their capacity to say new things or to give new information.

In the simplest cases, the significant variables in natural signs often are merely time and/or place. But even something as simple as tracks in the mud can include other significant variables as well. The size of the track may be a sign of the size of the animal, the distance between the tracks a sign of how fast the animal was moving, the angle or depth of the track may be a sign that the animal is pregnant, and so forth. Then size, and distance between, and angle or depth, are additional variables that help, structurally, to determine a more complex affair as the one signified.

In Connecticut, geese flying south are local recurrent natural signs of the approach of winter. Now consider what it is that recurs. *That* geese are flying south through Connecticut on November 25, 2003, is a natural sign *that* there will be winter in Connecticut soon after November 25, 2003. Exactly this "same sign"—another token of the same sign type—will recur, I suppose, if later that same day more geese fly south through Connecticut. But that kind of "recurrence" of "the same sign" does not teach us anything new. We already knew from the first sign that winter was approaching Connecticut. If natural signs recurred only in that sense of recurrence, they would be useless. Learning what a certain sign meant would not

1. Compare Wittgenstein: "That *'a'* stands in the relation *R* to *'b'* says *that aRb*" (*Tractatus Logico Philosophicus* 3.1432).

generalize; one could not learn anything from encountering "that same sign" again.[2]

By "recurring" in the phrase "recurring natural sign" we need to understand not that the same sign, in the above sense, *itself* recurs, but that the same sign-signified *relation* recurs. The interesting thing that recurs is the relation of the time of some geese passing through to the time of some winter. That is, what recurs is the applicability of a certain natural, structural, semantic mapping. The structure of the sign includes its own time and place. What is signified depends on the determinations of these determinables, on the arguments for these variables. The rule determines the determinate structure of the signified from the determinate structure of the sign. It fills in where and when a winter will occur given where and when the geese are flying.

Put another way, what "recurs" in the case of recurrent signs are other members of the same *system* of signs, where a sign system consists in a set of possible sign types, designatable transformations of which ("transformations" in the mathematical sense) correspond systematically to transformations of what they signify. Consider ε-tracks again. If, on a particular ε-track, we perform the (mathematical) transformation "move it ahead five years, move it south two miles, and cut it down to half its size" and perform the *same* operations on the place, time, and size of the quail indicated, we arrive at another (possible) natural sign in the same sign system. Notice that in this sort of case, times that represent times and places that represent places are perfectly ordinary ingredients of natural signs; they are not, for example, special "indexical" elements. Exactly similarly, size represents size in this example, but size surely is not an indexical. We can invent a special term for the case where a sign element represents itself. Call these sign elements "reflexive."

In this manner, "recurring natural signs" are really recurring relations between, or functions from, signs to signifieds. Call these functions "semantic mapping functions." Semantic mapping functions define isomorphisms between the set of possible signs in a certain sign domain and the set of their possible signifieds. Natural signs are

2. A friend asks: "But what is wrong with the other approach: type together all flockings of geese; that sign type then carries the information 'winter soon'?" What is wrong is that the phrase "winter soon" does not express a piece of information. It does not name a world affair. You have to place that phrase at a particular time and in a particular place in order to get it to refer to a world affair. To carry information is to refer to some state of affairs.

abstract "pictures" of what they represent; indeed, sometimes, as in the example just given, these pictures are not even particularly abstract. When natural sign systems have more significant variables than just time and/or place, it becomes more natural to call them not just "natural signs" but "natural representations," and to speak of the (local) "information" they carry. Thus a close relation emerges between an "informational" theory of signs and a "picture" theory of signs.

The semantic mapping function that mediates between a domain of signs and its corresponding domain (range) of signifieds can be a function that maps many different signs onto the same signified. Consider as an analogy the plus function in mathematics. Many different pairs of numbers may map by addition to the same sum, for example, 5 + 5, 4 + 6, 3 + 7, 2 + 8, and so forth. A single function may operate on a variety of different arguments or on different values of relevant variables to produce the same value. Consider various images of a three-dimensional object as reflected in a mirror. Depending on the angle of the mirror and its distance from the object, a variety of patterns of light reflected from the mirror will all correspond to the same object shape and color[3] in accordance with a single many–one rule of projection. There is a mapping from the domain of patterns reflected from the mirror into the range of shapes and colors that have caused these patterns, given uniform channel conditions of lighting and so forth. It is not that there is a probability of one, given a certain pattern reflected from the mirror, that a certain shape is present before the mirror. But there is a domain in which this mapping holds where channel conditions continue to exist or to recur (revolutions of the earth) for a reason.

Where one structure represents another structure, it is natural to speak of aspects of the first that correspond to definite aspects of the second as themselves representing or being signs of those aspects. So we say that the place of the track signifies the place of the quail and the size of the track signifies the size of the quail. But it is essential to keep in mind that this kind of sign–signified relation is only derivative. It exists only relative to the mapping of the *complete* sign to the *complete* signified, and relative to the larger sign system from which it has been abstracted. It is a serious mistake to suppose that the architectural or compositional meaning of a complex sign is derived by combining the prior independent meanings of its parts or aspects. Rather, the

3. The same partial shape, to be exact, allowing that not all the same parts of the object are reflected in each of these views.

meanings of the various significant parts or aspects of signs are abstracted from the prior meanings of complete signs occurring within complete sign systems. It is not that place means place and size means size and an ε-track means a quail and that when you combine these you have a sign meaning there is a quail of a certain size at a certain place. Obviously time does not mean time, nor does size mean size, not out of context. (Similarly, words do not have meanings first and then get combined into sentences. Nor does the ability to think begin, say, with the ability to think "horse" for horses and then other parts of propositions get added on later. But once again, I am ahead of my story.)

Let me illustrate these principles with some richer examples. I have said (chapter 3) that assuming a gas gauge is connected up correctly, the position of the needle on the gauge may be a recurrent natural sign of the amount of gas in that car. The hookup and surrounding conditions that cause this particular gas gauge to work right today will probably sustain themselves so that it continues to work right tomorrow. The reference class containing readings of this gauge, from the time it is initially installed until the time that it first breaks down, is the natural domain of a locally recurrent natural sign. The semantic mapping function for this recurrent sign is a function from positions of this pointer on this gauge at various times to amounts of gas in this car's tank at those same times. But, of course, unlike the case of the size of the ε-tracks in our previous example where transformations on track *size* corresponded to transformations on bird *size*, here transformations on pointer *positions* map to transformations on gas *volume*. (This semantic mapping function need not be a linear function, of course, to be a natural semantic mapping. In most cars, in my experience, this semantic mapping function is not linear.) In a similar way, every individual well-made undamaged measuring instrument, gauge, meter, or scope of any kind produces locally recurrent signs within its own individual domain that map by some semantic mapping function onto its own individual range of signified world affairs.

Now consider modern gas gauges more generally. Is there a reason why one can learn from experience with a few gas gauges how to interpret others? Of course there is. These gauges are, in general, purposefully made so that people who know how to read one will know how to read another. The idea of having a gas gauge and the rough design of the dashboard are copied from one model of car to another. There is a reason why learning how to read one gas gauge prepares one to read

another. Don't trust what looks like that needle when you go up to Mars, but here on Earth, the positions of needles on gas gauges all fall in the same roughly defined locally recurrent sign domain.

Notice, however, that an additional sign variable has been added when we move from the domain of a single car's gas gauge readings to the larger domain that includes readings of gas gauges on other cars. A "which car?" variable has been added. That the gas gauge pointer *of car* C_1 is partway up signifies that the gas tank *in car* C_1 is partly full. Just as a time can reflexively represent a time or a size reflexively represent a size, an individual can reflexively represent an individual. If you don't know which car the gauge is in, clearly you won't know which gas tank is partly full. The reflexive "which car?" variable is crucial. It supplies the subject term for the information supplied by the pointer, the subject term in this generalized sign domain being variable.[4]

Of course, pointing needles on gas gauges in this broader domain are not nearly as well correlated with gas levels in gas tanks as is the pointing needle on the working gauge of most individual cars. There are rather large differences in the semantic mapping functions that map needle positions to gas levels in different individual cars. Mine, for instance, reads a quarter full when it is about half full. Gas gauge readings, within this broad domain of recurrence, are not very accurate signs.

Here are some other examples of articulation within local natural signs. Owing to conditions prevailing and persisting in prevailing in most places on Earth, the world affair that is the pattern of sunlight, including wavelength, intensity, and direction, passing though each point of unoccupied space at each moment in time contains a great quantity of locally recurrent natural information. When the patterns passing through small spatially and temporally contiguous regions of points are pooled, the amount of information is vastly increased. Eyes, especially lens eyes, have been designed by natural selection to convert the locally recurrent natural representations found in the light surrounding an organism into natural representations having semantic mapping functions that are more user-friendly. The patterns of light typically striking the retinas of human eyes over small continuous

4. So far as I know, Dretske never addresses the issue of how the subject terms of propositions such as his "*s* is *F*" ever get represented either by natural signs or by intentional signs.

intervals of time are or contain naturally recurrent signs of a variety of kinds of world affairs.

First, there is a great deal of local information contained about the spatial layout of the immediately surrounding environment. Much effort has been spent by students of perception trying to describe, as exactly as possible, the natural semantics of these natural representations. Within naturally continuing domains, they often contain local natural information, for example, about the distances, sizes, shapes, colors, and textures of currently nearby objects. In these natural representations, as with the quail tracks, the time of the representation corresponds to the time of what is represented. And although the location of the representation is not the same as the location of what is represented, the second is a direct function of the first. The location is not strictly a reflexive element of the sign, but we might call it a "*relative* reflexive" (to echo "relative adjective"). Similarly, if one inch on a blueprint stands for one inch, length is a reflexive element of the blueprint sign. If one inch stands for one foot, length is a *relative* reflexive. The semantic mapping function transforms a determinate of the representation into a determinate of the represented that falls within the same determinable range.

A natural sign of a natural sign of an affair is itself a natural sign of the same affair. That is, Dretske's "Xerox principle" is preserved for local information (Dretske 1981). This assumes, of course, that what is meant by a "natural sign of a natural sign" is not merely a natural sign of a certain sign *form*, since the same sign form (say, a certain shape) may be a local sign of different things in different local domains. Rather, a natural sign of a natural sign has to signify the presence of a sign form *within a certain domain*. For example, a photograph of an ε-track may be a natural sign of an ε-track having once born a certain relation to a camera (see the paragraphs on photographs below), but if it has nothing in it to indicate the particular ε-track *domain* in which this ε-track occurred, it is not a natural sign of a natural sign in the sense intended.[5]

5. We can be more exact. Assuming that the picture is a sign of an ε-track having occurred on Earth during a fairly specific historical era, within which local domain such ε-tracks were correlated both with pheasant and quail, then we might say it is either a sign of a sign of pheasant or a sign of a sign of quail. ("These spots might mean measles or they might mean scarlet fever.") Depending on which, the picture itself is either definitely a sign of pheasant or definitely a sign of quail. So the Xerox principle still holds. (Thanks both to Gunnar Björnsson and to Nicholas Shea for requesting clarification on signs of signs.)

Recall now that certain combinations of size, shape, color, or texture may comprise local signs of the presence of a certain kind of object—of fox, for example, or of elm tree. Or they may combine to yield a local sign of the current presence of Jill or of Johnny. Here again, time stands for time, and the location of the properties shows the location of fox, say, or of Johnny. A retinal image occurring at location l_1 at time t may be a natural sign of a certain combination of size, shape, color, or texture at a certain location, and hence may be a natural sign of a natural sign, say, of fox, or of Johnny, being at l_2 (a closely related place) at t. But the time and place of the retinal image shows the time and place both of the properties and of what they in turn signify. So the effect is as though only part of the image-as-sign—the part that's left if we omit place and time of occurrence—was a sign of a sign—more accurately, a sign of a part of a sign. Then one sign appears to be embedded within the other. A sign of a sign of Johnny thus becomes a sign of Johnny, although neither sign, of course, is a complete sign apart from spatial and temporal context. The semantic mapping functions for such embedded signs result from applying the semantic mapping functions of succeeding signs on the route one after the other. The resulting function, if grasped, can also stand alone and be applied directly.

Call an affair signified by a sign, B, where B is in turn signified by another sign, A, a "more distal affair" signified by A. Then say that B is part of the "route" from A to that more distal affair. Given that the geese passing through is a sign that winter is on the way, and given that the presence of fresh droppings by the pond is a sign that the geese are passing through, the presence of the fresh droppings is a sign of the more distal affair that winter is on the way, and the geese passing through is part of the route from the droppings to this more distal affair. Similarly, a combination of the shape, color, texture, and location characteristic of fresh goose droppings occurring at a place may be a local sign of the presence of fresh goose droppings at that place, and hence signify the more distal affairs that the geese are passing through and that winter is on the way, the second affair being part of the route to the third and hence also to the fourth and furthest affair. And if certain patterns occurring at a certain place and time on the retina are a local sign of the occurrence of the kind of location, shape, color, and texture that are a local sign of goose droppings, then these patterns have the four previously mentioned affairs all as more distal signified affairs, each being part of the route to the next. Of course, different natural signs may signify the same more distal affair by different routes, or by

routes that are only partially the same. Gathering that there are quail in the wood by seeing their tracks and gathering this by seeing their stray feathers utilize retinal signs indicating different routes to a common distal affair. (Of course, retinal patterns themselves are not perceptions or intentional signs of anything, certainly not just as such.)

Natural signs may be interpreted at any level of embedding or at more than one level of embedding. Retinal patterns of the sort just described might result in the observer's recognizing certain colors and shapes but not in recognizing goose droppings as such, or in the observer recognizing goose droppings but not that geese are passing through or that winter is on the way. Moreover, since there is always a semantic mapping function that goes directly from a sign of a sign to the more distal affair signified, it is often possible to recognize distal signifieds without recognizing all or any of the signs on the route. We recognize colors and shapes, not patterns on the retina. Similarly, a creature may recognize foxes or Johnny without recognizing the colors and shapes that locally signify foxes or Johnny. It is well known, for example, that infants are able to recognize Mama, doing so, of course, only on account of her having quite definite perceptible properties that signify her locally, but that they do this years before they have concepts of any properties at all.[6]

Where retinal patterns occurring at certain times and locations are local natural signs of other things occurring at the same times and at related locations, the times and places of the retinal images are, once again, perfectly ordinary parts of these signs. There are no "indexical" elements in a retinal image. Especially important, the retinal representations of observed individuals are perfectly ordinary. They are just like natural signs of natural kinds such as *fox* and *elm tree*. What makes a retinal image be a natural representation of Johnny is not, for example, that it is caused by Johnny, but that it contains a local sign of Johnny, a sign of the presence of certain features, which are, in turn, a recurrent sign of Johnny within a certain locale. Parallel to this, a retinal image signifying the presence of *fox* will also be a sign of the presence of Reynard the fox, but not indexically, and not merely because it is caused by Reynard.[7] It will be a very local sign, correctly readable only by someone able to keep track of Reynard's very local domain. Indeed,

6. More extended discussion of these matters is in Millikan (2000), chapter 5.

7. It must, of course, be caused by Renard to be a natural sign of him, however, just as the spots on Johnny's face must have been caused by measles to be a natural sign of measles.

if there are lots of Reynard's twin brothers playing with Reynard, it may be that keeping track is not possible without, quite literally, keeping an eye on him.

On the other hand, keeping track of an individual is often accomplished by recognizing recurrent signs of a more general type exemplified by the individual. In a certain locale, all signs of fox may be signs of Reynard. Suppose you are trying to keep track of a squirrel that you just saw running up a tree in your yard. Call him "Scamper." You will track Scamper by catching various glimpses of squirrel shape or, just small-animal shape, then seeing where the branches are shaking, seeing or hearing where the hickory nuts fall, perhaps by hearing the squeaks, and so forth. Any sign of squirrel in this immediate vicinity is a sign of Scamper. In this manner you tell where Scamper is now and what he is doing. Alternatively you may be able to recognize Scamper by signs of the presence of a more specific kind of animal. He is the neighborhood squirrel with only half a tail.

Next consider photographs. Unlike a retinal image, a photograph, considered as a member of a certain local sign domain, contains no information as to the time or place of anything. It does, usually, contain the information that there was once a camera, c, and a time, t, and a spatial layout, s, such that s was in front of c at t. But that sort of information, taken just by itself, is completely general. It goes, as it were, under existential quantification. On the other hand, a photograph may contain the information that it was a fox, or that it was Johnny, that was in front of a camera, and that this object had such and such additional properties. It does this by being a sign of a kind of spatial layout, and hence a sign of properties, that are locally recurrent signs of the presence of fox or of Johnny. A photograph tacked to the cabin wall where I spent childhood summers showed a brown short-haired dog wearing a coat and tie standing on its hind legs beside a long-haired white dog wearing a dress and a lady's hat. It interested me enough that I still have not forgotten the fact that there once existed an x and a y such that x was a brown short-haired dog . . . and so forth. We seem to be peculiar among the animals in being able to record and make use of this very general kind of information (chapter 19 below).

A photograph may contain as a distal affair the local information that Johnny, at a certain rough age, was once flying a purple kite. But again, it is not merely because Johnny caused the picture that it contains a natural sign of Johnny. It represents Johnny because it shows characteristics that are distinctive of Johnny, distinctive at least in a local infor-

mation domain. Suppose the photograph was of a particular copy of the *New York Times* and showed nothing else. There would be exactly one individual copy of the *New York Times* that that photograph was a photograph of. But the photograph would not contain any information about that particular copy of the *New York Times*, because it could not inform us as to *which* individual copy it was a photograph of. Although both photographs and recurrent natural signs are "pictures" in the sense that certain significant isomorphisms determine their representational values, the "ofness" or "aboutness" of a photograph of an *individual* is not the same kind of "ofness" or "aboutness" as that of a recurrent natural sign of an individual. (Later I will argue that the ofness or aboutness of intentional representations of individuals is like that of recurrent natural signs in this way.)[8]

Now consider the retinal image produced by looking at a photograph. It too may contain the local information that Johnny was once flying a kite in front of a camera. But the semantic mapping functions that mediate between the image and the kite-flying affair are not the same as those for a direct retinal image of Johnny flying a kite. The sign route begins with a sign of the spatial layout of the immediate environment, but this layout is a natural sign only of the presence of a photograph, certain properties of which are signs of a spatial layout once in front of a camera, which in turn signify certain properties of Johnny, which signify Johnny—but not, of course, Johnny as being at the same place and time that the photograph is.

But, of course, the fact that all of this natural information is bombarding the retina does not mean that the organism behind that retina is capable of interpreting or using any of this natural information. Clearly, seeing what's in a photograph requires interpreting differently than seeing directly. Also, a hugely important difference between the photograph and a direct retinal image of Johnny is that the latter, but not the former, contains information about the spatial and temporal relation of the observer to the observed. Animals don't see anything in photographs, or if they do, they take what they see to be present. Their seeing capacities are designed only for use in immediate practical activities. They have no ability to use information about just what once

8. That is, looking ahead, what an intentional representation represents is not what caused it, or what would have caused it, or the thing whose properties it would have signified had the biological mechanisms producing it been operating properly. This is a crucial departure, as I see it, from Cummins's (1996) view, from Evans's (1982) view, and also, I think, from Dretske's view.

existed sometime and somewhere or other. Most animals don't see anything even in mirrors, for although mirrors can be used to guide immediate practical activity, special semantic mapping functions must be utilized for this, and most animals are not that flexible. The crucial functions by which relations of things seen to the seer must be interpreted are shifted when a mirror is used. At first interested, a kitten soon begins to interpret the image in the mirror just as a hole in information space. It no more sees anything there than you see the reflections of the people inside the train in the train window when you are concentrating on the scenery outside.

Equally important, so I reemphasize it here, interpreting a sign of a sign of a sign need not involve recognizing all, indeed any, of the signs along the route as such. Each sign along the route shifts the semantic mapping function. But each sign is a sign of each of its more distal signifieds in accordance with a resultant direct semantic mapping function as well, and this direct function may be the only relevant one in simple cases. The rabbit need not perceive, need not harbor an intentional representation of, what is happening on its retina, or what is happening to light impinging on external objects, or even what the color, shape, texture, motion, and so forth are that locally signify a fox, in order to recognize fox and take cover. It is possible that its nervous system should produce an inner intentional representation of fox directly from retinal stimulations. Similarly, it is in principle possible that a creature might recognize Johnny directly from retinal stimulations without going through the process of first recognizing the shape of his face. But here I have gotten ahead of my story again, for it is only intentional representations that can signify distal affairs without at the same time signifying all the more proximal ones in between, and intentional representations are the subject of the next two chapters. There I will argue that not just the origins but the uses of intentional representations are involved in determining their semantic values. If they are not *used* for representing any intermediate signs on the route from *A* to *B* they do not represent any of these intermediate signs intentionally.

I have said that embedding in natural signs is like embedding in natural languages. Two kinds of embedding occur in natural language. I will discuss only the simpler kind in this chapter. In these simpler cases, linguistic signs represent things not directly, but by intentionally representing properties that are natural signs of these things. More accurately, the linguistic signs accomplish this granted that they are

functioning properly, that is, functioning in the way that has accounted for their proliferation in a language community (chapter 2). To function properly, the properties they represent must actually be natural signs of something. In the more complex kind of embedding, to be the subject of chapter 7, linguistic signs intentionally represent other representations *as representations*, and the represented representations, if they are intentional representations, may or may not be signs of anything real. Linguistic signs of this latter kind are what produce the phenomenon of intensionality (with an "s").

Simpler cases of embedding in language involve what I will call "defining descriptions" of various kinds. Descriptions such as "the sixteenth president of the United States" are well known to philosophers under the name "definite descriptions." But there are many other defining descriptions that operate in much the same manner that are not naturally expressed using the definite article. Examples are "taller than Sally," "avocado colored," "a strong Greek accent," "moves like a karate blackbelt," "as close to Boston as New York." A characteristic of defining descriptions is that something like Donnellan's distinction between "referential" and "attributive" uses applies to them all (Donnellan 1966). For our purposes here, I am interpreting Donnellan's distinction as follows.[9] Recall the distinction between a public linguistic form's own function or purpose and a particular speaker's purpose in using it (chapter 2). In using a defining description, sometimes the purpose of the speaker will be accomplished whether or not the hearer understands or knows, independently, to what the description applies. It will be accomplished, that is, whether or not the hearer knows what the property mentioned in the description is a local sign of. Other times the speaker's purpose will not be accomplished unless the hearer does independently know to what the description applies.[10] Thus, although to be *taller than Sally* is to be taller than a definite height, a hearer may come to believe that Jane is taller than Sally without coming to believe, for a definite height, that Jane is taller than that. Donnellan's distinction concerns whether this matters or not, given a particular speaker's purpose. Similarly, a hearer may or may not know or understand what color avocados are, or what a strong Greek accent is like, or how karate blackbelts move, and this may or may not matter to a speaker's

9. I do not claim that Donnellan himself had just this in mind.

10. The notion "knowing independently what such and such is" is not an easy notion, however. For a full discussion of it see Millikan (2000), chapters 13 ff. See also Böer and Lycan (1986).

purpose in using a sentence that makes reference to these properties. Knowing to what a description applies is knowing, for any complete affair in which those descriptive properties are exemplified, within the relevant domain, what further affair is instantiated there. Compare: John is in the place and at the time that the peculiar contours of his face are instantiated.

It is worth adding to this way of articulating Donnellan's distinction a further distinction as well. There are times when a speaker's purpose will be accomplished only if the hearer both understands, independently, to what the description applies and also knows and keeps in mind the particular properties mentioned in the description. Then the speaker's purpose will fail if the hearer goes straight through to the furthest embedded affair represented and does not hold in mind the properties mentioned. Here is an example where understanding what properties are meant and knowing what instantiates them are both important: "I see that the smallest child has taken the largest piece of cake again!" (Millikan 1984, chapter 11).

Just as with embedding in natural signs, an interpreter of a defining description in natural language may recognize what properties are signified without recognizing of what they are the properties. Or an interpreter may recognize both what properties are signified and what these properties in turn signify. Or an interpreter may recognize the more distal affair signified without recognizing the properties that signify it, as when a small child recognizes that "the president of the Paleontological Society" is her mother without knowing what a president is, what a society is, or what paleontology is. The child may use a direct semantic mapping function, and this may be enough to fulfill a speaker's purpose in speaking, even though as a public language sign "the president of the Paleontological Society" is indirect.

Defining descriptions in natural language generally signify occurrences of properties, occurrences that are local signs rather than context-free signs of referents. Occasionally the local domain of the property occurrence intended is our entire universe. This does not make the property occurrence into a context-free natural sign. Context-free signs signify what they do by virtue only of natural or logical necessity, never merely by virtue of contingent truths of our universe. Often the properties signified by defining descriptions might have any of numerous significances depending on the domains in which they occur. The shape and characteristic motion of a cat in one local domain is a sign of Felix; in another it is a sign of Zeke. Similarly, being an

exemplification of *catkind* in the one domain signifies Felix, in the other Zeke. A hearer of the simple description "the cat" must gather from the context which local sign domain is the one within which the speaker refers to the occurrence of catness. Tracking local sign domains through language bears many similarities to tracking local sign domains in nature. ("Please bring the book on the table." "*Which* book on *which* table do you mean?") A lot of attention has been paid recently to the phenomenon of joint attention between participants in a conversation. To understand joint attention completely would be to understand how speakers and hearers manage to follow the focus of one another's minds during ordinary conversation so that the hearer's attention is successfully drawn to the local domains on which the speaker is focusing. This theme will be taken up again in chapters 10 and 11.

5 Teleosemantic Theories

Theories of the sort that I propose for intentional signs are often called "teleological theories of content." Accordingly, naturalist theories of the content of mental representation are often divided into, say, picture theories, causal or covariation theories, information theories, functionalist or causal role theories, and teleological theories, as though these divisions all fell on the same plane. That is a fairly serious mistake, for what teleological theories have in common is not any view about the nature of representational content. "Teleosemantics," as it is sometimes called, is a theory only of how representations can be false or mistaken, which is a different thing entirely. Intentionality, if understood as the property of "ofness" or "aboutness," is not explained by a teleological theory. Natural signs are signs of things and represent facts about things, but they cannot be false. To explain the possibility of falseness, then, cannot be the same as to explain ofness or aboutness.

The confusion began with, or at least passed through, a common interpretation of Franz Brentano's writings. According to this interpretation (Chisholm 1967),[1] Brentano held that what distinguishes mental phenomena from physical phenomena is that the former possess a property that he called "intentionality." He characterized intentionality in two different ways. He spoke of "object intentionality," "direction on an object," or "reference to a content." This was the capacity of mental phenomena to be *of* things or *about* things, as when a thought is *of* Johnny or *about* the fact that winter will soon arrive. But he also spoke of "intentional inexistence," which referred to the apparent peculiarity of objects of thought that these objects can be thought about, can be in or before the mind, even when they don't exist. One can think of

1. I am aware that Chisholm's interpretation is now considered questionable, but its influence has not abated.

Saint Christopher even though he never existed and one can think it is already snowing when it is not. Clearly, he argued, the relation between a thought that *does* exist and an object or situation that *doesn't* exist is not a physical relation. It is a peculiar "psychical" relation. Inside Brentano's term "intentionality," then, was trapped the theory that to explain how a representation could be of or about something is just the other side of the coin of explaining how it could be empty or false. Also trapped inside this theory, of course, was the idea that when one thinks emptily or falsely there *is* an object that is being represented in more or less the same sense that there *is* an object represented when one thinks truly. There is something called an "intentional object" or an "intentional content" that is present regardless of whether the thought is true or false.[2]

Teleological theories of content are best understood if we insist on dividing Brentano's "intentionality" into its original two aspects, treating them separately. Teleological theories have in common that they deny that there *is* any object at all that is being represented when one thinks emptily (say, when seeming to think about "phlogiston" or "the ether") or that there *is* any state of affairs or occurrence being represented when one thinks falsely. Similarly, there *is* no object, not even an inner one, being seen when one has an hallucination. Teleological theories all deny this for the same reason. They take it that mistaken representations, rather than representing peculiar objects, things called "contents," are merely representations that are *failing* to represent.

False representations are representations, yet they fail to represent. How can that be? It can be in the same way that something can be a can opener but be too dull and hence fail to open cans, or something can be a coffee maker yet fail to make coffee because the right ingredients were not put in or it was not turned on. They are "representations" in the sense that the biological function of the cognitive systems that made them was to make them represent things. Falsehood is thus

2. Brentano was surely mistaken, however, in thinking that bearing a relation to something nonexistent marks only the mental. Any sort of purpose might fail to be fulfilled, and hence might exhibit Brentano's relation, such as the purpose of one's stomach to digest food or the purpose of one's protective eye-blink reflex to keep out the sand. Nor are stomachs and reflexes "of" or "about" anything. The traditional reply, scouted in chapter 1 above, is that natural purposes are "purposes" only in an analogical sense, and hence "fail to be fulfilled" only in an analogical way. This response begs the question. It assumes that natural purposes are not purposes in the full sense exactly *because* they are not mental. It also fails to explain why analogical intentionality is not accompanied by analogical "aboutness."

explained by the fact that purposes often go unfulfilled. "What is represented" by a false representation is indeed "something that does not exist," because a false representation represents nothing at all. By turning intentionality in this way into just ordinary purpose, and by naturalizing the notion of purpose, teleosemantics yields a fully naturalistic resolution of Brentano's paradox about nonexistent objects of thought.

Here is another way to understand Brentano's problem and the point of the teleosemantic move. Compare the verb "to represent" with verbs such as "to see," "to hear," "to smell," and "to perceive." Gilbert Ryle called these latter verbs "achievement words" or "success words" (1949, p. 223). He contrasted them with "task" words or "search" words or "try" words. "Hunt," for example, is a try word, whereas "find" is an achievement word. "Look" seems to be the try verb corresponding to "see," "listen" the try verb corresponding to "hear," and so forth. Now consider the achievement verb "to know." What try word goes with it? There seem to be two: "to wonder whether" and "to believe." As you can hunt without finding, you can wonder whether without knowing whether, and you can also believe that without knowing that. Believing is aiming to know or seeming to oneself to know, but it is not necessarily succeeding in knowing. In exactly the same way, besides looking, one sometimes merely seems to see. Shouldn't there be a second try verb then to go with "see"? Where is the aiming verb that corresponds to "see" as "believe" corresponds to "know"? Or the aiming verb that corresponds to "hear" in this way? Where is the aiming verb that corresponds to "perceive"?

The difficulty is that we have no such separate aiming verbs. Instead, the same verbs are used over again. When delirious, you say that you "see pink elephants" even though you are surely not succeeding in seeing pink elephants, because there are no elephants there to see. You say that you "hear voices" even though you are not succeeding in hearing any voices because there are no voices there to hear. The verbs of perception are all equivocal in this way, and equivocity in language leads immediately to fallacies of equivocation in thought. For example, because in its achievement sense you can't "see" what isn't there, it is thought that when you see pink elephants there must be something there to see—not pink elephants, of course, but certainly *something*. Seeming to see is confused with actually seeing—not seeing the same thing, of course, so seeing what? Something "merely mental," an "inexistent object," for example, a "visual image."

Similarly, the verb "to represent" is equivocal. Used as an achievement verb, to represent something requires that there be something there to represent. Thus Brentano claims that representing something in thought requires that there exist something for the thought to represent. But "represent" is also used as a try word. You can represent a golden mountain even if there is no golden mountain to represent. This is what confused Brentano. The verb "to represent" collapses the distinction between succeeding and merely trying or seeming to succeed.

The teleosemanticist, in order to speak most clearly, should refuse to equivocate in this way. He or she should simply deny that you can see or think of or represent what doesn't exist, refusing to use the verbs "to see," "to think of," "to represent," and so forth except as success words. This avoids the confusion that results in the reification of special "intentional objects." Strictly speaking, you can't represent something that doesn't exist. In equivocating, what the philosophical tradition has previously done is to confuse actually representing with merely being in a mental state turned out by cognitive equipment designed to produce representations. Such a mental state may, of course, cause the mind to churn as though it were representing, but that does not produce actual representing.

Teleosemantics neatly disposes of Brentano's reified intentional contents. But by itself it says nothing at all about Brentano's "object intentionality." What teleological theories do not have in common is an agreed-on description of what representing—what "ofness" or "aboutness"—is. They are not agreed on what an organism that is representing things correctly, actually representing things, is doing; hence they do not agree on what it is that an organism that is misrepresenting is failing to do. To the shell that is "teleosemantics" one must add a description of what *actual* representing is like. When the bare teleosemantic theory has been spent, the central task for a theory of intentional *representation* has not yet begun. Teleosemantic theories are piggyback theories. They must ride on more basic theories of representation, perhaps causal theories, or picture theories, or informational theories, or some combination of these.

Failure to grasp this last point has led many to take a dismissive attitude toward teleosemantic theories. How, they ask, could the question whether my current thought is the thought that cats meow, or instead the thought that elephants are big, be a matter that is settled only by evolutionary history, or even by my past learning history? But a teleological theory, just as such, makes no attempt to explain what makes

your thought be a representation that cats meow or that elephants are big. Put it this way. You present your favorite theory of what a true representation is, for example, of what makes something be a true representation that cats meow. Then the teleologist subtracts nothing but adds one thing. The teleologist adds that for your true representation to be an *intentional* representation, it must be a function or purpose of the system that produced it to make representations. Otherwise, though it may indeed be a representation, it is not an intentional representation. Then the teleosemanticist proceeds to explain what a false representation is *given your view*. That is all teleosemantics amounts to.[3]

But there is a catch (of course). The following requirement will have to be placed on the theory of representation you present. Representing things will have to be something that it might sometimes benefit an organism to do. Otherwise it will be a mystery why any organism would contain systems designed to make representations, and the teleosemantic move will not be possible. Surely such a requirement is not unreasonable. But I will argue that in the end it turns out to be surprisingly restrictive. Indeed, in the end, it may not be possible to formulate a coherent teleological theory *exactly* in the way just described. But I will argue that we can come very close.

A common way of glossing teleosemantic theories of representation has been to say that they claim that "the function" of an intentional representation is "to represent" or "to indicate" something. But that way of expressing the teleosemantic idea invites serious confusion. For in the sense of function intended, the functions of things are *effects* that these things have, namely, effects they have been selected for causing, or that their producers have been selected for producing things that will cause, and so forth. Consider, for example, using an informational theory of content as the base on which to build a teleological theory of intentional representation. Natural information is carried by natural signs. To represent, in the base sense, will then be to be a natural sign. But it can't be the function of an intentional representation to *be* a

3. For example, the teleologist, *merely as teleologist*, has no immediate problem with the idea that Davidson's Swampman is truly representing everything that Davidson would have been representing truly had Davidson lived on in Swampman's place. It is just that he is not intentionally representing these things. His true beliefs are representations, but not intentional representations. And where Davidson had *false* beliefs, Swampman is not representing anything at all, either in the achievement or in the trying sense. He is not failing to achieve his purposes on any level of purpose (chapter 1). (There is a wrinkle, of course. Swampman's beliefs about his past and about who his relatives are seem problematic on any view of representation.)

natural sign. It can't cause itself to be a natural sign or cause itself to indicate something. For example, it couldn't cause itself to have been caused by something. Taking the classic example of the firing of a fly detector in a frog's eye, it couldn't be a function of that firing to have been caused by a fly. Functions don't work backward. What, then, is the informational teleologist to say? Several different kinds of things might possibly be said, and to keep these kinds separate makes all the difference when thinking about teleological theories in general.

The teleologist might claim that the function of the apparatus that *produces* intentional representations is to produce basic representations, whatever one takes basic representations to be. For example, the teleologist who takes natural signs as basic representations might say that the function of the systems that produce intentional representations is to produce natural signs. Intentional representations are purposefully produced natural signs. That is one possibility.

A second possibility might be to claim that the function of an intentional representation is to have some sort of effect that retroactively makes it into a basic representation. A basic representation is then defined by its effects, as a terminal illness is defined by its effects. Perhaps it is designed to have the effect of correctly representing something *to something else*, an organism or an interpreting apparatus within an organism. The teleologist will then have the job of explaining, without going round in circles, what it is for something to have the effect on something else that will constitute its being interpreted. What kind of reaction is interpretation? Possibly a functionalist theory could be given here. What makes a basic representation be a representation is how it interacts with other representations and how these together produce the organism's behavior.[4] Possibilities one and two might also be combined.

A third possibility is less obvious. Let me give some background first. When a trait has been selected for by natural selection, or by learning, it has been selected for some effect that it has. (If it was selected not owing to some effect it had, but merely because it correlated with the occurrence of some helpful happening that was not its effect, then it has been "selected" but it has not been "selected for" [Sober 1984].) That is why a thing's functions (in the sense meant here), or its pur-

4. This would be a good choice for the teleologist who thinks that Swampman has thoughts. But such a functionalist theory would have to ignore the causes of representations, and look merely at their effects.

poses, are effects that it has, not its causes. On the other hand, for any trait to have had an effect systematic enough to have caused its selection, there will generally be some explanation of how it caused that effect that is also systematic. That is, it will not be that every time it caused this effect, it did so by an entirely different mechanism. This need not be true in the extreme. Just as certain traits of an animal may have been selected for because sometimes they served one helpful function and sometimes another, as one's hair both protects one's head from abrasion and also keeps it warm, there could be a trait that always served the same function during its history but did so by the operation of several alternative mechanisms. But it would be impossibly unlikely that there would be more than a few mechanisms by which that function was effected. I call predominant mechanisms of this kind "normal mechanisms" for performance of their functions by these traits.[5]

A normal mechanism for performance of a trait's function will pretty invariably involve the presence of other things that act in cooperation with it, acting on it or being acted on by it, and it will involve the presence of various supporting conditions. In the absence of these supporting things or conditions, probably it will not be able to perform these functions. But it is always possible that a trait should cause some proper effect, an effect it was selected for, by accident in some cases. It is possible, for example, that my reflex eye-blink, though caused only by a passing shadow, might nonetheless succeed accidentally in keeping a piece of sand out of my eye.

Here, then, is the third possibility for the teleologist. The teleologist might claim that when the systems that produce and/or use intentional representations perform the tasks they were designed to perform and perform these tasks by means of their normal mechanisms—let us just say "in a normal way"—then the intentional representations are basic representations—whatever "basic" representations are taken to be. Notice that it need not follow from such a theory that it was a function or purpose of the system that produces intentional representations to make things that are basic representations. That would merely be how they normally manage to serve their functions.

5. In earlier writings I have I referred to them as "normal explanations" or "Normal explanations" for performance of a trait's functions. This caused some confusion, since many think of an explanation as being a set of propositions rather than what these propositions are about. For a thorough treatment of the notion of a Normal explanation, see Millikan (1984), chapters 1 and 2.

The reader will be right in suspecting that I have introduced this third option so as to embrace it, and right in suspecting that I will use local natural signs for my base representations. But intentional signs will be defined independently by reference to the functions of their producers. It will follow from this that they are also designed to have certain effects on their consumers, the organisms or parts of organisms that use them.

6 Intentionality

The teleologist needs a base theory of the representing relation on which to build his description of intentional representation. Fred Dretske (1986, 1988, 1995) placed his theory of natural signs and the natural information they carry at the base of his teleological theory (see chapter 3 above). Some items, Dretske claimed, have the "function" of carrying natural information, and when they do this they come to represent intentionally. They become intentional representations, representations that can be false (Dretske 1986, 1991). Dretske's "functions" are what I have been calling "purposes," at least close enough for my purposes.[1] If we take it that a plain representation, one that is not intentional in Brentano's sense, is just a bearer of natural information, then Dretske's theory of intentional representation is a neat example of a completed teleosemantic theory of the sort described in chapter 5.

Now I have claimed that there exists practically no natural information of the kind Dretske describes. If intentional representations were to be made by organisms out of natural information, it would have to be out of a more user-friendly kind of natural information. Let me introduce my suggestions on intentional representations by asking whether Dretske's program can be carried out using local natural information as a base, and if not, why not. In particular, let us ask whether locally recurrent information might be used in this way. Might intentional representations be just things designed to carry local natural information? Might they be things produced by systems whose purpose is to produce locally recurrent natural signs?

The proposal on the table here puts no definite requirements on the probability that the systems designed to produce natural signs actually succeed in doing so. They need only to succeed often enough to offset

1. At least they appear to be in Dretske (1995).

their own production costs (energy and resources used) plus any negative effects resulting directly from failures. The rabbit, for example, can afford to be mistaken that a fox is near many times, so long as it also takes a fox to be near whenever one actually is. Having an intentional natural sign–producer that is extremely fallible may be much better than having none at all. It is necessary only that the costs of being wrong are lower than the costs of totally unrelieved ignorance. Similarly, an organism's capacity intentionally to represent certain kinds of states of affairs does not depend on its ability reliably to discriminate those states of affairs. Putting things bluntly, no vestiges of verificationism remain under the proposed analysis.

But the teleologist who uses natural information at the base of his theory must explain at least why producing natural signs might sometimes be useful to the organism that produces them. How might an organism benefit from the production of natural signs? Why might a system within it have been selected for that job? It will pay to be very careful here. We need to distinguish a useful effect that also happens to result in the production of natural signs as a side effect from the production of the signs being itself a useful effect. For example, the production by the body of calluses where the wear is has the useful effect of protecting the skin from further damage, and where the calluses are is a natural sign of where the wear has been. But the fact that the calluses are a natural sign of where the wear has been is not, of itself, of any use to the body. The disposition to produce calluses was not selected for its effect of indicating where the wear has been. If I dig ditches to channel the water across my property when it rains, then if my purpose is fulfilled, the ditches will be natural signs of where the water will flow when it rains. But my purpose was not to produce natural signs of where the water will flow when it rains. Perhaps geese have been designed by natural selection to react to frosty nights by flying south, the useful effect being that they fly south just before winter arrives. If this design works right, a side effect will be that frosty nights are a natural sign that geese will soon be flying south, and geese flying south will be a natural sign that winter will soon arrive. Both of these natural signs are the result of the operation of natural selection on geese, but the production of neither sign was selected for. Neither the calluses nor the ditches nor the frosty nights nor the geese flying south will be intentional signs, then, under the proposed analysis.

On the other hand, consider the distinctive clucking sound that a mother hen makes when she finds food. The clucking is a locally recur-

ring natural sign that the hen has found food. Further, the mother hen's chicks respond to her calls by running to her and thus finding the food. Indeed, her disposition to call has been selected for having this effect on her chicks. The call is not merely a natural sign of food that happens to be used by the chicks. It has been purposefully produced in order to serve the chicks as a sign of food. On the proposed theory, her call is an intentional sign.

The lesson seems to be that if an intentional sign is the same as a purposefully produced natural sign, it must be designed to function as a sign *for some kind of interpreter*. The teleologist who claims that intentional representations are produced by systems whose function is to produce natural signs must also claim that intentional representations are designed to have the effect of correctly representing something to some other organism or interpreting system (combining the first and second possibilities for a teleological theory suggested in chapter 5). This accords with the governing idea, suggested in chapter 3, that a useful notion of natural sign should define a category useful to the natural epistemologist. Signs are things apt for use by sign-users. Intentional signs are signs purposefully produced for use by sign-users. It follows that besides a theory of what natural signs are, we also need a theory of what sign use is.

A refinement is now needed for this proposed description of intentional signs. Suppose that I purposefully shoo away flies with a flick of my hand. My hand flicks cause the flies to depart because they serve as a natural sign of danger to the flies, and I intend this. Are my flicks then intentional signs? We might solve this problem in an easy way by saying that they may be intentional signs, but they are not cooperative intentional signs, and that when people talk of intentional signs they usually have cooperative intentional signs in mind. Cooperative intentional signs are produced by systems designed to make natural signs for use by cooperating interpreting systems. That is, the sign-maker system and the sign-using system must have evolved or been designed to function symbiotically. Cooperative intentional sign-makers must be designed to cooperate with interpreting systems that have been designed, in turn, to cooperate with them. A cooperative intentional sign will always stand midway between two systems that have been designed to cooperate with one another. Perhaps one system is a clucking hen and the other a listening chick, or perhaps one system comprises certain systems in the brain that make signs for other systems in the brain to use. However it goes, that the sign should be a natural sign

must be of concern to *both* systems. What helps to proliferate the chicks also helps to proliferate the hen, and what helps to proliferate one half of the brain generally helps to proliferate the other. By "intentional sign" I will now always mean cooperative intentional sign unless I indicate otherwise.[2]

But is it true that the intentional representations in one's head are, in general, recurrent natural signs of the affairs they are about? A locally recurrent natural sign must fall within some natural domain within which the signs coincide with their signifieds all for the same reason (chapter 3). But the perceptual systems of an animal may rely on numerous alternative and quite independent natural cues in constructing even elementary perceptual representations. For example, ocular disparity (convergence), partial occlusion of one object by another, tautness of the muscles that focus the eyes (accommodation), and atmospheric haze are all used in depth perception, jointly or alternatively. It is quite possible, of course, that any one of these natural signs of depth, taken by itself, bears a strong enough correlation to depth to be of use by itself to an organism, and hence that the set of inner representations of depth derived by each of these various methods is, in turn, a separate set of locally recurring natural signs. But suppose we generalize the principle involved. Consider how many quite independent ways you may have of recognizing one of your parents, or your spouse, or your children as manifested to you in perception. You know them by the look of their faces or of various of their other body parts from a hundred angles. You know them by their postures, their mannerisms, their clothes and other belongings, their voices, their handwriting, by their characteristic linguistic expressions, by traces of their habits (who leaves the lid off the peanut butter in the afternoon?), and so forth. To each of these various kinds of natural signs, you make the same cognitive response, forming another representation token representing some fact about this same person. Numerous different kinds of natural signs of the same thing have all been converted into a common intentional currency, but each by a different method.[3]

2. Kim Sterelney (1995, 2001) overlooks the requirement of cooperation for producer and consumer laid down for intentional icons in Millikan (1984, 1993).

3. Sterelney (1995) wishes to *require* this kind of flexible tracking of information for intentional representations. One can use one's terms as one likes, so long as the ideas are clear and useful.

Further, it seems likely that advanced perceptual systems, like connectionist perceptrons, do not rely mainly on single recurrent high-quality natural signs. Often they are moved by the cooccurrence of a diversity of very weak symptoms of a thing, to form a representation of its presence. Further, prior knowledge of present state conditions and of natural laws or uniformities is often involved when a human takes one thing as a sign of another. Recall the beeping briefcase from chapter 4. Intentional sign production, when taken entirely generally, seems quite different from recurrent natural sign production as defined in chapter 3. In the case of humans at least, a multiplicity of quite different mechanisms, operating in accordance with a wide variety of different principles, is what leads to coincidence of intentional signs with what they intentionally signify.

But perhaps we don't have to be so straitlaced. The second requirement for locally recurrent natural signs was that there must be a reason why the mechanism producing coincidence between sign and signified continues to exist or to repeat itself throughout the domain of the sign. That the correlation continues must not be an accident. Take the limits of the sign domains for perceptions and beliefs to fall inside the organisms that produce them. Surely it might be said that the correlations between intentional signs and what they intentionally signify extends throughout that domain "for a single univocal reason."[4] The convergence is no accident. There are systems designed by selection processes, processes that take place on a variety of levels, that see to this correlation.[5] Let us relax the criteria for locally recurrent natural signs, then, so as to include these inner representations.

So far I have been able to defend a neo-Dretskian program that would explain intentional representations as purposefully produced

4. This is how Nick Shea put it to me. He nudged me into taking this more flexible position on locally recurrent signs, whereas I had been vacillating. The decision is, in a way, merely verbal. But some ways of using words help one to think more clearly than other ways do. There is one comment I feel obliged to add, however. In Millikan (2000) I argued that there is no proof that inner representations that are understood as representations of the same by the cognitive systems need to be carried by similar vehicles. But the various tokens of a recurrent natural sign do have to be similar to one another. But even so, inner representations would still be local natural signs in the sophisticated sense that the mitten in the path in chapter 3 was a natural sign. The cognitive systems do seem to have systematic ways of determining what they are local signs of in this more sophisticated sense. (How one knows what one's inner signs are about is also discussed in Millikan 2000, chapter 13.)

5. For conception, these systems are discussed in Millikan (2000), chapter 7.

local natural signs. But there is one problem for such a theory that I think cannot be solved.

If the intentional representation–producers have as their purpose to produce natural signs it must be because they will be aided by systems that use these signs to guide them in some kind of activity that is productive for the sign-making organism. Intentional signs must have ways of earning their keep. How will these signs be used? What will constitute their being correctly "interpreted"? In the obvious case, they will be used to guide their consumers in activities that succeed only by taking account of or conforming to the affairs that are signified by the natural signs. That is, these activities, whatever they are, will succeed by normal mechanisms (chapter 5) only because the effect of the signs is to adapt these activities to the existence of the signified affairs. Variations in the world must correspond to variations in the sign that produce adaptive variations in the activities of the sign's interpreters or consumers. Given the way the consumers' activities are designed to vary with the sign, then, there will be some determinate semantic mapping function[6] by which the sign must correspond to the world if its consumers are to perform their functions normally.

The difficulty is that that is *all* the sign's consumers need in order to perform normally. The consumers can do their jobs perfectly so long as the signs they consume correspond to world affairs by the required mapping function. It doesn't matter to them how the signs they use were produced, so long as they map onto world affairs the right way. Further, if we focus clearly on the function of the sign-producers, carefully distinguishing their function from the normal mechanisms by which they fulfill this function, we see that their function is only to produce for their consumers what the consumers need. Their function is only to produce representations that correspond to world affairs by a certain mapping function. Their purpose or function is not to achieve this in any particular way. Thus they might sometimes fulfill their function by accident, as the eye-blink might keep sand out of the eye by accident when triggered by a sudden shadow. But it is the way the signs were produced that determines whether or not they are natural signs. So it is not a purpose of the intentional sign–producers to produce natural signs. When they perform their functions by their normal mechanisms they produce natural signs. When they perform

6. I would say "semantic rule" if it were not that the notion of a "rule" tends to have prescriptive overtones that I wish, by all means, to avoid.

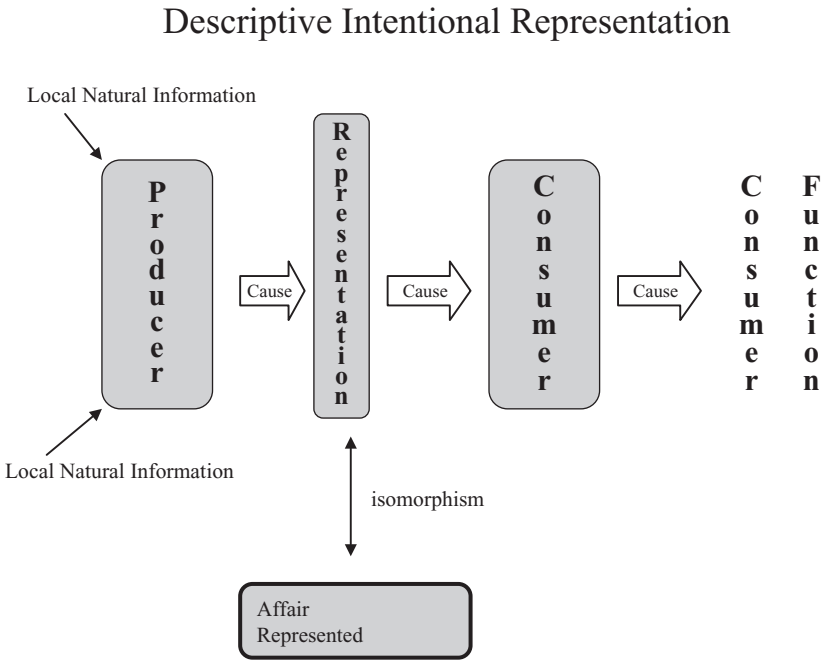

Figure 6.1

their functions by accident they produce only true intentional signs. Putting things intuitively, these representations are true but don't constitute knowledge.[7] But the purposes of both the producer and the consumer may be fulfilled anyway.

What then is the upshot for a teleological theory of intentional representation? The need for dialectic is over, I think. Let me just lay down my position.

Most theories of representation deal with descriptive representations only—with representations that purport to represent facts. But directive representations are certainly equally important, as well as a third kind that I label "pushmi-pullyus"[8] because they both describe and direct. Figures 6.1, 6.2, and 6.3 diagram these three kinds of intentional representations. In every case, when the production and use of these representations proceeds by normal mechanisms, they are local natural

7. For defense of the position on knowledge suggested by this cryptic remark, see Millikan (1993), chapter 12.

8. After Hugh Lofting's charming double-facing creature by that name. For a general essay on pushmi-pullyu representations, see Millikan (1996).

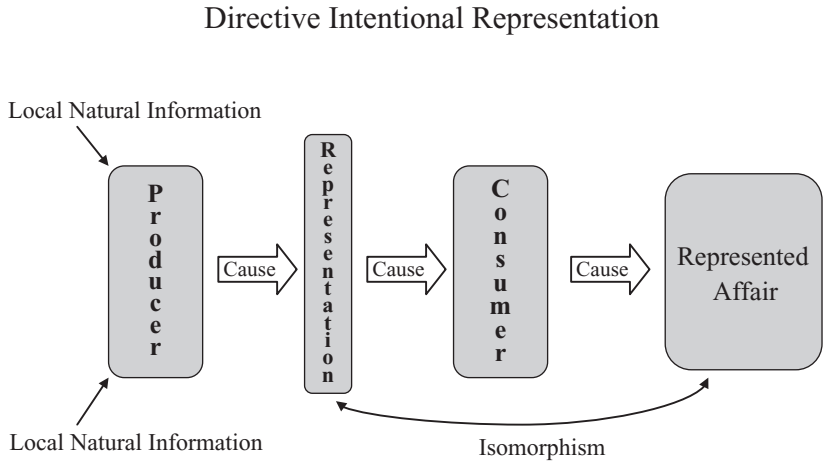

Figure 6.2

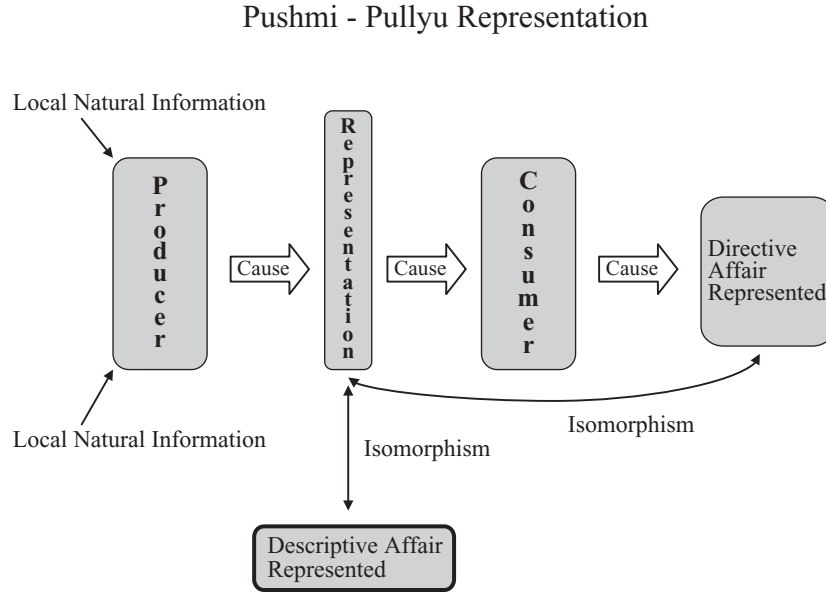

Figure 6.3

signs. (That the directive representations are natural signs is evident if you keep in mind that natural signs can be signs of future affairs as well as past affairs, and that local signs can be flanked by things indistinguishable that are not signs.) The theory might thus be described as a sort of informational theory or natural sign theory exemplifying the third possibility for a teleological theory of intentional representation described in chapter 5. Or it might be taken as a sort of functionalist theory, combining possibilities one and two from chapter 5. In the peculiar case of human belief and desire, part of the functional role concerns use of these representations in inference, prior to their eventual effects on action. That is part of how they are "consumed" or "interpreted." The theory might also be taken as a sort of picture theory, because intentional representations, like recurrent natural signs, necessarily come in systems involving a domain of possible signs running isomorphic to a domain of possible representeds (chapter 4).

In the diagrams, causation runs from left to right, and the single lines indicate mapping relations or isomorphisms. Isomorphisms are logical rather than causal relations, of course. Where only single lines are drawn there may be no significant causal relations. It will be seen, then, that the theory of descriptive representations is not a causal theory of intentional representation. This is because recurrent natural signs are not always causally related to the affairs they signify (as was explained at the end of chapter 3, using the magnetosome as an example).

In each of the diagrams there is a producer and a consumer. These will have been designed to cooperate with one another. Perhaps each is a separate organism, usually conspecifics. Or perhaps they are two parts or aspects of one organism. What the consumer does helps the producer, and what the producer does helps the consumer, and this is no accident but rather the result of some kind of selection or learning that has operated on both together. The presence of each is part of the normal mechanism by which the other fulfills its functions. In each diagram the producer produces a sign that will be true or satisfied only if it maps onto some affair in accordance with a definite mapping function determined by a history of joint successes of producer and consumer (or their ancestors).

For descriptive signs, that the sign maps in this way is a condition that is required for the consumer to perform its tasks, whatever they are, by the mechanisms normal for it. The content of the descriptive sign is not determined by the tasks its consumer performs. It is determined by what the sign needs to correspond to if the consumer is to

perform its tasks in its normal way. The producer's job is merely to make a sign that corresponds in the right way to a world affair. If it does this in its normal way, by its normal mechanisms, the intentional sign it makes will also be a local natural sign.

For directive signs, that the sign maps in the right way will be a result of the consumer's activity. The consumer's job is to cause the sign to map in this way by producing or causing a corresponding affair. Its job is to obey the producer's orders. But equally, it is the job of the producer to give orders that will benefit both it and the consumer. Only in that way can such a cooperative pair be selected for. Similarly, in the case of a descriptive representation, the consumer's job must be such as to use the representation in a way that will benefit both itself and the producer. Whatever more concrete jobs that consumer has, it must have this effect (or have had this effect) often enough to ensure selection of the cooperative pair.

Thus the cooperation between producer and consumer in production of natural signs can be accomplished in either of two basic ways. First, it might be that the producer is the one primarily responsible for making the sign correspond to the world. Then the sign vicariously guides the consumer in relation to the signified as the consumer performs some task mutually beneficial to itself and its producer. These are descriptive intentional signs. They are designed to stand in for world affairs, typically affairs outside the organism, and to vary according to these world affairs, controlling the animal's internal or external behavior as needed to adjust to these world affairs. Second, the consumer may be the one primarily responsible for making the world correspond to the sign. Then the producer's job is to make the sign be such that when the consumer has produced the signified world affair, the result is mutually beneficial to itself and to the consumer. These are directive intentional signs. Directive signs guide the consumer in the production of world affairs that vary according to how the signs themselves vary. They are blueprints for what is to be constructed or brought about. Inner imperative signs are represented purposes of the organisms that harbor them, as represented purposes were discussed in chapter 1. They represent what it is their purpose to bring about.

But the most interesting kind of representation is the pushmi-pullyu representation. It is also the most primitive. Consider again the hen's food call. It is at once descriptive *and* directive. It is the hen's job to make the call coincide with the time and place of some food, and it is

the chick's job to make the call coincide with the time it approaches that place. The call is false if there is no food; it is unsatisfied (not complied with) if the chicks do not come. Almost all animal signals are of this kind. For example, bee dances tell at once both where the nectar is and where the watching bees are to go. The vervet monkey's leopard, snake, and flying predator calls tell what kind of predator is near and direct the response appropriate to that predator. In part IV I will discuss pushmi-pullyu representations in detail and explore the question of how and why natural creatures should have evolved the capacity to form intentional representations also of more differentiated kinds.

On the theory proposed, intentional representations always come with propositional attitudes attached. It is essential to them that they have some kind of function, that they are designed for a particular kind of use. Frege's notion of sense, which implied that you can first represent a proposition and then add an intentional attitude to it, has done a lot of damage, I believe. There are not and could not be intentional representations that lacked attitude. There are no intentional representations without purposes, and having a purpose guarantees attitude. Activities such as hypothetical thinking, for example, or just thinking of possibilities, are extremely sophisticated activities, and ones that are possible only for a creature that sometimes uses the results in the production of ordinary descriptive and directive representations. It is because thoughts of this kind have the function of sometimes turning into more basic kinds of representations that they can exist at all.

In *Knowledge and the Flow of Information*, Dretske noted as a remarkable fact that our percepts can carry "information about a distant causal antecedent . . . without carrying information about the more proximal members of the causal chain . . . *through which* this information . . . is communicated." The percept "skips over (or 'sees through') the intermediate links in the causal chain in order to represent . . . its more distant causal antecedents" (1981, p. 158). And he worried about how abstract representations are possible, ones that carry only the information that, say, an object is triangular and not also that it is isosceles or equilateral. He introduced a special process called "digitalization" to solve this latter problem. Both problems are solved at a stroke, however, if we carefully distinguish intentional representations from natural signs. Local natural signs of distal affairs carry local information about all of the more proximal affairs on the route from them to those distal affairs (chapter 4). But a natural sign that is also an intentional sign will carry only some of that information intentionally. The

information it carries intentionally is only the information it has been selected for carrying, that is, only the information that is used by its cooperative interpreters. This information may be very abstract, and it may be about very distal affairs. If its consumers are designed to use only the information that something is triangular, then that is all the information that it carries intentionally. If they are designed to use only the information that a predator is near, then it need not carry information about any more proximal affairs intentionally, such as patterns on the retina or properties that are local signs of the predator (chapter 4). Similarly, of course, not every stimulus that an organism discriminates on the way to producing intentional representations is itself intentionally represented. Nor must an organism be able infallibly to discriminate the distal objects, properties, or kinds that it intentionally represents from similar ones. It needs only a fallible capacity to use some natural signs or other of these things under some conditions. Possibly it gets things wrong a large part of the time.

Dretske's example of the magnetosome (chapter 3) illustrates these principles nicely, though, again, it was not introduced by him for this purpose. The magnetosomes of northern hemisphere bacteria discriminate magnetic north, which, when the normal mechanisms associated with their functions are in place, corresponds to geomagnetic north, to the direction of deeper water and to the direction of lesser oxygen. Of what, Dretske asks, is the magnetosome's orientation an intentional sign? Does it signify magnetic north, geomagnetic north, deeper water, or lesser oxygen? It is, of course, a recurring natural sign of all four. But it is an intentional sign only of lesser oxygen. This is because it needs, and needs only, to coincide with lesser oxygen to serve its purpose. Aerate the deeper water with oxygen and the direction the magnetosome points continues to be a natural sign of magnetic north, of geomagnetic north, and of deeper water, but the bacterium dies. Place the bacterium in southern hemisphere waters and the direction it points continues to be a natural sign of both magnetic and geomagnetic north, but the bacterium dies. These other natural signs are not what interests the bacterium, or rather, not what interested natural selection in selecting magnetosomes to build into the bacterium. Corresponding to magnetic north, to geomagnetic north, and to deeper water is merely the normal mechanism by which the magnetosome manages to point to lesser oxygen. Of course, Dretske is right that the magnetosome that directs the bacterium in the wrong direction because someone holds a bar magnet overhead is not broken or malfunction-

ing. In that sense it is functioning perfectly properly (Dretske 1988). But it doesn't follow that it is succeeding in performing all of its functions, any more than a perfectly functional coffeemaker is performing its function when no one has put any coffee in it. Very often things fail to perform their functions, not because they are damaged, but because the conditions they are in are not their normal operating conditions.

Take another kind of example: Although each of the signs that emerges during sensory perception carries natural information about affairs at many levels of distality, each lies on its own level of intentional representation. For example, the edge-detector cells in early vision represent edges, not light intensity gradients across the retina. Their function is appropriately to guide internal acts of identification of contours and shapes, given the presence of certain edges. That the edges are where the detectors say they are is sufficient for them to do their jobs properly. Whether or not it was indeed gradients across the retina that caused the edge-detectors to be properly aligned with edges on a given occasion, and hence whether or not they carry natural information about gradients, is not relevant to the intentional information they carry. Similarly, the representations of contours and shapes that are produced from the edge-detectors are intentional representations neither of edge-detectors (they are not intentional signs of signs) nor of edges. And the representations of interesting objects that are produced from the representations of shapes are representations not of object shapes but of fully identified objects.

Similarly, the English sentence "It is raining" is a recurrent natural sign that the speaker believes it is raining. But it is not an intentional sign that the speaker believes it is raining. Its memetic function, derived compositionally from the combined memetic functions of its significant components, is to produce beliefs that it is raining, not beliefs that speakers believe that it is raining (see chapter 2; also Millikan 1984, 2001a,b). But of course it may be read as a natural sign by a hearer with the appropriate cognitive skills and in this way produce the belief that the speaker believes it is raining. That is another matter entirely.

I have emphasized the sense in which the teleological theory I am proposing is akin to an informational theory of representation and the sense in which it is akin to a functionalist theory. Let me also emphasize its kinship with picturing theories of representation. Like locally recurring natural signs, intentional signs are always members of a domain exemplifying a *system* of possible signs, the entire system running isomorphic to the domain of its signified affairs. The notion of

a sign, I suggested, is best developed as at root an epistemological notion (chapter 3). The embedding of a sign in a system of signs affords an interpreter a capacity to learn new things from signs, or in the case of inner signs, to perceive or think new things with signs. The value of a system of representation lies in its productivity. This depends, in turn, on there being some kind of isomorphism, in the abstract mathematical sense, between the domain of the signs and the domain of their signifieds.

There is no need to place any limit, however, on the complexity of the semantic mapping functions that might map intentional representations onto their representeds. Isomorphisms can be defined by functions that are as bizarre, as gruelike, as you please. A bizarrely coded secret message from a CIA agent may be as much an "icon" or "picture" that maps onto a certain world affair in accordance with a generalized semantic mapping function as any sentence or diagram. Signs must be things apt for use by sign-users, but sign-users can be very idiosyncratic in their habits. For example, if mental representations are systems of brain happenings or brain states that map onto represented world affairs, no a priori limitation is implied on the kinds of brain happenings or states involved or on the complexity of the mappings employed. Every representation is in some kind of code. The complexity of the code is irrelevant. On the other hand, any intentional representations in the brain would of course have to come with inner interpreters that knew how to read them, that is, interpreters that could be guided by them reliably to fulfill further functions. Simple codes relying on only a few principles, if they were also highly productive, tapping into rich natural isomorphisms between the domains of the signs and the signifieds, would seem much the most likely to be preferred by natural selection.

Notice that in describing intentional representation as closely analogous to the representation of natural signs, I have made no reference to inference or rationality or to representations being "calculated over." Opposing a central current in contemporary American philosophy, I claim that rationality is *not* "the mother of intentionality" (Dennett 1987). Rationality will play an important role in the discussion of inner representation before this essay ends, but rationality is a characteristic only of systems using representations in certain ways, and most intentional representations are not used in this way.[9]

9. In this disclaimer lies my main disloyalty to my teacher Wilfrid Sellars. It should be clear that in many other ways I am very deeply indebted to him.

The theory of intentional representations that I have presented in this chapter is explained in a rather different way, and with some differences, too, in terminology, in Millikan (1984), chapter 6. Many critical questions about the theory as presented there have come and gone, but there are two that seem to cling. Let me conclude this chapter by saying a word about them.

In Millikan (1984), chapter 6, I said that the content of a descriptive or "indicative" representation was determined by what its consumer needs the representation to map onto if it is to perform all its functions in accordance with a "most proximate Normal explanation." "Normal explanations" are what I have been calling "normal mechanisms" here. Karen Neander (1995) has objected that among normal conditions that must be mentioned, say, for the male hoverfly's female-detecting systems to carry out all of their functions, are that the female is fertile and that she won't be eaten before she reproduces. But if it were a function of the hoverfly's signal-producers to signal when a fertile female who won't soon be eaten passes by, there would have to be a normal mechanism by which these producers had historically performed this function (chapter 5). There would have to be a systematic way that it managed to produce representations of fertile females who were not about to be eaten. For this sort of simple perceptual device, the explanation would have to be that the hoverfly's perceptual systems were sensitive to some kind of recurrent natural sign of the affair they were to represent. But hoverflies do not encounter recurrent natural signs that the things causing images on their retinas are fertile or that these things won't be eaten. The point is not that the hoverfly has no way of discriminating natural signs of females that are fertile and won't be eaten from other images crossing its retina. The point is that there simply are no such signs crossing its retina. The domain in which the hoverfly operates is one in which the chance that the shadow crossing its retina, assuming that it is of a female hoverfly, is also of a fertile female not about to be eaten is no higher than the chance of any *arbitrary* female hoverfly being fertile and not about to be eaten. By contrast, assuming that it is the shadow of a hoverfly, the chance of the shadow being that of a female is considerably higher than the chance of an arbitrary hoverfly being female. This is because only female hoverflies cruise, whereas male hoverflies hover. Similarly, given *where* the male hovers, the chance of the shadow crossing his retina being that of a hoverfly rather than of some other small particle of matter is also very much raised. He purposefully hovers in a recurrent sign domain where

such a shadow is very likely to be that of a female hoverfly. (An unarticulated image crossing the retina can produce an intentional sign that the thing spotted is both a hoverfly and female, just as the bee dance can represent nectar, hive, and sun without mentioning any of them explicitly.)

A second query concerns the possibility that there might be biological systems whose jobs are to produce false representations. For example, there is some evidence that people who are overconfident are more successful at performing certain tasks than people who evaluate their skills more accurately. First we should notice that any such system responsible for producing overconfidence would not have been selected for the fact that it produced false beliefs, but for the fact that it produced lots of confidence. For example, if a person were always 100 percent accurate at performing a certain task, there would be no profit in his believing falsely that he was less competent than that. Falseness itself could not be the point. More important, however, we should note that many biological systems ride piggyback on systems developed earlier for other purposes. Systems whose jobs were to distort certain beliefs would have to ride on more general systems whose basic jobs were to produce true beliefs. Otherwise there would be no standard mapping rules according to which the distorted beliefs were designed to map world affairs hence according to which they were false.

7 Intensionality

By stressing in the previous chapters that all complete signs signify complete world affairs, I may seem to have implied that complete signs are always translatable by sentences. To see why this is wrong we need to understand how signs are used to represent other signs. The difficulty lies in the fact that the only direct way we have to speak of what nonsentential signs represent is by misleadingly comparing them with sentences.

In chapter 4 I discussed the way defining descriptions work—by intentionally representing natural signs. Intentional signs may also be used to represent other intentional signs. Sometimes this produces the phenomenon philosophers call "intensionality" (with an "s"), as I will explain. There are also some cases in which intentional signs are used to represent natural signs that are not as straightforward as the cases of defining descriptions. These can give rise to a different (perhaps previously unrecognized) form of intensionality.

Wilfrid Sellars claimed that the form of expression "'X' means *Y*"—as in "'Hund' means *dog*" or "'rouge' means *red*" or "'Chicago est grande' means *Chicago is large*"—does not assert a relation between an expression and some other entity, say, a property or a world affair. To understand the meaning of the "means rubric" is to understand what its characteristic purpose or *function* is. Its function is to produce in the hearer a disposition to use the expression "X" in the same way that the hearer already knows to use the expression "Y" in his or her home language. Thus, so long as the two expressions matched up in the "'X' means *Y*" formula play, as Sellars put it, "the same role" in their respective languages, the form "'X' means *Y*" is used correctly. The beauty of Sellars's account is that it works just as nicely for expressions that obviously are nonreferring, as in "'et' means *and*," "'arret!' means *stop!*," "'Hélas!' means *Alas!*"—or "'signifier' means *mean*"—as it does

for "'Hund' means *dog*." On the other hand, taken by itself, this account does nothing to clarify what in the world a "linguistic role" is. It portrays the expression "'X' means *Y*" as a translation rubric but, taken alone, it tells us nothing about what it is for one expression to be a good translation of another.[1]

Similarly, in "On Saying That," Davidson (1968–1969) claims that a sentence such as "Galileo said that the earth moves" is true just in case uttering the words inside the "that" clause of this sentence makes the speaker and Galileo into "samesayers." But again, we are left wondering just what it is for two speakers to say the same thing. It seems reasonable to see Davidson's account as strictly comparable to Sellars's, however. To be "samesayers" is to use expressions that have the same "linguistic role"—whatever that is.

Davidson's analysis is easily applied also to the expression ". . . says to. . . ." In asserting "Mother says to wear your leggings," the speaker claims to be a samesayer with Mother. Mother has said words having the same linguistic role that "wear your leggings" would have if actually used, rather than merely displayed, in the present context to the present hearer. Davidson's analysis is also applicable to forms such as ". . . believes that . . . ," ". . . intends to . . . ," ". . . wishes that . . . ," and so forth, on the assumption that beliefs, wishes, intentions, and so forth are mental representations and that a mental representation can, in some sense, "play the same role" as a linguistic representation. That, I believe, was how Sellars saw the matter. He could do so because "playing the same role" was, for him, very much a matter of more or less (1963, chapter 6). Both men suggest that we represent a representation by holding up another representation that is similar to it in relevant ways. I will argue that which ways are relevant ways is generally determined pragmatically rather than being grammaticalized, and that this results in intensionality.

I have drawn attention to language forms—phonological structures, words, syntactic forms, aspects of prosody, and so forth—as reproduced entities (chapter 2). They are memes, with natural purposes that may differ from or "cross over" the immediate uses to which individual speakers put them. Given the description of intentional representation offered in chapter 6, public language forms (types) are intentional representations just when fulfillment of their functions or purposes by normal mechanisms, which entails the collaboration of

1. See, for example, Sellars (1963), especially chapter 6.

trained cooperative hearers, requires that they coincide with affairs in the world according to established semantic mappings. Their continued reproduction has depended on their having served cooperative purposes of speakers and hearers often enough, and this has depended in turn on correspondence between them and world affairs onto which they have mapped by rules to which both speakers and hearers are adjusted. Broadly, then, my suggestion is that the best sort of translation of a language form will match both its purpose, that is, its linguistic or memetic function, and also its semantic mapping function (note the two different senses of "function"). Generalizing this, the best sort of translation of any intentional sign will match both its purpose and its semantic mapping function.

If this is right, the possibility of saying (showing) *precisely* what an expression in another language "means" by the method of samesaying will depend on the availability in one's home language of an expression having both a matching purpose and, if the expression is an intentional representation, also a matching semantic mapping function.

Consider, for example, the intentional signal the rabbit produces when its predator-detectors fire. The rabbit thumps its hind feet smartly on the ground. The natural purpose of this is to trigger a reflex that causes its relatives (rabbits, not just any bystander) to freeze or take cover. There is no literal translation of that particular pushmi-pullyu danger-thump into English or French. "Danger!", for example, tells of danger to humans, and "Rabbit danger!" does not have the function of sending any rabbits to cover. Rather than directly saying "the rabbit thump means . . . (*so and so*)," to achieve accuracy we must dispense with the "'X' means *Y*" formula and set to *describing* its purpose—as I did just above. Similarly, in trying to explain what the formula "'X' means *Y*" itself means, Sellars didn't use samesaying. He *described* what its function is, what reaction it is used to cause in a hearer.

Turning now from linguistic function to semantic mapping, the linguistic function of the "'X' means *Y*" formula will be performed in the normal cooperative way only if "X" has the same linguistic role in its language that "Y" does in its. Saying "'X' has the same linguistic role as 'Y'" is thus a way of stating truth conditions for that formula. But it does not follow, for example, that the sentence "'Hund' means *dog*" *means the same* as the sentence "The word 'Hund' plays the same role as does the word 'dog.'" For these sentences do not have the same linguistic function. Very small children do not have concepts of words, as

fully understanding the latter sentence would require them to have. For example, they will claim that since there are no ghosts, "ghost" is not a word (Susan Carey, private correspondence). But they acquire the ability to react appropriately to the verbal "'X' means Y" formula very early on. Similarly, the truth condition of a sentence asserting identity, "A is B," requires the word "A" and the word "B" to have the same referent or extension, but understanding the sentence does not require thinking about words (Millikan 2000, chapters 10–12). The job of an indicative sentence is not always to cause a belief with the same truth condition that the sentence has (Millikan 1984, chapter 12; 2001b). The functions of intentional signs can come apart from their satisfaction conditions.

Besides the difficulties that concern matching at once both linguistic function and semantic mapping function, there may be difficulties in matching semantic mapping functions just taken alone. For there can be crucial differences in the articulation of representations that have, nonetheless, identical truth or satisfaction conditions. Semantic mapping functions are not the same things as truth or satisfaction conditions, and they can easily come apart from them. This is because a sign is, essentially, a member of a system of signs, and the same piece or aspect of the world—the same truthmaker—can be represented in sign systems that are not isomorphic to one another. I have said there is no translation of the rabbit's danger-thump into English or French because these languages contain no forms with the same primitive function. Another reason is that the semantic mapping function that aligns rabbit-thumps with the affairs that satisfy them does not articulate these affairs in a way that parallels the semantic mapping function for any English or French sentence.

The rabbit-thump sign has exactly two variables, time representing time and place representing place. Rabbit-thumps are articulated the way stoplights are. Move the time and place of the red light and that moves the time and place to stop. Contrast a rabbit-thump or a currently lighted red light with the sentence, "Stop here now!" The sentence is articulated so as to contrast with "Stop over there now" and "Stop over there in an hour," and with "Sit here now" and "sit over there tomorrow," and so forth. It is also subject to a negation transformation: "Don't stop here now!" The rabbit-thump is not a member of any such system of signs. No transformations of it tell of times other than their own times, or of other places, or of things other than rabbit danger, or of when or where there is no rabbit danger. So to say "Freeze

or take cover here and now!" is not to samesay accurately with the rabbit. Would "Freeze or take cover!" come closer to samesaying with the rabbit (ignoring that its function is not to affect rabbits but humans)? But "Freeze or take cover!" contrasts with just "Freeze!" or just "Take cover!" and also with "Hop or jump!"

More interesting, there is no negation transformation of the rabbit-thump parallel to "Don't freeze or take cover!" Nor is the rabbit-thump transportable into other contexts, as the word "Freeze" is in "Freeze when I blow the whistle!" Still, the satisfaction conditions of the rabbit-thump are expressible in English. The indicative satisfaction conditions are that there is danger to rabbits at the time and near the place of the thump. Those are the thump's truth conditions. Its imperative satisfaction conditions are that the rabbits now freeze or take cover. To express satisfaction conditions you do not have to samesay. But satisfaction conditions do not reveal semantic mapping functions. To tell what the truth conditions are is not to reveal the significant articulation of the representation.

Compare the dance of the honeybee. It represents the current location of nectar relative to the bees' hive and the direction of the sun. But there are no transformations of it that would tell about nectar location relative to objects other than the hive and the sun, or about the location of anything other than nectar. Its references to the nectar, the hive, and the sun are all implicit. Only the reference to the angle between the nectar and the line from the hive to the sun is explicit. No English sentence with the same truth conditions approaches this degree of inarticulateness. I cannot tell you where the nectar is relative to the bees' hive and the sun without explicitly mentioning at least nectar, explicitly mentioning or describing the relevant hive, and explicitly mentioning the sun. An English sentence with the same truth conditions is subject to significant transformations that will tell instead of the relation of nectar to hive and moon, or of nectar to the Eiffel Tower and the moon, or the relation of peanut butter to hive and sun, and so forth.

Nor is the language of the bees "systematic," in the way Fodor and Lepore use that term: "If . . . a language can express the proposition that aRb, then it can express the proposition that bRa" (1992, p. 146). The bees have no way of saying that the sun lies at a certain angle between the hive and the nectar. It is true that bees sometimes represent with their dances the location of water, if water is much needed or, when they are swarming, of suitable places to build a new hive. But there is nothing in the bee dance itself to indicate this shift to a

different semantic mapping function. The shift is recognized in the same sort of way that the domain of a local sign (or of a defining description) is tracked. The bee has to independently "know," as it were, which local domain this sign is in, for the dance itself doesn't say. (The bee has to understand pragmatics as well as semantics!)

A sentence that represents the location of nectar relative to the sun and the bees' hive will also be subject to a negation transformation. It will contrast with a sentence representing that there is no nectar there. But bees have no way of saying where there isn't any nectar, so don't bother looking. Other bee dances tell of nectar at other places, but to say there is nectar one place does not contradict that there is nectar another.

Having distinguished between semantic mapping functions and satisfaction conditions, we can apply this to the philosopher's notion "proposition," which hovers between these two. Volumes could be written about this ambiguity and the trouble it has caused, but let me just cite one example. Fodor and Lepore say the sentences of a language are, in general, "isomorphic" to the "propositions" they express: If a sentence expresses the proposition that John loves Mary then there will be elements corresponding to John, to loves, and to Mary (1992, p. 147). In a footnote they then remark that this is debatable for the sentence "It's raining," which they take to correspond to a proposition about place and time as well as rain. And they say that this seems not to be true of certain idioms as well. Having thus waded in and stubbed their toes, they complain that the waters in this area are muddy. What they are stumbling into here is the ambiguity in the notion of a "proposition." Semantic mapping functions are different from truth or satisfaction conditions, but the notion "proposition" hovers between. Sometimes it comes to rest on one side and sometimes on the other. The notion of "the proposition expressed" presupposes that semantic mapping functions, which are determined by "compositionality" in the broad sense, that is, by "architectural structuring" (chapter 4), are the same as truth or satisfaction conditions—but they are not.

Also consider Evans's "generality constraint" in this connection. Evans held that in order to think of a thing it is necessary to know *what* one is thinking of, that this requires that one have a "concept" or "Idea" of that thing, and that a concept or Idea is a general ability that "makes it possible for a subject to think of an object in a series of indefinitely many thoughts, in each of which he will be thinking of it in the same way" (Evans 1982, p. 104). That is, it is not possible to think of a thing

unless one can represent it as embedded in many alternative kinds of states of affairs. Something like this may be right as a requirement for having a *concept* of something (Millikan 2000, chapters 13 and 14), but it would be a mistake to claim that this was a necessary condition on all inner or on all intentional representation. The bees, for example, surely represent that there is nectar at a certain angle off the line between sun and hive without expressing detachable concepts of the nectar, the sun, and the hive. Moreover, inner pushmi-pullyu representations such as thirst, hunger, and pain represent occurrences of inner states at times and direct appropriate action without further articulation. Thirst, for example, is not an articulate desire for water, but its truth condition is that the body needs water. It directs drinking, but it is not an articulate directive like "You drink water now!"

Thus it is that the public language representations of humans, and surely also human beliefs, desires, and intentions, may differ quite radically from more primitive inner and outer representations having the same satisfaction conditions, representations that are either used by humans below the level of explicit belief, desire, and intention, or used by other species. Descriptive sentences in all human languages have, at a minimum, a subject and a predicate and are sensitive to negation transformation. These properties set them far apart from a host of simpler representations, undoubtedly including many kinds of inner representations that help to govern human behavior on levels lower than that of rational thought. An important part of the story of the evolution of cognition, with which I will be concerned in Part IV, concerns the emergence of various new forms of articulation, as well as new functions, for inner intentional representations.

It is on this dimension that sentences differ also from other sophisticated intentional representations such as maps, charts, graphs, and diagrams. The correctness of an ordinary map or diagram may *entail* the truth of various sentences. But the affairs mapped in common are projected by semantic mapping functions that articulate these affairs against quite different contrasting possibilities. The space of significant transformations surrounding each of these different kinds of representations is entirely different. Each resides in what early Wittgenstein might have called a different "logical space." And, of course, this dimension is also often relevant when comparing different linguistic expressions to one another. In *Philosophical Investigations*, Wittgenstein remarks on the difference between saying or meaning that the broom is in the corner and saying or meaning that the brush and the stick are

in the corner. Or consider the difference between believing you are drinking water and believing you are drinking H_2O.

Having come this far we can more easily understand the phenomenon of intensionality. Sellars and Davidson each pointed out in his own terms that we represent signs by displaying other signs that are like them. That is, we "portray" or offer portraits of signs we wish to talk about by holding up similar signs. Notice that this does not make the portraying sign into an indexical. Its kind stands for another of the same kind just as the place of quail tracks stands for quail in the same place and the size of the tracks stands for quail of the same size. The kind is a merely a "reflexive" element of the sign (chapter 4). But almost always, the portrait is like the original only in certain respects—respects that happen to be relevant to the communicative purposes of the moment.

First, the possibility of using samesaying to express *precisely* the properties of a sign that is not in one's home language depends, as I have said, on the availability in one's home language of an expression having both a matching semantic mapping function and, if the sign to be represented is an intentional sign, also a matching purpose. But as Sellars pointed out, "playing the same role" is very much a matter of degree, and matches that are far less than perfect are often perfectly serviceable. Indeed, often only certain properties of the sign to be portrayed concern us, so that the portraying sign needs to be like the portrayed sign only in very limited respects. Second, there are times when we want to convey information about more than the role of a representation. Sometimes the very words the speaker used make a difference. Then samesaying may require using the very same words, or words that are like them in relevant physical or etymological respects. But unfortunately, when one sign is held up to portray another, which respects of likeness are the relevant ones on the given occasion is not generally, as linguists put it, "grammaticalized." One usually relies on pragmatics to be sure that the hearer understands what aspects of the sign held up are the ones being attributed to the represented sign.

It is customary to describe intensional contexts as contexts in which coreferential terms cannot be substituted for one another without possible change of truth-value. For example, although Bernard J. Ortcutt may be the same man as the man that Ralph has seen in the brown hat, since Ralph may not know this is so, "Ralph believes that Bernard J.

Ortcutt is a spy" may have a different truth-value than "Ralph believes that the man in the brown hat is a spy." This is supposed to show that "Ralph believes that . . ." is an intensional context. But in fact it is often possible to substitute coreferential terms inside ". . . believes that . . ." contexts with no risk of changing truth-values. This is because the purpose of holding up a portrait sentence within a "believes that . . ." context is often merely to portray aspects of the reference of the believer's thought. For example, if I say to you "Ralph thought that our venerable dean was a spy," the fact that Ralph has no idea that either Bernard J. Ortcutt or the man in the brown hat is a dean, let alone our dean, has no effect on the truth-value of my sentence. The phenomenon here is not that one cannot substitute coreferential terms without change of truth-value, but that the grammar alone does not prove that one can. Whether one can or not is a pragmatic matter.

I suggest that an intensional context, described in a more general and illuminating way, is merely a context in which one sign is held up to portray another but where grammar alone does not tell what kind of likeness is intended. Consequently, grammar alone does not tell what other signs might be substituted for the sign held up without altering the import. The phenomenon concerns not merely substitution of coreferential terms but various other substitutions as well, such as substitutions of single words or of phonemes or of words with different etymologies.

Consider "John said the Earth moves." The most typical reading of this sentence would be to take "the Earth moves" as a portrait of the semantic role of the sentence that John uttered. "John said the Earth moves" would then be taken to convey about John the same thing that "Galileo said the Earth moves" truly conveys about Galileo, even though Galileo did not speak English. But "John said the Earth moves," uttered by anyone, or written by someone who does not care about philosopher's conventions with quotation marks, could also be used to portray John's very words, the vehicle as well as the semantic role of his representation. Indeed, it is even possible to use that sentence to portray the vehicle of John's representation only, forget its role. Suppose that I am aware that my hearer does not understand much English. Perhaps John and I are just helping her with English phonemes, and I am just repeating John's sentence for her with clearer pronounciation. Nothing in the form of the sentence "John said the Earth moves" shows which of these interpretations is intended, a

portrait only of phonemes, or a portrait of certain definite meaningful words, or a portrait only of semantic content, or some combination of these.

If I say instead, "John said *that* the Earth moves," perhaps the grammar indicates that the embedded sign I hold up portrays only aspects of the meaning, and not of the vehicle of the sign talked about. I may say it if John spoke German or Italian rather than English. Yet this is not always the case either. Consider "John kept insisting *that* there were many more Greeks than Hellenes." Given that the Greeks *were* the Hellenes, here it must be that the sign vehicle John used is being portrayed too. The situation that would normally be portrayed by the "that" clause in "John insisted that there were many more Greeks than Greeks" is quite different. Result? The phenomenon called "intensionality." There are contexts in which single words having the same referent cannot be exchanged without "changing truth-value," because exchanging them changes what kind of sign would be taken to be portrayed, a sign typed by vehicle or a sign typed by meaning. That the context is one of this sort is a matter not of the grammar alone, but of pragmatics.[2]

Now compare "John said that your mother is brave" with "John said that the governor is brave," as said by Tom to little Willie. Assume that Willie's mother is the governor. Recall that defining descriptions such as "your mother" and "the governor" can be used by speakers to serve any of three functions (chapter 4). It may matter or it may not matter whether the hearer knows of whom the properties mentioned in the description are natural signs in the relevant domain, and it may matter or it may not matter whether the hearer understands and keeps in mind the properties mentioned. Now unless John, the original speaker, had been speaking to Willie himself, presumably he would not have used the words (the vehicle) "your mother." It is pretty clear pragmatically, then, that John's exact words are not being portrayed. So what is being portrayed? Did John refer to the governor as Willie's mother, or not? Likewise with "the governor." Did John refer to Willie's mother as being the governor of something or not? Whether the sign used to portray the intentional sign John used follows the same "route" (chapter 4) to the further affair represented by John's sign or whether

2. I do not take predicates having the same extension to be the same kind of phenomenon as descriptions having the same referent. I assume a realism about properties. My remarks about coreferential descriptions should not be taken to generalize to merely coextensional predicates.

it doesn't is not shown in the grammar of these sentences. But it may well be clear from context what aspect of the sign John used was central to John's purpose in using it, or what aspect is central to what the current speaker wants to convey. For example, although John didn't refer to the governor as being Willie's mother, it may be important to Tom, the current speaker, that he convey to little Willie exactly that it is *Willie's* mother who is admired by John. That may make little Willie proud.

More generally, unless the pragmatic context clearly indicates otherwise, names and defining descriptions that appear within "said that" contexts are not usually taken to portray either the vehicles or the descriptions used by the original speaker. Rather, they portray more distal affairs that are signified by the vehicles or by the properties represented on the route to these distal affairs. The names or descriptions that are used are chosen to convey these distal affairs in whatever way the current listener will most easily understand within the conversational setting. But again, it is clear that there do exist some contexts in which coreferential terms cannot be exchanged without changing truth-value *under at least some pragmatic circumstances*. They cannot be exchanged without changing what kind of sign a sentence held up as a portrait would normally be taken to portray. These are intensional contexts.

Since the memetic purposes of public language forms are never exactly the same as the natural purposes of inner representations such as perceptions and thoughts—obviously they do not do exactly the same jobs—no exact translation of what any inner representation "means" can be given in a public language. It is impossible strictly to samesay with someone's belief. That may be one reason why there is something unnatural about saying that thoughts and perceptions "have meanings," even on the assumption that they are inner representations. You can't *say* what they mean in a straightforward way. Certainly you couldn't *portray* exactly what they mean taking their exact functions into account. But there is another and deeper problem about representing thoughts. There has been considerable controversy recently about whether ordinary language embodies a sort of "folk theory" about thoughts, a theory that implies that thoughts are representations in people's heads that can be fairly accurately portrayed using ordinary sentences. Advocates of this view usually claim that the folk also believe that our behaviors result from causal interactions, corresponding to inferences, among these inner sentencelike

representations.[3] Happily, I don't need to be concerned here with how "the folk" explain behavior. But it does seem clear that the way we talk about our intentional mental states does, rightly or wrongly, assume that these states are enough like sentences to be portrayed by holding up sentences. In the second half of Millikan (2000) I argued at some length that sentences may well be a very misleading model for thoughts. But if we want to understand only how *representations* of thoughts are employed in ordinary language, we can withhold judgment on what thoughts themselves are actually like. We can examine merely how they are portrayed by ordinary language, given our use of expressions such as "believes that . . . ," "fears that . . . ," "intends to . . . ," "hopes to . . . ," and so forth.

Thoughts are portrayed as being much like inner *sayings*. Typically, I believe, they are thought of as analogous to sentences the thinker would use to express these thoughts candidly.[4] And the way ". . . says that . . ." and ". . . says to . . ." are used corresponds very closely to the way "believes that . . . ," "desires that . . . ," and ". . . intends to . . . ," and so forth are used. Although the sentences held up as portraits in these contexts are intended to bear a resemblance to sentences the speaker might candidly have used, just how close a resemblance depends on the pragmatic context. In particular, defining descriptions (chapter 4) captured inside these contexts may be intentionally used to portray descriptions the thinker in question would himself have been disposed to employ, or they may be used merely referentially.

The default assumption is probably that what is portrayed is the reference. The descriptions are chosen mainly for the purpose of getting the hearer to understand the references, given the hearer's background and the current pragmatic context. Descriptions the original thinker might have used to express his thought, should these even be known to the speaker, probably play a role only occasionally, when it happens

3. It is often assumed, further, that the folk take these causal interactions, these inference dispositions, partly to determine the very meanings of these inner representations. But this last description of how the folk think about thought is implausible, given the history of philosophy of mind, for no one made this suggestion prior to the twentieth century.
4. It does not follow, I hasten to add, that ordinary people think they are "individuating" thoughts by reference to a one-to-one correspondence to possible sentences. The thinker might have many different ways at her disposal of expressing the same thought, for example, many ways of expressing thoughts of the same subject by using different defining descriptions. There is no reason to attribute to the layperson a philosopher's conflation of a way of recognizing something with a way of thinking of it. On this confusion, see Millikan (2000), chapters 8 and following.

that something more turns on them. So although contexts in which intentional attitudes are portrayed are usually contexts in which coreferential terms can be exchanged without changing truth value, this is not always the case. Under some pragmatic circumstances, changing the descriptions will change the understood truth-value. The ancients believed that the morning star was seen in the morning, but they did not believe that the evening star was seen in the morning. Contexts in which intentional attitudes are portrayed are intensional contexts.

There are even circumstances in which the sign held up inside the "that" clause following an intentional verb portrays its own *vehicle* as one the thinker would use to express his thought. Using the same example again, it is clear what "John firmly believes that the Greeks were more numerous than the Hellenes" has to mean, at least in a context where speaker and hearer both know that the Greeks *were* the Hellenes. It has to mean that John would have expressed his thought using the very words "the Greeks" and "the Hellenes." Compare also "John was quite sure that a new spigot would cost more than a new faucet!" Given that spigots are the same thing as faucets, it is pragmatically clear that John was a speaker of English, or that he at least had these two words of English. (Even if one believes in Fregean senses, it is completely implausible to suppose that the public words "spigot" and "faucet" express different Fregean senses. Equally, I suggest, for "John was quite sure that Cicero was born before Tully.")

Modal contexts are intensional contexts. I would like to defend the view that modal contexts too are best analyzed as containing representations of representations. Talk of possible worlds is just disguised talk of representations. Though I think this view is correct, the argument for it will have to wait for another occasion.

Descriptions of causes, of natural explanations, and of natural purposes are sometimes cited as creating intensional contexts. But this, I believe, is an error. Consider, for example, the difference between "What caused the fire was that Herbert's youngest child was playing with matches" and "What caused the fire was that Billy was playing with matches." That the child playing with matches belonged to Herbert surely was not relevant to causing the fire. But this merely exemplifies Donnellan's distinction again (chapter 4). A hearer has to gather from context whether the speaker's purpose is, or is in part, to convey information about the properties mentioned in the description or whether these properties are to be understood as relevant only as a sign of something else (the referent) about which information is offered.

Talk about causes is not talk about signs. Similarly, descriptions of natural purposes are subject to Donnellan's distinction.

In this chapter I have discussed phenomena that give rise to linguistic contexts that are "intensional" (with an "s") in the modern classical sense. Notice that there has been no reference to any relatives of Quine's "creatures of darkness," to "intensions" as understood in the tradition of Carnap, nor to any relatives of Fregean "senses." I have been describing only differences in purposes and differences in semantic mapping functions that map either natural or intentional signs onto the extensional affairs that they signify. The "intensional" has been explained in completely extensional terms.[5]

5. Once again I must mention Wilfrid Sellars. He was as concerned as Quine to escape this kind of darkness, and my way of escaping has made use of paths that he originally hewed.

III

Outer Intentional Signs

8 Linguistic Signs Emerge from Natural Signs

Chapter 4 described some similarities between locally recurring natural signs and public language signs. Chapter 6 described similarities between natural signs and intentional signs. Given this background, there are observations that fairly cry out to be made about similarities in the interpretation of natural signs and public language signs. Part III is about these similarities.

Explicitly represented human purposes emerge from more primitive levels of purpose and then submerge again, for their final phases are implemented at these more primitive levels (chapter 1). Similarly, conventional language signs emerge from natural signs and never break entirely free of them. The most primitive intentional signs used between organisms are signals such as the rabbits' danger-thump, the call that the hen makes to her chicks when she has found food, the posture of the angry cat, its lashing tail, the dog's playbow and its wagging tail, the mating dance of the stickleback fish and of the Canadian goose, and so forth. Many communicative body postures and facial expressions in humans are of a similar nature, being endogenous in origin (Elkman 1980). According to ethologists, communicative animal signals appear to have evolved from preparatory movements or "intention movements," made at the onset of some activity. Originally these movements served accidentally as cues to conspecifics. They were just natural signs, either of the animal's ensuing activity or of the activity's normal stimulus, signs, for example, of danger, food, readiness to mate, or of hostile or friendly feelings. If it was advantageous for the animal that its conspecifics, or at least its kin, should be aware of what it was readying to do, or aware of the stimulus that induced this readiness, these intention movements slowly became more genetically enabled, stereotyped, and exaggerated. Similarly, conspecifics interpreting the sign slowly became tuned not just by

learning but by genetic selection to react to the sign appropriately. That conspecifics should be aware of what an animal is ready to do or of the stimulus that provoked this readiness is often useful to both animals. For example, if the bird makes an obvious signal as it prepares to flee from a predator and this causes the whole flock to scatter, it helps the original bird not to be individually noticed by the predator. Certainly it helps the bird for its partner to know when it is ready to mate, and so forth.

The evolution of intentional signs from natural signs also occurs quite rapidly through a similar ratcheting process involving learning. According to Tomasello, "the available evidence suggests that ontogenetic ritualization, not imitative learning, is responsible for chimpanzees' acquisition of communicative gestures," where

> in ontogenetic ritualization a communicatory signal is created by two organisms shaping each others' behavior in repeated instances of social interaction. For example, an infant may initiate nursing by going directly to the mother's nipple, perhaps grabbing and moving her arm in the process. In some future encounter the mother might anticipate the infant's impending behavioral effects at the first touch of her arm, and so become receptive at that point—leading the infant on some future occasion still to abbreviate its behavior to a touch on the arm while waiting for a response. . . . (Tomasello 2000, p. 176)

Similarly, the human mother who sees her baby reaching for something may hand it to him, from which he soon learns simply to hold out his hand toward something he wants. Consider, as more sophisticated examples, how members of a sports team, or of a string quartet, or dance partners, may fine tune their coordination through practice, without an explicit understanding or awareness of the subtle signs they are using to accomplish this. It is possible that the emergence of a certain amount of intentional human signing (not necessarily involving explicit intentions) may have originally evolved in this way in connection with mutually beneficial social activities such as cooperative hunting, warfare, the creation of environmental structures like shelters and fortifications that benefit all, and so forth. Intentional signs originating in this way might then be passed on among humans by imitation.

If this is right, however, one would expect there to be a vagueness in some cases between production of a sign merely as a preparatory movement or natural sign and production of it as an intentional sign, and also a vagueness in some cases between instinctive recognition of

the sign's significance and recognition as a result of past experience and learning. For example, there would be a vague line sometimes between animal signals that were still merely natural signs and those that were intentional. And there would be a vague line between natural signs that help tune members of sports teams and string quartets to one another's doings and intentional signs that serve this function. This is one way, then, in which intentional signs emerge, but only gradually, from natural signs.

Having become intentional signs, animal signals do not lose their character of being also natural signs. They continue to be locally recurring natural signs in the sense defined in chapter 3. Even if the signaling animals or humans were to come to produce the signals through a large number of mechanisms, for example, on the basis of various different inductive and abductive inferences, still these signs could be read exactly as are recurrent natural signs. It makes no difference to the interpreter how various or how complicated the mechanisms of sign production are so long as the signs correlate well enough with corresponding world affairs within some trackable domain. (False intentional signs remain, of course, as entirely unlike natural signs. There are no false natural signs. And intentional signs that are true only by accident are also unlike natural signs, for if reading them produces true belief or productive action, that is sheerly accidental.)

In chapter 2 I discussed two ways that words and sentences can acquire proper functions or purposes, emphasizing that these different origins of purpose sometimes cause linguistic tokens to conflict in purpose. First, there are the purposes of public word types and syntactic form types as replicating memes having their own linguistic functions. These functions are, as it were, the current survival values of these linguistic types, given current cooperative speaker uses and hearer responses to them that encourage their proliferation. Speakers in the language community are adapted to an environment in which hearers are responding, sufficiently often, to the forms speakers produce in ways that reinforce these speaker productions. And the hearers in the community are adapted to conditions under which speakers, sufficiently often, produce these language forms in circumstances such that making conventional responses to them aids hearers. These are the *conventional* functions of the language forms. These functions sustain the conventions of responding to these types in certain ways and, reciprocally, of producing them for certain

purposes.[1] Since within each language community speakers have been designed by learning to use certain linguistic forms to produce certain responses in hearers, while hearers have been designed, reciprocally, to respond in expected ways to speakers using these forms, these forms are produced and used or interpreted by mechanisms designed to cooperate with one another—the requirement for being cooperative intentional signs (chapter 6).

Insofar as these public signs need to map onto affairs in the world in order to complete their memetic functions normally, they also fulfill the mapping requirement for being intentional signs. As will be remembered, descriptive intentional signs can serve their full functions by normal mechanisms only if they already map correctly, whereas the functions of directive intentional signs are to produce world affairs onto which they then will map correctly. Pushmi-pullyus work both of these ways at once ("In this classroom, Johnny, we raise our hands when we want to speak"). The various grammatical moods of sentences each has a variety of conventional functions. These forms are, as it were, polysemantic in function (Millikan 1984, chapter 3; 1998; 2001a). For example, the indicative mood in English sometimes conventionally serves a directive function ("You will report to the commanding officer tomorrow at 6 A.M. sharp!") and sometimes conventionally serves a pushmi-pullyu function ("The meeting is now adjourned").[2] Like intentional animal signals, when descriptive forms are true and not by accident, and when directive forms are complied

1. By saying these usages are "conventional," I mean only that they are replicated or in some way reproduced and that their forms are largely arbitrary with respect to their functions. Other forms could have done the same jobs had they been precedented instead. For a careful explication and defense of this analysis of what makes public language forms "conventional," see Millikan (1998).

2. See Millikan (1996). On this analysis, question forms, whether indicated by syntax or by tonal inflection or both, are directive (as is traditionally supposed).

It is not the function of every directive public language form to produce an intention; nor is it the function of every descriptive form to produce a belief. There are ways of functioning other than through the production of desires, ways that result in effects systematically isomorphic to the language forms that produce them, and there are ways of functioning other than through the production of beliefs, ways that require language forms to correspond systematically to certain kinds of world affairs. An example of the latter is the form "'*X*' means *Y*," discussed earlier in chapter 7, which has truth conditions but whose function is not to produce a belief. Another is the form ". . . exists" (Millikan 1984, chapter 12). Also, there are functions that ride piggyback on more basic functions, as is the case with hypotheticals and various of the modals. But I must leave all of that aside here. (Gunnar Björnsson at the University of Stockholm has been investigating the functions of various of the modals from this point of view.)

with and not by accident, public linguistic forms are locally recurrent natural signs as well as intentional signs. And like locally recurrent natural signs, one must track their natural domains in order to read them. For example, the indicative mood serving a descriptive function occupies one domain, the indicative mood as conventionally serving a directive function occupies another. Each use of the mood proliferates relatively independently (Millikan 1984, ch. 4).

The second way that words and sentences acquire proper functions or purposes is derived directly from speakers' purposes in using them. The purpose of a speaker in producing a linguistic token lends it a derived proper function or purpose (chapter 2). Functions derived from a speaker's purpose in using a public linguistic form may or may not accord with the form's own purpose, depending on whether the speaker is using the form in a conventional or a nonconventional way. Consider, then, tokens that are used in nonconventional ways, for example, fresh figures of speech or fresh Gricean implicatures. The mere fact that a speaker uses a linguistic token with the purpose of producing a certain response in a hearer is not enough to make that token into a cooperative intentional sign token having that purpose. A cooperative intentional sign has to be produced by a device designed to cooperate with its interpreting devices, the interpreting devices being, reciprocally, designed to cooperate with devices designed like it. Does it follow that language forms used in nonconventional ways are not cooperative intentional signs?

Both the impulse to communicate and the capacity to understand that others are attempting to communicate seem to be built into human children but seem not to be attainable, or seem attainable only in very small measure, by members of other species.[3] Only human children spontaneously follow the gaze of others, understand pointing and other gestures intended to call their attention to what another attends to. Children deprived of the use of a public language by deafness spontaneously invent ways of communicating (Kegl and Coppola 1999). They spontaneously suit their gestures to the natural capacities for interpretation of the people around them. The deaf child attempting to make her wants known through exaggerated preparatory movements, through miming, and through attention-drawing gestures creates cooperative intentional signs (in our technical sense), for the systems in the child responsible for the impulse to communicate in this fashion have

3. For discussion, see Donald (1991), pp. 132–134.

certainly coevolved with the systems in her human interpreters capable of coming to understand these communications. These two systems have been designed to cooperate with one another. Neither is found in other species, or not to any significant degree. When the child is successful, the signs she uses have as purposes to produce definite responses on the part of the interpreter, and the interpreter also understands and cooperates entirely purposefully. However, signs of this sort are not recurrent natural signs.

The elements of successful communications are repeated by the deaf child, becoming stereotyped or conventionalized over time, and are more and more easily understood by the child's caretakers and peers. Thus the child's signs slowly move from having purposes derived only from intention in use with a cooperative hearer to acquiring conventional memetic functions. They slowly become elements of a local public language. They slowly become recurrent intentional signs, which implies that they are also, in normal cases, recurrent natural signs. In a similar manner, nonconventional uses of public linguistic signs in normal conversation that are spontaneously understood by the hearer are intentional signs carrying content intended by the speaker. These uses also readily become conventionalized, producing recurrent intentional signs. And just as a vague line can exist between animal signals that are still serving mainly as natural signs and those that are fully intentional, there is often a vague line between unconventional uses of language that are slowly becoming conventional and fully conventional uses, producing recurrent conventional signs.[4]

There is an interesting similarity between many nonconventional uses of conventional linguistic forms and the communicative miming of deaf children. The deaf child who uses mime to communicate does not actually engage in the activities mimed but does count on the interpreter's understanding of what activities these are. Miming is like picturing; the ability to see the mimed event in the miming resembles the ability to see a pictured object in its picture. Similarly, the speaker who turns language to other than conventional purposes involving figures of speech or Gricean implicatures does not use these bits of language for their conventional purposes (or not for their conventional purposes alone) but does depend on the interpreter's understanding of what these conventional purposes are. It may be that sarcastic, hyperbolic, and joking uses of public linguistic forms are best understood as pre-

4. This point is expanded in chapter 11, and more so in Millikan (2001a).

tending or miming the conventional uses of these forms. This suggests that there may also be a close relation, at least in some cases, between reading nonconventional uses of conventional signs and reading natural signs too. But I will not pursue this matter here. Signs used nonconventionally certainly appear not to be recurrent natural signs in any straightforward way, and hence would seem not to be readable in the same way that recurrent natural signs are read.

Returning to conventional uses of linguistic signs, notice that the transition from recurrent natural signs to recurrent intentional signs does not affect the kinds of semantic mapping functions that apply. So it does not affect the kinds of abilities needed to interpret these signs. The slow transition from an anticipatory movement produced as a by-product and read merely as a natural sign to its purposeful stereotyped intentional form has no effect either on its semantic mapping function or on the ease or difficulty of tracking its domain. Recurrent intentional signs produced for an interpreter designed to read them might also be read as recurrent natural signs by an independent interpreter not designed symbiotically to interpret them. Thus I might read the rabbit-thump as a sign that something is threatening the rabbits, or read the bee dance as an indication of where to find nectar. Nor does it matter to the purposes of the chick whether its mother's food call is merely a recurrent natural sign, or also an intentional sign. Just as you or I can learn what a black cloud means or what the geese flying south means, so might a Martian learn to read conventional human language signs, even mistaking them, perhaps, for being genetically determined (like bee dances) or locally recurrent natural signs (like preparatory movements).

Of course conventional human language signs could be read this way only when true or satisfied. Similarly, black clouds mean rain only when it actually rains. To interpret a locally recurrent natural sign you have to stay within the boundaries of its domain. To interpret a recurrent intentional sign, you also have to track its domain, which is extended through the medium of competent, reliable, and sincere speakers of the language who have learned from one another. That is not always an easy thing to do.

I propose now to argue that, in fact, conventional signs used for their conventional purposes usually are read in exactly the same way that natural signs are read. This may be a somewhat unintuitive claim. In the remainder of part III I will spell this claim out in detail, discussing a number of its consequences, along with consequences of the thesis

that the distinction between natural and conventional signs is not sharp but graded.

Chapter 9 argues that understanding language is at root just one more form of sensory perception. One hears what goes on in the distal world through the medium of other people's perception and speech transmission systems just as one sees what goes on in the distal world through a transmitting medium such as television, or through the medium of normally surrounding light. This perception is "direct" in a strong sense that I will explain.

Chapter 10 argues that just as no intentional representations of retinal images intervene between physical objects and the seeing of those objects, no representations of speaker intentions in speaking need intervene between world affairs spoken of by speakers and hearers' understandings of those world affairs. Just as one can interpret information filtered through a pair of binoculars currently trained and focused on a certain domain without having any understanding of what is inside the binoculars and why they work, one can interpret information filtered through another person's perceptual and cognitive systems currently trained and focused in a certain direction or on a certain domain without knowing anything about mind mechanics. Conventional language signs possess their own reference class domains, just as natural signs do. These domains have to be tracked exactly as do the domains of natural signs (unless one happens accidentally to live one's life largely or exclusively within the domain). Tracking these domains often involves following in the wake of the eye gazes or mental gazes of speakers, but need not involve thinking about speakers' intentions.

Chapter 11 is about the semantics–pragmatics distinction. I will interpret this as the distinction between what is conventional in language use and what is, instead, a matter of communicative cooperation between an individual speaker and an individual hearer. I will argue that there is a very wide and vague boundary between these extremes. The distinction rests on statistics over individual psychological processing techniques, which may vary widely, not merely among speakers of the same language, but for the same individual speaker on different occasions.

The distinction between semantics and pragmatics has sometimes been interpreted as separating the contribution to meaning of language proper from the contribution to meaning of language context. In

chapter 12 I will argue that this is a confusion. It arises through a failure to understand the exact way in which communication conventions concern not only phonology, syntax, and lexical conventions but also context. A proper understanding of how context functions within the conventions of language leads to a new interpretation of the semantics of indexicals, demonstratives, and referential descriptions.

9 Direct Perception through Language

I will argue that understanding language is simply another form of sensory perception *of the world*. I have already argued that perception is a way of understanding natural signs or, better, of translating natural signs into intentional signs. So this will help pave the way to the view that understanding language is very much like understanding natural signs.

A sign of a world affair that in turn signs a second world affair may itself be a sign of that second affair (chapter 4). Similarly, if the second sign is a sign of a third—and so forth. And there is always a direct semantic mapping function from the first sign to the last affair signified. A certain sound may signify that the dehumidifier has come on when heard from our bedroom at home, and this in turn may signify that the local power failure is over (a frequently recurrent sign in the rural area where we live). In our summer cabin an indistinguishable sound may signify that the refrigerator has come on, in turn signifying that we are not yet out of propane. To interpret these signs, you must be sensitive to the sign domains they inhabit. But in this case, the domain in which the sound signals the dehumidifier and the domain in which the dehumidifier signals the power are the same. Likewise, the domain in which the sound signifies the refrigerator and the refrigerator signals a nonempty propane tank are the same. A child then might simply hear the dehumidifier sound *as* the sound of electric power coming on, not being aware that it is the dehumidifier that produces what she reads as a sound of electric power, or the child might hear the sound of the refrigerator directly *as* a sign of propane and hence as a sign that we won't go to town today for gas.

Similarly, suppose that certain patterns on the rabbit's retina are natural signs signifying the presence of a fox and these patterns mediate between the fox and the rabbit's awareness of the fox. It will

not be necessary that the rabbit possess an intentional representation of the retinal patterns in order to recognize the fox. Certain patterns traveling up its optic nerves will typically be *natural* signs of what is happening on its retina, but they will not be *intentional* signs of this. An intentional sign signifies only what it is *used* to represent when it operates normally. The optic nerve patterns are not used to guide either the rabbit's behavior toward its retina or any inferences about its retina. It does not matter to the proper functioning of these patterns that they have been produced in a normal way from retinal images, rather than, say, by experimental electrodes, so long as they coincide with a fox. The images on its retina are of no more concern to the rabbit than the direction of magnetic north is to the anaerobic bacteria we encountered in chapter 6, and for exactly the same reason. It is true, of course, that I can think about the images on my retina if I like, perhaps even know quite a lot about them. But when it comes to perceiving a fox, I have no more need to represent retinal images intentionally in the process than does the rabbit.

Exactly similarly, the words that the dog hears when its master says "Go for a walk?" may move the dog directly to the expectation of a walk, that is, to an anticipatory mental representation of a walk, without the mediation of intentional representations of the master's intentions, or of the sounds its master has made, or of the words its master has spoken. You and I are capable of mentally representing those intentions, those sounds, and those words, but the dog, probably, is not. Certainly he has no need to. The question naturally arises, then, whether it is necessary for you or me to harbor intentional representations of the sounds, the phonemes, and/or the words that mediate when someone calls out to us with the message, for example, "Dinner's ready!" Why should it be more complicated for us than for the dog?

Well, I think it is more complicated. First, let me explain why it has to be more complicated. Then let me explain why, despite these complications, there still is an important sense in which, in routine cases, perception *of the world* though the medium of language can sensibly be called "direct perception."

Consider for a moment the connectionist network, VisNet, which was trained to recognize each of seven different faces, each presented through photographs at nine different angles in succession (using Hebbian induction and a trace rule), each succession of nine views of the same face being presented to the network 900 times. The result was

that the network was able to recognize each of these faces with great reliability from each of these nine angles (McLeod et al. 1998, pp. 294ff.). This is quite an achievement for a contemporary connectionist net. But if such a net were now presented with an eighth face, it would take it just as long to learn to recognize that new face as it took it to learn each of the seven previous ones, indeed, perhaps longer, because of interference from traces already laid down. A face-recognizer built like VisNet would be like a brain that relied on semantic mapping functions that went directly from retinal stimulations to inner intentional signs of, say, Johnny, without first passing through a stage that represents the objective shape of Johnny's face and then interpreting this shape as a sign of Johnny (chapter 4). Or it is like the child who relies on a semantic mapping function that goes directly from hearing the words "president of the paleontological society" to an inner sign for Mommy, without first going through a stage at which the words "president," "paleontological," and "society" are understood (chapter 4). Or, of course, it is like the dog who goes directly from hearing the sounds "Go for a walk?" to expecting a walk without going through the process of first recognizing the phonemes, then the words these represent, then the syntactic structures involved, moving finally to an interpretation of what they represent.

If you have need to learn many new faces quickly, simulating VisNet is not a good way to go about it. It is more efficient if you first put in place the ability to recognize *any* shape, or at least any facelike shape, as such, from *any* arbitrary angle and at *any* arbitrary distance. It is better if you first put in place the capacity called "perception of shape constancy." Once you have the general capacity to recognize same-shape-again, for pretty much any shape, from pretty much any perspective, learning to recognize Johnny involves only learning which face-shape is a recurrent sign, in your locale, of Johnny. Learning to recognize Sally uses the same general capacities over again, with the minor variation that you must learn which face-shape is a local sign of Sally. Similarly, if only the dog could learn first to recognize each of the various phonemes in his master's language through the variety of their possible acoustical manifestations, then learn to recognize the words that the presence of various strings of these phonemes may express, then learn to recognize the syntactic forms that can be expressed by various strings of these words, then add a grasp of the semantic mapping functions from words-plus-syntax in his master's language onto what these sentences conventionally represent, it would not take

him dozens and dozens of exposures to learn the significance of each new one of the trivial number of expressions of which he may (roughly) grasp the meaning by the end of his life.

(I have spoken here of the phonemes as *expressing* words, not as composing words, and of strings of words as *expressing* syntactic forms because words considered *as* signs do not equal strings of phonemes. Exactly the same phoneme strings can compose different words. Words and syntactic forms are [aspects of complete] signs and, as we know, exactly the same physical type, depending on the domain it is in, can compose quite different signs, signs of quite different things.)

That human speech perception is routed through the recognition of phonemes and only later of words is evidenced, for example, by the fact that all same-sounding words are primed when a word is presented, even if they have quite different meanings (Swinney 1979). Words must be recognized by context after recognizing the phonemes expressing them. And the same is often true of syntactic forms. Just the parts and ordering of parts of which they are composed, their surface forms, do not identify them. Context, often including inner context (the particular words within the syntactic arrangement), is required to keep track of the meme families that are syntactic forms. This is evident in examples of syntactic ambiguity and mood ambiguity, for example, in the sign "Recycle cans and waste paper" and in the sentence "You will attend university," which might be descriptive (a fortune teller's prediction) or might be directive (Father's orders).[1]

There seems to be a good reason, then, why humans, in understanding human language forms, should go through a process in which intentional representations of things that are signs of signs and so forth are formed, passing through a number of layered stages of intentional representation in the process of translating public language signs into inner representations of world affairs. Similarly, neurological evidence suggests that ordinary visual perception involves the translation of gradients of luminance across the retina as signs of various rudiments of visual form, such as lines or edges with a particular orientation, right angles, ocular disparity, directional movement, color edges, and so forth, which are then interpreted as signs of such constancies as shape, mentioned above, and of a variety of other property constancies, such

1. For more detail on this point, see Millikan (1984), chapter 4. What I call "memes" in this essay are there called "members of first-order reproductively established families." The relevant sense of "word," in the present context, is what was there called a "least type" of word.

as constancies for color, size, texture, quality and direction of movement, and so forth.[2] Taking another example from neurology, Zipser and Anderson (1988) have developed an extremely plausible connectionist model of the way representations of the direction of light points relative to the eye are combined with representations of the orientation of the eye within the head to yield representations of the orientation of the light points relative to the head, and they have demonstrated a very close match to the activation profiles actually found among single neurons in the posterior parietal cortex apparently responding to light–head angles. That is, the evidence is that the brain reads signs of light–eye relations coupled with signs of eye–head relations as signs of light–head relations in a *systematic* way. It does not learn to recognize each light–head relation separately. It uses representations at various levels and translates from one level to the next systematically.

Why is this kind of arrangement, then, not an example, exactly, of "*indirect* perception"? Why isn't the obvious conclusion to be drawn that even the most basic representational functions of the brain involve it in making inferences—inferences prior to the perception of objects in space and prior to the perception of objects represented through language? It is easy to imagine falling into a verbal dispute over this issue. What I will do is argue that there are at least three important differences between the way the brain reads natural signs and signs of signs, and so forth, of the world, and the way indirect perception was traditionally understood to go. The result, I hope, will be to convince the reader that the terms "direct" and "indirect," as they were traditionally understood, fail to have any useful application in the realm of perception. Then I will introduce my own suggestion about a sensible use of the terms "direct" and "indirect."

Traditional theories of "indirect perception" contrasted the perception of objects in the outer world with a different kind of perception that they took to be *direct* perception, namely, perception of sense impressions, or sense data or sensations. The idea was that these sense impressions were the first objects intentionally represented by the mind, and that the mind had to perform inferences in order to move to representations of anything outside. Perception of impressions was typically assumed to be not only direct but also infallible. One couldn't represent one's own sense impressions to oneself wrongly.

2. See, for example, Norman (2000), but the general idea here is ubiquitous in the neurological literature on perception.

Perception of outside objects, on the other hand, was fallible, subject to illusion. But on the description of perception of the outer world we have been considering, no such contrast can be drawn. The mind–brain does not begin by representing either happenings on the sensory surfaces, like light gradients across the retina, or inner things, like visual impressions. The first steps in perception involve reacting to natural signs of features of the outer objective world by translating them into inner intentional representations of those outer features, for example, of edges, lines, angles of light sources in relation to the eye, and so forth. These are outer things, or relations to outer things, not inner objects. Nor, of course, does the brain represent nor the mind become aware of the vehicles in the brain that do the representing. It does not represent its own representations during perception. If direct perception has to be of inner things, there simply is no direct perception, at least none that is involved during the process of perceiving the outer world. Moreover, on every level, inner intentional representations are fallible. Nothing is direct in the sense of being epistemically "given" to the mind. Interpretation of signs is always fallible, chancy, in very principle.

The perception of outer-world objects is not "indirect" either, not in the classical sense. For the traditional view of "indirect perception" was that representations of the outer world were derived by the use of inference. I have described the derivation via perception of a representation of, say, Johnny, or of the affair represented by a sentence one hears, as involving a series of fallible *translations* from one sign or signs into others. Taking an affair as a sign of another affair is reacting to the first affair by translating it into an intentional representation of the other. It doesn't matter whether the first affair, the one taken to be a sign, is a natural sign or an intentional sign. In either case, sign interpretation is just translation. Where the first sign is an intentional sign, this can be put by saying that interpreting it is reacting to the sign vehicle by translating it into or deriving from it another sign vehicle. The question that arises, then, is whether there is a significant difference between translation and inference. For inference also seems to be reacting to one sign vehicle (or perhaps a pair) by translating what it (or they) signify, or at least a portion of this signification, into another sign vehicle.

Here are two differences between the interpretation of signs, and of signs of signs, postulated to occur within perceptual processes, and the traditional paradigm of inference, a movement from beliefs to further beliefs. First, translations cannot easily be modeled after any of the rec-

ognized forms of inference, deductive, inductive, or abductive. This is because they don't seem to have major premises, either explicit or suppressed. Consider the move from seeing the shape that is like that of Johnny's face to mentally representing the presence of Johnny. The premise "all shapes like that coincide with presences of Johnny" is not presupposed. It may very well not be true, in fact, and the perceiver may be quite prepared to discover that it is not. Somewhere in the world is another boy whom one could not tell from Johnny by a glance at his face. Nor is the premise "most face-shapes like that coincide with presences of Johnny" presupposed. What is relied on here is not an implicit major premise, but a fallible ability you have to track a kind of sign domain. The capacity to track, or the luck to happen to remain within the domain of a certain local sign, is not helpfully modeled as a major premise, any more than your luck in not unexpectedly hitting the wrong kind of patch of ice or sand when trying to keep upright on a bicycle would be helpfully modeled as a major premise.[3]

And there is another difference between paradigm inference and the transitions from sign to sign that take place during perceptual processing. Primitive inner signs, such as the sign indicating the angle of the light point relative to the eye, are dedicated intentional signs. They are likely to have just one or two jobs to do, for example, interacting with signs of the position of the eye in the head to produce signs of the relation of the light to the head. That is all. The classical model for inferences, on the other hand, comes from the realm of belief. It is assumed that beliefs form a system of representations such that any representation in that system could, in principle, interact with any other, either directly or through intermediate beliefs. None is isolated, nor are any groups of them isolated from the main body of beliefs. Further, their interactions do not take the form of set algorithms, but may branch in any of numerous directions from a given starting point or points. Beliefs are not, as such, dedicated in advance to any particular purposes. They may help to serve any of a wide variety of special purposes, not determined for them in advance by the systems that have

3. Another possibility might be to try to model the major premise behind the translation of one sign into another as an implicit identity judgment. "That shape there *is* Johnny's face." That there actually are no such things as identity judgments, that is, that identity sentences don't correspond to any mental representations, is argued in Millikan (2000), chapter 12. Identity sentences are an example of language forms that have truth conditions but whose functions are not to cause mental representations that have truth conditions.

fashioned them. They are designed to interact with other beliefs and desires in a flexible manner, acquiring more specialized purposes depending on their inner intentional environments.

For these reasons I think that assimilating the translations that take place during ordinary perception of objects to inferences will mislead rather than enlighten. There is no such thing as "direct perception," in the sense originally intended by those claiming that ordinary perception was "indirect," and there is no such thing as "indirect perception," in that sense, either. If a distinction is to be drawn, then, between arriving at a representation of the outer world "directly" versus "indirectly," it will have to be drawn in another way. Perhaps it would be best to drop the terms "direct" and "indirect" altogether in this context and stick to describing "just the facts, Ma'am, just the facts." On the other hand, it is useful to have a term to describe moves from sensory inputs to representations that do not pass through stages sensibly called "inference" but involve at most only translations. I propose to save the term "direct perception" for this purpose.

My job now is to argue that coming to believe something by being told it is so, in the typical case, is the formation of a direct *perceptual* belief. Forming a belief about where Johnny is on the basis of being told where he is is just as direct a process (and just as indirect) as forming a belief about where Johnny is on the basis of seeing him there.

You and I have concepts of edges, of various shapes, of phonemes, of words, and of sentences. We can use these concepts in developing theories, for example, of how objects and properties are perceived and of language perception and production. Very small children do not have concepts of such things as edges, shapes, phonemes, words, and sentences. For example, as mentioned in chapter 7, a small child will tell you that "ghost" is not a word because there aren't any ghosts. Most children are unable to segment words into their component phonemes or to recognize phonemes as recurring entities until about five or six years (Liberman et al. 1974). Certainly they have no concepts of phonemes as such. But that their brains employ representations of edges, shapes, phonemes, and words by the time they are beginning to talk seems unquestionable. Their ability to recognize new faces and to understand what is said to them clearly depends on these capacities. This is also true, it appears, for morphemes. Children will generalize from reading, say, the "-er" in "higher" and "bigger" to reading "smaller" and "smarter" but not to "weather" and "bicker" (B. Byrne 1996). Yet probably even most adults do not have concepts of mor-

phemes. Not every capacity to produce intentional signs to be used on the way to producing further intentional signs has the right use or a general enough use plausibly to be considered a "concept."

There is evidence that when we hear someone speak, normally what is said goes directly into belief, exactly as when we observe some event happening directly (Gilbert 1993). We do not first understand what is said and then evaluate whether to believe it. Rather, we first believe what is said and then, if we are not under too much cognitive stress, we may think it over critically and reject it. Subjects who are under too much cognitive load—say, they are trying at the same time to count backwards from 1000 by threes—strongly tend simply to believe whatever they hear. In general, there seems to be no reason to suppose that there is only one particular level of distality at which each sensory modality perceives "directly," in the sense I have defined.[4] When you watch television, you usually directly see what is depicted. You see the newscaster's face and what he is wearing and you hear what he is saying. But if you change your frame of mind slightly, say, you are wondering whether to purchase this TV set or not, you may stop seeing and hearing at this level and instead concentrate only on the quality of the reception. Or you may see the pixels on the screen flashing in patterns, especially if the reception is not very good, or if you are a repair man trying to diagnose the set's maladies. Similarly, as an adult, you can directly perceive the phonemes or directly perceive the words being uttered by a speaker if you want to. But usually you perceive only the world affairs spoken about. Just as what we see is dependent on the depth at which we focus our eyes, the distality of what we directly see or hear depends on where we focus our minds.

Depending on the external media through which information is transmitted for perception, the very same world affairs may appear to the same sensory organs in different guises. Although we have surprisingly good color constancy perception under a variety of lighting conditions, colors that are perceived as the same objective color do not have the same appearance under all conditions. Nor do shapes that are perceived to be the same objective shape have the same appearance from all angles. Ringing bells and clacking sticks are easily recognized for what they are whether heard through air, underwater, in a sound-absorbing chamber, or in an echo chamber, but they do not sound the

4. I am using "distal" here as it was used in chapter 4, to refer to the length of a chain of signs of affairs that are themselves signs.

same under these different conditions. Rain does not sound the same when heard falling on the roof, on earth, on snow, and on the water, even though it may be directly perceived as rain through any of these media. Exactly similarly, rain has a different sound when the medium of transmission is the English language ("It's raining!"). And it sounds different again when the medium of transmission is French or German. What world affairs sound like when transmitted through language depends on the language community you are in.

Why is the notion that understanding and believing what is said to you is just another level of natural sign reading and, often, just another form of direct perception so unintuitive? One reason is that what is given to us in ordinary perception is always given as in some rather definite current relation to us. It is given as happening at the time we perceive it, as happening relatively nearby, and often as bearing quite an exact spatial relation to us. This kind of information is needed to guide action, for how one can presently act on a thing always depends on its present relation to one. Ordinary perception is for immediate action, whereas what one learns through language is not typically used that way. Usually I am not told what exact spatial and temporal relations the world affairs being presented through language have to me here and now. Let us take a careful look at this difference between ordinary perception and perception through language.

One of the many traits that seem to distinguish us rather sharply from other species is an enormous flexibility in learning to read new recurrent signs of affairs at different levels of distality and defined by different kinds of mapping functions. For example, learning to comb hair in a mirror is very easy for us, but interpreting what is in mirrors is not possible at all for most other animals. Some have thought that this has something to do with the development of a "self concept," but there is no coherent argument for this. Why should seeing a part of one's body through a mirror and correctly interpreting it in the way necessary to guide one's reaching and touching behavior require a self concept, if seeing a part of one's body in the normal way for the same purpose does not? Nor should it be thought that mentally representing the relation of something to oneself, perceiving the relation of something to oneself, requires that one represent *oneself* explicitly. Recall, for example, that bees represent the relation of nectar to hive and sun without explicitly representing either the hive or the sun (chapter 7). What is required in using a mirror is only that one accommodate governance of one's perceptions and guided motions to a new semantic

mapping function in taking account of the relation of seen objects to oneself. In the rearview mirror, I directly see that there is a car behind me. The car behind guides my motion in relation to it appropriately and directly. Only a very few of the higher primates are capable of making this shift to seeing things in mirrors.[5] For example, as mentioned in chapter 4, when a kitten first sees itself in the mirror, it tries to approach the other kitten, to smell it and touch it. Failing in this, it then tries to look behind the mirror. Finding nothing there, it immediately adopts the attitude that mirrors produce merely holes in information space. You cannot see anything through a mirror any more than though murky water. It is quite impossible to interest a kitten in its reflection after this first disappointing encounter.

Consider what is required to understand what a photograph represents. There is some question to what degree animals other than humans can learn to "see" anything at all in photographs. Pigeons can be taught to sort photographs into those that picture trees versus people versus water and so forth (Herrnstein et al. 1976). They are rewarded for doing this, of course, so the mental representations they derive from the pictures have a function. And it is very likely that their capacity to do this kind of sorting rides piggyback on their prior capacities to recognize actual trees, people, and water. But it seems out of the question that they actually acquire any *information* from these pictures about what is pictured. This is because a photograph contains no information about the relation *to the current observer* of what it depicts. And what use would a pigeon have for a mental representation, derived from a photograph, that there once existed, *sometime and somewhere or other*, a tree looking just so in front of a house looking just so? Similarly, just as a kitten soon stops bothering with mirrors, most animals soon stop looking at television. If they don't stop looking, it is likely that they misinterpret what they see, as did the small daughter of a colleague of mine who asked "Daddy, how did you get in there?" after watching her father on local TV.

That human perceivers can retrieve information from photographs and television depends on their capacity to use information about distal affairs that are not represented or yet understood as having definite and useful relations *to themselves*. Exactly similarly, information

5. Interestingly, according to Epstein, Lanza, and Skinner (1981), pigeons are able to locate spots on their bodies using a mirror. The authors take this to be proof that recognizing parts of one's body in a mirror does not require a self-concept.

presented through human language forms does not typically include information about the relation to the hearer of the affairs presented. The capacities required to understand human language include, then, not only a marvelous flexibility in accommodating new semantic mapping functions, but also the capacity mentally to represent, which requires having some use for (chapter 6), information that does not include the relations *to you* of the things the information is about. But if you are willing to extend notions of perception far enough to cover "seeing" the newscaster on TV even though you do not perceive his relation to you, the extension to "hearing" the news events through his recounting is the same kind of extension. There is no shift in directness of perception, but only a lessening of content in what is perceived. Information about relations to self have dropped out.

But there is another important reason why believing what is said to you may not seem to resemble direct perception. The reliability seems to be quite different. Recall the discussion of verbs of perception in chapter 5. These verbs equivocate, posing sometimes as achievement verbs and other times as verbs merely of aiming or trying. This is one of the factors that may make it seem that ordinary perception of objects and affairs in the world is far more reliable than acquiring information through language. You cannot see what isn't there, but you certainly can hear what isn't true. But, of course, you also *can* see or hear what isn't there—"hearing voices," for example—just as you can hear what isn't true. When we speak of "hearing" that such and such is the case through the medium of language, however, the aiming sense of "hearing" predominates, whereas ordinary "hearing," like "seeing," is more likely to be meant in an achievement way. "Yesterday upon the stair I saw a man who wasn't there"[6] has a very peculiar ring. So does "Yesterday upon the stair I heard a man who wasn't there." But "Yesterday I heard there was a man on the stair but there was no man there" is quite straightforward.

There is a reason for this. Ordinary perception is indeed considerably more reliable than what one hears said, at least under common circumstances. It is not easy to fool ordinary perception. To create strong perceptual illusions generally requires a good deal of knowledge about the perceptual mechanisms and often quite special equipment, for example, of the kind optometrists have in their examination rooms. But surely these illusions should not be classed as indirect per-

6. The verse continues, "He wasn't there again today; My God I wish he'd go away!"

ceptions just because they are deceiving. There is no such thing as infallible perception of anything. Suppose you have new lenses with a new strong correction for astigmatism so that the sidewalk in front of you looks curved or wavy. Do you start perceiving or "observing" the world "directly" again only after adjusting to the glasses? It is reasonable to say that you see the news commentator on television, but what if the filmstrip he shows you is dubbed or outright faked? Dubbing of films is currently the rule rather than the exception. Is there a difference of kind between believing what you apparently see when a film as been dubbed and believing what you hear someone say when it's false? In the modern world, if you want to believe only what's true, you often have to apply heavy filters to other methods of perception as well as to perception through language.

The picture I want to leave you with, then, is that coming to believe, say, that Johnny has come in by seeing that he has come in, by hearing by his voice that he has come in, and by hearing someone say "Johnny has come in," are normally equivalent in directness of psychological processing. There is no reason to suppose that *any* of these ways of gaining the information that Johnny has come in requires that one perform inferences. On the other hand, in all these cases it is likely that at least some prior dedicated representations must be formed. Translations from more primitive representations and combinations of these will be involved. If one insists on treating all translation as a form of inference, then all these require inference equally. In either event, there is no significant difference in directness among them.

10 Tracking the Domains of Conventional Signs

I want now to argue that just as no intentional representations of retinal images intervene between physical objects and the seeing of those objects, no representations of speaker intentions in speaking need intervene between world affairs spoken of by speakers and hearers' understandings of their words.[1]

When conventional signs are true or satisfied and when this has come about in the normal way, conventional signs are locally recurrent natural signs. True, tokens of the same conventional sign may have diverse etiologies, through different people's perceptual systems and cognitive systems. They differ from more ordinary recurrent natural signs in that there will usually be numerous kinds of causal paths to their production, depending on the ways that different speakers have managed to translate diverse prior natural signs into a uniform medium of thought and expression. But there are reasons why the same linguistic form continues to coincide with the same kind of represented affair over a certain domain—it is no accident—and we have decided to take that as the primary criterion for a locally recurrent sign (chapter 6). Assuming that this step in the production of a conventional sign has been accomplished through normal mechanisms—the speaker is not confused, does not lie, and so forth—then reading a conventional sign is mainly a matter of tracking its natural domain, that is, determining what reproducing family it has been copied from. Compare tracking the bird species from which a particular ε-track was derived (chapter 3).

Defining descriptions (chapter 4) need to be mentioned here separately. "The dog," for example, is not a conventional sign, say, of Fido. But it can be used as a cooperative intentional sign of Fido. To be an intentional sign it must be produced by a speaker whose ways of

1. My target, of course, is the contemporary neo-Gricean school of pragmatics.

speaking are cooperatively tuned to fit the capacities of hearers whose ways of interpreting are, in turn, tuned to fit the sign-making dispositions of speakers (chapter 8). I am claiming that hearers' capacities to understand speakers are continuous with their capacities to understand natural signs. Normally the hearer will be able to understand the speaker's token of "the dog" as referring to Fido only if this token, given the domain it is in, is a natural sign of Fido. But "the dog" need not be a locally recurring sign of Fido. Perhaps it has never been used to refer to Fido before and never will be again. Recall the discussion of tracking Scamper the squirrel in chapter 4. You track Scamper by recognizing squirrel signs, which in turn are signs of Chipper, given that they occur in a certain domain. A general capacity to track the presence of *squirrel* becomes a specific capacity to track Scamper when restricted to a certain domain. Indeed, although squirrels squeak only under special circumstances and it could be that Scamper does so only once in his life, still a certain kind of squeak may be a natural sign of squirrel and hence, in a more restricted domain, of Scamper. Similarly, "the dog" is a recurrent local sign, in a fairly wide domain, of dogs, and hence, in a much more restricted domain, may turn up as a natural sign of Fido. Fido is manifesting himself in the sort of way that dogs commonly do through the medium of an English speaker.

In cases where the speaker is wrong or confused or lies, and so forth, normally a linguistic sign will still be a natural sign, but not a natural sign of the same thing of which it is an intentional sign. Perhaps it is a sign of the speaker's mental state, of what she believes, or of what it is her intent to induce the hearer to believe, and so forth. It is not, however, a conventional or intentional sign of any of these. Similarly, a conventional linguistic sign that is true for normal reasons is a natural sign also of what the speaker believes, of what it is the speaker's purpose to communicate about, and so forth, but it is not an intentional sign of any of these. Interpreting language signs as natural signs of speaker beliefs and intentions is a rather sophisticated activity. It requires the interpreter to go well beyond tracking and recognizing the memetic family of the sign and what the sign would normally represent. Especially, the hearer must possess concepts of mental states. Most children don't acquire these before age four or five, by which time they have been using and understanding language for several years. The translation of conventional language forms directly into belief or intentional action does not require concepts of mental states. It is a fairly simple, straightforward affair.

Optometrists ask their patients to look through a stereoscope and tell, for example, whether the dog has jumped all the way through the hoop, halfway through the hoop, or whether just its front legs have gone through. Children believe they really are looking at a picture of a dog jumping through a hoop. They assume they are looking at whatever it is they apparently see. For them, "seeing is believing." Adults, on the other hand, may be aware that they are not looking at a single picture at all, but looking at a picture of a dog with one eye and a picture of a hoop with the other. They may have no tendency at all to believe they are actually seeing a picture of a dog jumping through a hoop. In this situation, for them, seeing is not believing. Indeed, if they have had a lot of experience of the right kind, they may be able to tell, straight off, by looking, that their left eye is stronger than their right eye, or vice versa. It does not follow that when looking at a picture of a dog jumping through a hoop in the normal case, they first judge that it looks as if a dog were jumping through a hoop, then judge that they are not looking through a stereoscope, and that there are no other unusual influences disturbing the light that arrives at their eyes, finally concluding that they are actually seeing a picture of a dog jumping through a hoop. Similarly, the fact that there are situations in which understanding what someone says does not result in believing it, indeed, in which *by* understanding what is said one tells straight off that the speaker is confused, wrong, lying, or joking, does not imply that in the normal case one first judges that the speaker is not confused, lying, or joking, and then moves to believing what the speaker says. Normally, hearing that *p* is believing that *p*—possibly not in the purely statistical sense of "normally" (although I would argue that that is so too), but believing that *p* is the default. It is what happens when nothing intrudes.[2]

It does not follow that understanding conventional language forms is merely "decoding" (cf. Sperber and Wilson 1986; Origgi and Sperber 2002). Only signs that have been copied from one another or from the same models to mean the same coexist within the same conventional sign domain. Conventional signs have domains that must be tracked, just as the domains of local natural signs do. Or if they don't have to be tracked, this is purely contingent. It is a lucky fact if the interpreter

2. For a careful defense of this position and a fuller discussion of the theory that communicative speaking involves having "Gricean intentions," see Millikan (1984), chapter 3.

never happens to stumble into another domain where the same surface form means something different, or never stumbles on a speaker using the form who is misinformed, lying, or joking.

There is no such thing, for example, as a pattern of sound that is intrinsically dedicated to one function. Just as nothing can prevent some new disease from developing that causes spots that look just like measles, nothing can prevent conventions arising from diverse sources that accidentally cross, producing homonyms or equivocal words, phrases, or sentences. Putting this another way, there can be no such thing as a metaconvention that determines when a sound pattern has been produced in accordance with a convention, or in accordance with this convention rather than that where two conventions cross. That a token results from convention-following, from reproduction, rather than from some other source cannot itself be a matter of convention, nor can it be a matter of convention that it came from this source rather than that (Millikan 1998, 2003).

The division between semantics and pragmatics has sometimes been interpreted as resting on whether the context of the sign has to be considered in determining the meaning. But if what is meant by "the sign" is merely the sound or the written pattern, clearly its context is always relevant. It must be recognized as part of a certain language, for example, or as one from among several homonyms, or as representing one from among various familiar senses of a certain word, and so forth. The domain from which it comes may be recognized by the country in which it is spoken (hearing "surgery" or "bonnet"—am I in England or America?) or the language of the words that surround it, or the accent in which it is spoken, or by knowing the background of the person speaking, or by knowing the meanings of surrounding words such that it fits with them to make a complete representation, or makes a representation of something that might come to a speaker's mind in this context, and so forth. "Hit me" means one thing when playing blackjack and another when being instructed in boxing. "Break a leg!" conventionally means *good luck* but, again, only when copied within a certain kind of context from prior such uses in similar contexts. Similarly, where tokens of sentences identical on the surface are conventionally used to perform different speech acts, as when the indicative mood is used for giving orders in the army or the interrogative mood used for making requests, though these usages are thoroughly conventional, they can be tracked only through context.

A particularly clear example to show the continuity of language-sign tracking with natural-sign tracking is the tracking of proper names. How do you know which John is meant when somebody says "John"? That is, how do you know which of the dozens of memetic families of referring "John" tokens this particular token has been copied from? Just as you may have to take into account what part of the country or the wood you are in to recognize quail tracks, you may have to take into account with whom the speaker is acquainted, or where the speaker has just been, or what general domain he or she has in mental focus, say, family or work or the last hunting trip, to know which "John" domain this token comes from.

Does taking into account the domain the speaker has in mental focus require thinking about the speaker's mind? The tradition of pragmatics following Paul Grice's work assumes that understanding a communication always involves that the hearer should first "understand what the speaker intends to communicate," and that this implies that the hearer knows, say, that the speaker intends her to believe such and such, or intends her to do such and such.[3] One upshot of this tradition has been a great deal of interest in the question of which if any of the nonhuman animals possess a "theory of mind," for it is assumed by many that this would be necessary for human-style communication. But if you read the phrase "understand what the speaker intends to communicate" transparently, it does not imply that the hearer thinks about the mind of the speaker at all. It describes the content of the hearer's understanding, but not necessarily by using a description of that content that the hearer herself would employ or understand (chapter 7). It means, merely, that the hearer thinks *the same content* that the speaker purposefully communicates. Let me try to make plausible that this is the correct reading of "understand what the speaker intends to communicate" in this context. That is, no thoughts about other people's minds are necessary in order to grasp their meanings during ordinary communication using conventional forms in conventional ways.[4]

3. —or intends her to believe he intends her to believe such and such or to do such and such . . . and so forth. . . .

4. I believe this is largely true for nonconventional uses as well, such as metaphors and other figures of speech, Gricean implicatures, and so forth. The argument for that appears in Millikan (forthcoming) in the chapter entitled "Purposes and Cross Purposes."

Recall the discussion in chapter 9 of what is involved when one sees things in a mirror. How do you reidentify what you see through a mirror? For example, how do you identify the car you saw through the rearview mirror a moment ago with the car you now see passing on your left and moving in front of you? You may identify it in part by its style and color, of course. But different individual cars may have the same style and color. This style and color will represent reappearance of the same car for you only insofar as you are also tracking this car. Your ability to reidentify depends on the fact that the car projects itself in a continuous path through space and time and that you do the same, a way that makes the relation between the two of you over a short period of time predictable within limits. Not that you couldn't make a mistake. One blue Ford seen in the rearview might, in principle, suddenly be passed by an indistinguishable blue Ford that now passes you without your noticing the switch. How one *normally* manages to reidentify is the point; infallibility is always irrelevant. The point is that your ability to track, over space and time, the domain in which presence of a blue Ford will signify *that same car again* does not require you to have any understanding of how mirrors work, of light energy, of principles of reflectance, and so forth. You don't have to know anything *about* mirrors. You only have to recognize what place, relative to yourself, you are seeing when you look in the rearview, and be able to continue tracking places roughly continuous with that same place as you move your glance through the side window and then through the windshield.

Similarly, how do you tell what you are seeing through a pair of binoculars? Which of the birds lined up on the telephone wire just seen way over there with the naked eye is the one you are now seeing through the binoculars? Perhaps you tell a little bit by what it looks like, but you wouldn't be using the binoculars if you already knew what it looked like in full detail. You can tell there is probably a bird there without the binoculars, and with the binoculars you see something that is definitely a bird. But which bird? One that is over there in the same direction the binoculars are pointing. That is the beginning of your tracking. But there seem to be a number of birds over there lined up on the telephone wire, and you can't tell at exactly which one the binoculars are pointing. So you look at the context that surrounds the bird. It is just in front of the Y on the lower branch of the small maple tree just behind the wire. If you can also see that Y without the binoculars, you may be able correctly to track that bird from seeing it with

the naked eye to seeing it with the binoculars. Or perhaps you can see that it is the third bird over from the left and you can see which is third also without the binoculars. The bird's general properties, the bird's general direction in relation to you, the direction the binoculars are pointing, and certain features of the bird's context all combine to enable tracking and hence (re)identification. And if you now move the binoculars slightly to the right, you may identify the next bird on the right a bit more easily than you did the first bird, by the relation it bears to the first. Again, you don't have to know anything at all about the principles behind the operation of lenses, not even that there *are* lenses in the binoculars, to be able to do all this.

Now consider the phenomenon—clearly a very important one—of joint looking. Very early, infants follow the gaze of another person trying to communicate with them. Soon they also use and comprehend pointing and other showing gestures quite spontaneously. In doing so they are following where the other person's attention is directed. This is like observing in which direction the binoculars are pointed.[5] When talking to an infant, where an adult's attention is directed is likely to be where the subject of the adult's conversation is to be found. This helps the infant correctly to identify what she sees with what she is hearing about. What kitties look like when projected through the medium of ambient light is one thing; what they sound like when projected through the medium of another person's speech is different. But the infant learns to recognize these signs as of the same thing. Kitties can be identified and learned about in various ways, through various media (Millikan 2000, especially chapter 6). The infant learns what kitties look like in various postures, what they feel like, the sounds they make, and what they sound like through language. There seems no reason why this last would require that the infant employ a theory of mind or concepts of mental states. Why would the infant need to understand the innards of minds any more than it will need to understand the innards of binoculars?

Now consider how you might identify what you see in a photograph or a home video if you can't recognize it straight off by its appearance. If you know where the roll of film was taken, you will know something

5. According to Gomez (1991), arguing that there is no evidence that gorillas have a theory of mind, gorillas reared with humans understand that looking at another's eyes is a means of controlling useful causal contact. They check to see that a human is attending to the same thing that they are, but for them this is merely part of a causal link needed for an effective interactional process.

about the domain from which the natural signs it contains emanated, and this may be how you recognize that, yes, that is the Ely Cathedral. Or since the photo falls between the picture of Uncle Robert and the picture of Aunt Sally, it must be an angle on their house in Little Falls. Or, not knowing in advance where the roll was taken, you may recognize baby Willie by recognizing older sister Jane who holds him up, and older brother Tom squinting there in the back. Similarly, watching a home video you identify many of the things you see by their context and by spacetime continuities over short periods, as when reidentifying the blue Ford seen in your rearview mirror a moment ago. You identify by recognizing the general domain on which the camera was focusing and/or by first recognizing some of the things in the domain directly. Again, this requires no knowledge and no thoughts about what makes cameras work or what is inside them. Similarly, I may correctly read the volt meter as telling the voltage of my battery by knowing that it is my battery to which it is connected but without having any interest in what is inside volt meters or how they work. These acts of identification require, merely, knowing how to track various kinds of sign domains.

In exactly similar manner, you recognize which John it is that is manifested through a token of the word "John," by knowing with whom the speaker is acquainted, or where the speaker has just been, or what general domain the speaker is in the middle of talking about, and hence on what domain his inner binoculars are focused, what domain he is drawing verbal pictures of for you. And it is by tracking rather than by thinking of speaker beliefs and intentions that you recognize the domains on which speakers are focusing when they use brief definite descriptions such as "the dog," "the boy," "the table," "the lake," and so forth. Knowing on what domains they are focusing, you may know which individual boys, dogs, and so forth are the ones for which occurrence of the properties mentioned in the descriptions are natural signs. That is also how you recognize the domains of quantifiers, as in "Then everyone went out to see the sunrise" (all the people?—*which* people?) and in "Some people were complaining about the food" (some of the people?—some of *which* people?). Because yours and the speakers's perceptual/cognitive binoculars are similarly constructed, tending to distinguish the same figures and ground when focusing on the same domain, tracking is also the way you recognize which pairing relation is being expressed by the possessive, for example, whether "John's book" is the one he wrote, or the one he owns, or the one he is looking

at or carrying, or the one he has nominated for the Pulitzer prize. And that is how you know from hearing just the sentence "John is too small" for what John is too small. It is in this sort of manner that you interpret what the speaker's words refer to—not "what the speaker has in mind," for there is no need for you to think of that which is meant under a mental description. No "theory of mind," no representation of the speaker's beliefs and intentions, is required for this.

Understanding language is seeing the world through the cognitive systems of another person who has learned, been trained, been calibrated, to make manifest in a uniform way, things in the world on which she focuses. To know what is manifested through the conventional speech of another, one may have to know on what this human instrument is focused, what it is currently wired up to. But one needn't know anything about its insides.

11 Varieties of the Semantics–Pragmatics Distinction

The distinction between semantics and pragmatics has traditionally been troubled. At least three different broad criteria have been used in drawing it, with very poor overlap among them.

First, the study of semantics has been equated by some with the study of truth or satisfaction conditions as opposed to the study of the "force" of linguistic utterances. All matters concerning how language functions, what it does or what people do with it, would then fall under pragmatics. In chapters 6 and 7, I discussed the distinction between semantic mapping functions (these are functions in the mathematician's sense of function) and linguistic functions (these are purposes in the sense of chapter 1), claiming that all complete intentional signs are associated with both kinds of functions. And it follows from the discussions in chapters 2 and 8 that both of these quite distinct kinds of functions characterize conventional and nonconventional intentional signs alike. On this interpretation, the differences between tokens of morphologically and syntactically identical sentences used to perform different speech acts, as when the indicative mood is used to give orders rather than to make assertions, or when the interrogative mood is used to make requests rather than to ask questions, are automatically considered pragmatic distinctions. But the differences among the syntactically distinct indicative, imperative, and interrogative moods should then also be considered pragmatic distinctions, for these moods have distinct ranges of linguistic function determined according to context in separate ways. These differences, however, have not generally been considered "pragmatic." (On the other hand, they have not generally been termed "semantic" either.) The distinction between semantic mapping functions and linguistic functions as I have defined it is a pretty clear one, I believe, and it seems best not merely to duplicate this distinction with the terms "semantic" and "pragmatic." To talk

just of "satisfaction conditions" versus "force" might be better here, while acknowledging that the satisfaction conditions and the force of a particular linguistic token are always derived from each of two sources that may not coincide.[1] There is always the memetic function or public use of the token's type, and there is always the speaker's purpose in using the token. Public linguistic force and satisfaction conditions of a sentence token may each either follow conventional usage or be diverted instead by speaker purposes into nonconventional channels.

A second tradition equates the study of semantics with the study of the meanings—both the semantic mappings and linguistic functions—that language forms have considered apart from the contexts in which they occur. Pragmatics then studies what language context adds to contextless meaning. I argued in chapter 10 that because conventional language forms have domains in the same way that locally recurrent natural signs do, there can be no such thing as interpreting a conventional linguistic sign apart from its context even when it is used strictly conventionally. As Sperber and Wilson (1986) put it, interpreting language is never strictly "decoding." For example, context must be used in order to determine whether one is hearing an instance of the word "bolt" as in "Please don't bolt the door" or as in "Please don't bolt out the door," and in determining whether "You will marry Regina" is an instance of the convention that uses indicatives for issuing decrees or of the convention that uses indicatives to make predictions (fortune tellers).[2] If we take this proposed way of distinguishing semantics from pragmatics quite strictly, all such distinctions will count as pragmatic distinctions. But of course not all of them have generally been considered to be pragmatic distinctions. Perhaps the matter could be clarified by drawing a distinction among different *ways* in which context contributes to a hearer's understanding. Perhaps the use of context to distinguish which conventional linguistic form is being used should simply be ruled out of pragmatics. Then taking the distinction, for example, between decrees and predictions to be pragmatic (as has been

1. Note that "force" in this context is not illocutionary force. The public linguistic function of the type and the speaker purpose of a linguistic token both include perlocutionary aspects of function that, when the normal mechanisms accounting for proliferation are engaged, accord with the speaker's and hearer's purposes. (These purposes are not usually *explicit* intentions.)

2. That this last sort of distinction is indeed a distinction between two conventional forms, the surface indicative form being polysemantic, is argued in Millikan (1984), chapter 4, and in Millikan (1998).

customary) is just an error. This would leave within pragmatics the use of context to determine the domain on which the speaker's attention is focused (chapter 10) and the use of context to determine what it is the speaker's purpose to convey when a language form is used nonconventionally. But it would also leave within pragmatics the interpretation of all indexicals and demonstratives, no matter in how strictly conventional a manner they were being used.

A third way of drawing the semantics–pragmatics distinction considers semantics to be a study of the conventional aspects of language use whereas pragmatics is taken to study aspects of communication achieved in nonconventional ways. Again, much of the study of how a hearer tracks the domains on which a speaker's attention is focused will count as part of pragmatics, for surely this tracking is done in large part not by following conventions. And again, of course, the study of how hearers interpret forms that are not being used conventionally will fall inside pragmatics. But there is a sharp difference between this way of interpreting the semantics–pragmatics distinction and the second way mentioned above, or so I will argue. For although it is clear that if a hearer picks up a speaker's message without relying merely on conventional aspects of the speaker's usage the hearer must be using context in order to do this, so that any study of nonconventional language uses would have to make heavy reference to context, the converse does not follow. It does not follow that whenever context must be considered in order to interpret a message, the usage is nonconventional. I have already mentioned, of course, that context is nearly always needed for telling which conventional signs are being used. More interesting, there are many conventional ways of using context as a proper part of a linguistic sign. Chapter 12 will be partly about that.

Whether one adopts the second or the third way of interpreting the semantics–pragmatics distinction, the distinction between semantics and pragmatics will often turn on whether the linguistic form being interpreted is being used in a conventional way or in a nonconventional way involving one-off communicative cooperation between an individual speaker and an individual hearer. In this chapter I will argue that the line between conventional and nonconventional uses of language is vague in the extreme, so that the semantics–pragmatics distinction is necessarily vague as well. The conventional–nonconventional distinction rests on statistics that would have to be gathered over individual ways of psychological processing of linguistic forms, but these ways

may vary widely, not merely among speakers of the same language, but for individual speakers on different occasions. What is a matter of convention for one speaker may not be for another, or for the same speaker at a different time. Further, there may be a number of alternative ways in which the same language form serving the same function may be understood or processed psychologically while preserving the same conventional outcome in understanding.[3]

I have remarked on the vagueness of the line between natural signs and genetically determined intentional signs such as animal signals and certain human facial and bodily expressions (chapter 8). Human facial and bodily expressions, such as smiling and frowning, are especially interesting because during the development of a particular human individual, what is in the first instance genetically based may later come under some control first by unconscious learning systems, and later by conscious systems. Thus smiles emerge naturally in infants and are naturally understood by infants, but they can be unconsciously reinforced just as eye-blinks can, and they can also be produced with conscious intent to communicate. In this sort of case, it is quite easy to grasp how the gradual progression goes from purely natural signs, through genetically determined intentional signs, through unconsciously learned memetic signs, to consciously reproduced memetic signs.

A second sort of gradual transition occurs from signs such as the deaf child's early signing, or new innovative uses of public language forms, into conventional signs, that is, into signs with stable memetic functions (chapters 2, 8). A conventional sign is one that is being used by speakers and hearers to serve a certain cooperative function because it has successfully served that function before.[4] Its use and comprehension are reproduced, not newly invented. The speaker relies on the hearer's familiarity with the particular use rather than relying only on his or her general interpretative capacities. But, clearly, one can be more or less familiar with a use, and one can more or less invent a use, prompted more or less strongly by prior experience. Also, some speakers may reinvent the same use that others copy, and some hearers may already be familiar with uses that others need to grasp afresh from context. Consider, for example, half-dead metaphors, which are often

3. The wide fuzzy band that separates the conventional in language from the nonconventional is discussed at more length in Millikan (2001a). The discussion here is much abbreviated.

4. This thesis about linguistic convention is fully argued in Millikan (1998).

much more dead for some people than for others. Or consider certain common uses of phrases or sentences to perform speech acts arising originally from Gricean implicatures. Is "Can you reach the salt?" a literal question, or is it a literal request? Once you grasp that for a usage to be conventional is just for it typically to recur on account of precedent (Millikan 1998), the debates about what is "said" (that is, conventionally signified) versus what is only pragmatically "implicated" takes on a clearer meaning. But it also becomes clear why the line between these two is wide and fuzzy in very principle. The transition from nonconventional uses of language to fully conventional uses is a gradual, largely statistical matter.

A third sort of vagueness between what is conventional and what is not in language use derives from vagueness about what should count as having reproduced a language element with the same meaning. The development of human languages clearly depends heavily on constraints laid down by the phonological structures and the grammars of each language, for these delimit quite sharply what is to count as a correct copy or reproduction of a language form (chapter 2). But there are no such formal constraints on what is to count as having copied the same use or the same meaning again. This is a very serious issue, which I discuss at more length elsewhere (Millikan 2001a, forthcoming). Here I will mention just two kinds of vagueness that concern meaning.

First, language is by no means always understood compositionally, and hence its meanings are not always copied compositionally. Chunking, the reproduction of whole phrases and sentences, may be the norm rather than the exception, particularly with children and the poorly educated. Children learn at least five to nine new words each day from age eighteen months to six years (Waxman 1991; Clark 1991; Byrnes and Gelman 1991)—Chomsky says, "about a word an hour from ages two to eight with lexical items typically acquired on a single exposure . . ." (Chomsky 1995, p. 15). Surely they are capable of learning a similar number of phrases. Large-scale reproduction of phrases accounts for the subtle distinction between idiomatic and unidiomatic usages of a language: "That makes good sense but it's not how a Frenchman would say it." Thus for a native English speaker, rivers and caves have mouths but buildings do not; bottles and violins can have necks but dumbbells and meadows do not; and in the expression "in my neck of the woods" the word "neck" is not even read as the same word. When speaking on the telephone I say "This is Ruth Millikan" but when introducing

myself in person I say "I am Ruth Millikan"—why the difference?—and so forth.

Recall now that different signs may articulate exactly the same world affairs quite differently (chapter 7). When language is chunked in the mind of the user, the articulation is understood differently than when language is grasped compositionally. The result is that differing interpretations of what this chunk means, and hence what it would mean in other contexts, are readily made. Nice examples of this are historical changes in understanding of the grammar of certain phrases, for example the change in the grammar of "going to" from verb-plus-preposition to verb-with-auxiliary and the change in the verb "will" from a simple verb taking a direct object to an auxiliary verb (Roberts 1985; Roberts and Roussou 2003). Another kind of change of this general type is exemplified by the classic oxymoron "The dog went to the bathroom on the living room rug." Anyone who has had to read papers of first-year university students, in the United States at least, has seen hundreds of chunked expressions, unparsed, hence half-understood, hence misused. Asserting that people *ought* always to understand language compositionally does not help, of course. Indeterminacies concerning the actual meanings of public language forms, the ways these are actually sifted and shifted through public usage, are not touched by such moralistic or aesthetic sentiments.

Second, whether a phrase or sentence type is copied alone to have a certain meaning, or whether it-plus-its-context, either linguistic or non-linguistic, is what is copied, may often be an indeterminate matter, a merely statistical matter resting on quirks of individual psychological processing. Consider the difference between "Have you been swimming?" and "Have you been in Antarctica?" Is "Have you been . . . ?" (a) a polysemantic phrase that sometimes means *have you just been*? and other times means *have you ever been*? Or (b) does it always mean just "Have you . . . ," whether any time, or some particular time, being a matter not of semantics but of one-off pragmatic understanding between speaker and hearer dependent on their personal mutual knowledge? Or (c) is there a *convention* that when used with verbs that denote frequently recurring events it means "Have you just . . . ," whereas with verbs denoting events that don't typically recur it means "Have you ever . . ."? Further, whichever of these ways the hearer manages to interpret the phrase, is it clear that the speaker has to have reproduced it with the same understanding?

Taking a second example, when the adjective "red" is used in the combination "red hair" does it (a) *conventionally* mean something different than when used in most other combinations? (What counts as red hair would never count as red if one were speaking of a dress or a flag.) Or (b) does "red *X*" always just mean "red for an *X*" (Wheeler 1972)? That is, is "red hair" more like "red" in "red herring" or more like "long" in "long hair" versus "long road" (quite a different length)? And however the hearer interprets this phrase, is it clear that the speaker has to have reproduced it with the same understanding?

The very fact that context must be used in tracking the domain of a conventional sign produces a sort of insistent vagueness in the conventional–nonconventional distinction. For if a certain kind of context is sometimes an obvious indicator of the memetic family from which a language form token comes, speakers will learn sometimes purposefully to place their tokens in just such a context in order to be understood more reliably. Soon placement in such a context will tend toward a conventional manner of communication. One speaker copies it from another, and hearers learn to interpret it quite automatically. At what point will we say this sort of careful placement in context is no longer a matter of pragmatics, but part of the conventional semantics of a language? I see no reason to suppose that answers to questions of this sort are somewhere written. Very likely the statistics on actual psychological processing among native speakers, whether different speakers or the same speakers at different times, are quite scattered.

A clear case of the unclarity of the conventional–nonconventional distinction concerns demonstratives. Consider how demonstratives must have evolved. Suppose that the hearer understands what you are talking about, in part, by where you are looking (chapter 10). Then the hearer is taking the direction of your look as if it were a pointing gesture. Indeed, perhaps your looking actually is a pointing gesture, that is, you purposefully and obviously look at something in order to draw your hearer's attention to it, to single it out for the hearer. What was originally merely a natural sign of the domain from which a speaker's language signs were emerging readily becomes a purposeful sign or indicator, used by the speaker. Some disposition to this has either been selected for genetically, or learned in an environment where cooperating hearers are sensitive to direction of look. Generalizing this, speakers learn to leave a trail that allows hearers to track the focus of their attention. As they become more sophisticated, they become more

aware of ambiguities and possible misinterpretations, purposefully filling in more context where needed, painting in, as it were, what bird sits on the telephone wire beside the John they are talking about—where he comes from, his profession, or his last name. (Certainly it is true that having some understanding of the mechanics of other people's minds can help here, but, perhaps more commonly, this resembles checking to see if one can easily read one's own handwriting as one proceeds down the page.)

To the degree that direction of look comes to be purposefully used by a speaker, direction of look becomes an aspect of public linguistic signing. Similarly, drawing the hearer's attention to what you are talking about by picking it up and showing it to the hearer constitutes another aspect of a sign system that emerges gradually from a nonintentional or natural background. The human hearer has a natural tendency to notice and examine what another human is handling and to notice what newly appears in front of himself. Recalling how anticipatory movements of animals, originally read as natural signs, can gradually turn into intentional signals (chapter 8), it becomes clear how various forms of pointing naturally emerge from anticipatory movements that preceded showing as these become stereotyped and exaggerated. In accord with this, there are cultural differences among the ways pointing is done. Sometimes the index finger is used, sometimes the middle finger, sometimes, in conventionally defined contexts, protruding the lips is used (Sherzer 1973). Pointing gestures are a good illustration of the vagueness that can occur between natural signing, intentional but nonconventional signing, and fully conventional signing.

A conventional sign is one that would be unlikely, or much less likely, to be employed and to be understood, or to be understood so easily, were it not being reproduced from examples of prior usage. Conventional usages are memes that are proliferating themselves owing to a particular function they are serving. Because some people have reacted to a sign in a certain way, others have been encouraged purposefully to use the sign to produce this reaction. And insofar as speaking and understanding is, in the first instance, a cooperative enterprise, furthering the interests of both speakers and hearers, that speakers are purposefully using the sign to produce certain reactions encourages hearers to reproduce these reactions, which are then produced more reliably. The sign has attained conventionality when its use for a certain purpose is proliferated not just through genetic transmission, and not

only because speakers are finding it a useful tool, but because, causally because, others are using it or have used it before. The conventional sign is being reproduced or "copied" for a certain function, not discovered or invented anew by each producer–consumer pair. Thus conventionality is clearly a matter of degree. Ways of pointing to or, more generally, ways of demonstrating or, still more generally, ways of drawing attention to the sign domain from which one's linguistic signs are emerging come in all degrees of conventionality. This kind of vagueness may apply as well, of course, to figures of speech, implicatures, and other extensions of usage that are slowly moving from being entirely innovative, through being somewhat familiar, to being handled automatically without parsing or derivation of meaning from compositional structure.

12 Demonstratives, Indexicals, and a Bit More about Descriptions

It is easy to assume that only those aspects of language that involve phonology, morphology, syntax, and perhaps aspects of prosody are strictly conventional. Call these aspects "narrow linguistic aspects." In chapter 11 I mentioned that besides narrow linguistic aspects, many ways of indicating or pointing are conventional. In this chapter I will argue much more generally that narrow aspects of linguistic signs are often merely parts or aspects of more complete conventional signs. Besides pointing gestures, various other aspects of surrounding context are often just as conventional parts of the linguistic sign as its narrow aspects are. Call these "wide aspects," and call the whole sign that includes its wide aspects the "wide linguistic sign." Grammatically separate terms called "indexicals" and "demonstratives" are indicators whose functions are to show explicitly how wide aspects are functioning within the wide conventional signs of which they are a part. There are also wide linguistic signs that contain no such specialized narrow linguistic parts. It is for this reason that defining pragmatics as the study of how context helps to determine the meanings of sentence tokens yields quite different results from defining it as the study of nonconventional aspects of linguistic usage.

Consider first a strongly conventional way of demonstrating what you are talking about. You draw an arrow to it. Suppose that the arrow points to a square you have drawn on the blackboard, and you say "This is a closed plane figure" as you draw the arrow to point toward the square. Exactly what is it that composes the conventional sign in this case? Not only the sentence you utter and the arrow you draw, I will argue, but also the square itself, *the very thing demonstrated*, is a part of the conventional sign. Let me explain.

Recall that when you read a gas gauge, the car it is in stands for itself (chapter 4). Similarly, when you see the label "poison" on a bottle, the

bottle stands for itself. It is a reflexive sign of itself, just as the place of the ε-track signifying quail is a sign of itself, and just as one inch on a blueprint may stand for one inch on the model to be built (chapter 4). A complete sign always involves a whole world affair that stands for another whole world affair. The affair that is *the-label-"poison"-attached-to-a-bottle* represents the affair that is poison being inside the bottle. The bottle is part of each of those affairs, the affair that is the sign and the affair that is the signified. Unlike the ε-track and its time and place, however, the positioning of a label on a bottle and the word "poison" itself are aspects of a conventional rather than merely a natural sign. First let me illustrate how the square drawn on the blackboard is like the bottle, each conventionally standing for itself. Then I will show how the placement of the little word "this" inside the sentence "This is a closed plane figure" is functioning to help the semantic mapping.

Going back to animal signs for a moment, consider, say, the mating dance of the male stickleback fish. It indicates to the female that that very male is now ready to mate with her—not some indeterminate male, but *that very male*! If we recall that instinctively produced animal signs are not merely natural signs but intentional signs as well (though not, of course, conventional signs), the dancing male stickleback clearly functions as a reflexive intentional sign of himself. Similarly, some bees dance horizontally such that the direction of the dance relative to the hive and the current location of the sun represents that very same direction as the direction of nectar relative to hive and sun. The orientation of the dance in its wider environment intentionally signifies that same orientation of the nectar in its wider environment. Thus, not just the dance itself but its orientation within the wider environment is part of the intentional sign produced. Feral members of *Felis Domesticus* often do not bury their feces but leave them exposed in conspicuous places as a sign of their dominance within a territory. That their domesticated kin usually bury their feces is a sign of submission, most likely to the humans that feed them (Milius 2001). Where the feces are placed is certainly an integral part of the intentional sign produced, and not merely part of its environment.

Similarly, although we usually think of language signs as composed merely of phonological units, words, syntactic forms, and so forth, often the situating in a context is a reproduced element too and is equally an aspect of the conventional sign. Conventional signs can include aspects of the environment as proper parts, just as natural signs and animal signs can. Then the distinction between a sentence and

certain parts of the environment in which it is embedded is not semantically significant. When a label is placed on a bottle, it is a matter of convention that the bottle stands for itself, given that there is a label on it. Suppose I wave my arm across the landscape before me and exclaim "Wow! Breathtaking!" I am labeling the landscape, which stands for itself. Thus that the square to which the arrow points represents itself is not peculiar. The fact that the square lies both outside of the sentence proper and outside of the conventional indicator, the arrow, that points to it does not preclude it from being just another aspect of a conventional sign.

What may at first seem peculiar, however, is that the little word "this" in the sentence "This is a closed plane figure" seems also to be standing for the square, so why would the square be needed, redundantly, to stand for itself? The word "this," considered merely as part of the English language, does not, of course, represent either this particular square or squares generally. Compare it with an x on a map that is obviously intended to represent something or other to be found in that location on the mapped terrain. So you consult the map's key to find out what kind of thing it is that is in that location. In the key, next to another x, you find the word "hostel." The x on the map shows the place where the something is, and the key tells what something it is that is in that place, namely, a hostel. Similarly, the word "this" in "This is a closed plane figure" shows by its grammatical place where there is something in semantic space, while the arrow pointing to the square is a key that shows what it is that occupies that semantic place. The word "this" holds a place for the square in the affair represented exactly as the x on the map holds a place for the hostel. Another interesting case for comparison might be the stories that sometimes appear in children's magazines in which pictures of various denoted objects have been substituted for the names of these objects in sentences about them. The pictures appear in certain grammatical places and the pictures show what occupies those places. For amusement, compare also Quine's "Giorgione was so called because of his size" (Quine 1960, p. 153) or, more perspicuously, "Giorgione was called that because of his size."

Where an aspect of context is part of a wide conventional linguistic sign and there is a filler or variable entered into the syntax of the narrow linguistic part of the sign that holds a place open for that aspect of context to fill, we can borrow linguists' terminology and say that the grammatical position has been "lexicalized," and, since it has been given a grammatical place, also "grammaticalized." For example, in

"Wow! Breathtaking!" the subject has not been lexicalized and grammaticalized, whereas in "Wow! *That* is breathtaking!" it has been. In "Careful! Poison ivy!" reference to place has not been lexicalized and grammaticalized, whereas in "Careful! There's poison ivy around here!" reference to place has been lexicalized and grammaticalized.

Phrases and sentences differentiate themselves, standing out from the rest of the natural world as indicating the presence of signs designed to be interpretable. When you recognize what you hear as a sentence, you expect it to be part of an interpretable sign, interpretable, at least, by someone. Whether or not the context in which a narrow linguistic form appears constitutes an aspect of a wide conventional sign having a fully conventional meaning may often be a matter of degree (chapter 11). But where a grammatical placeholder has been inserted, it is clearly a matter of convention that *some* aspect of context is intended to fill out the sign, either in a completely conventional way, or in a way that the speaker tries to indicate in a nonconventional manner to the hearer, relying on nonconventional methods of tracking. An example of the latter kind might be a token of "Aunt Nellie saved this one" said of a wildflower pressed between the leaves of a book that I hand to you closed, but you understand me because we have just been looking at and talking about various other flowers preserved in this way.

In chapter 4 I argued that the time and place of a natural sign such as a quail track could be significant aspects of the sign but were not "indexical" elements. Time and place of the quail track signify time and place of the passing quail in the same way that the size of the track signifies the size of the quail. That was perhaps rather high-handed of me; more of a decision on usage, designed to help underline certain similarities and differences, than a statement of prearranged fact. In the linguistic case, there is a fairly clear distinction between aspects of the sign that are freely created in accordance with general conventions of phonology, morphology, syntax, and, perhaps, aspects of prosody—the narrow linguistic sign—and aspects that can clearly be designated as a "context" in which these free aspects are placed. In the case of purely natural signs, there is no distinction of this sort to be drawn. All parts or aspects of a natural sign are just more parts or aspects of the sign proper. Similarly, however, all parts of a linguistic sign placed in a context that completes it in a conventional way are just more parts of that conventional sign. No radically new principles of signing are introduced when the aspects of the context of a narrow linguistic sign are

conventionally incorporated within it. I am suggesting that we reserve the term "indexical," then, for a special kind of sign, namely, for lexicalized and grammaticalized elements of a public language whose job it is to indicate explicitly how elements of context are to be positioned within the mappings of wide conventional signs.

The difference between a demonstrative and an indexical is, then, that the indexical conventionally indicates not only the syntactic-semantic place the context is to fill, but also fully indicates in a conventional way what the relation is between the indexical word and its outside filling. No further indicator such as a gesture or additional indicating context is required. I will unpack this claim.

The time and place of a linguistic sign token can be considered to be an aspect of nonlinguistic context, as distinguished from freely created aspects of the sign. This is because every linguistic sign token has to have *some* time or place, but since this time and place are not always significant aspects of the sign, in the particular cases where the time and/or place do carry part of the meaning, one might consider this additionally meaningful element to be a wide or contextual addition, rather than a narrow linguistic element. Similarly, every linguistic sign has some producer, and almost all are produced to be interpreted by somebody. But, typically, the identities of speaker and hearer are not relevant to the meanings of linguistic signs. So in the particular case where the identity of the speaker or of the intended hearer carries part of the meaning, one might consider these elements also to be a wide or contextual addition rather than narrow linguistic aspects of the sign. The uptake of context into a wide conventional sign, considered simply as such, is no different in principle from the way the time and the place of the quail track is taken up into the full natural sign concerning the quail. However, when reference to its time or place is lexicalized and given a grammatical place within the narrow linguistic sign, convention leaves no room for question about exactly what contextual aspect completes the conventional sign. Let me supply some contrasting examples.

Suppose that you are a surgeon and I am your assistant, and during an operation you direct "Scalpel!" then "Scissor!" then "Suture!" Who is to do what with the scalpel, then the scissor, then the suture, and when and where is not lexicalized, but it is determined by the context in an entirely conventional way. Similarly, if I say "It's raining," the place at which I say this conventionally determines the place of the intentionally signified rain. For example, the following dialogue is not

possible *within* the conventions of English (which is why it could be a joke).

"It's raining!"

"Where?"

"In Tahiti."

"It's raining," standing alone, simply is not a way you can conventionally say, in English, that it is raining somewhere or other. Taking another example, it seems that the Machiguenga Indians have no proper names for one another but use kinship names instead (Snell 1964). Thus one must know the identity of the speaker in order to know who is being spoken about. In English, "Mama" and "Daddy" work that way. It is fully conventional, a convention that has a very wide domain, that the person referred to by these terms depends on who is related in a certain way to the one speaking. But reference to the person speaking is not lexicalized. Compare with the one-word sentence "Scalpel!" the five-word sentence "Now hand me the scalpel!" Here positions for time, for the person to whom the scalpel is to be handed, and for being handed rather than, for example, tossed or thrown, are lexicalized and put in their proper places in the grammatical structure. Further, the words "me" and "now" leave no room for doubt about exactly what aspects of context are part of the sign. Similarly, if someone says "It is raining here," reference to place is lexicalized, and if the child says "Where's my mommy?" reference to the child is lexicalized. Further, in contrast to cases in which lexicalization involves a demonstrative, there is no freedom concerning exactly what the words "here," and "my" are to stand in for. "Here" has to stand in for the present place and "my" has to stand in for the speaker. Thus words like "I" and "you," "here," and "now" are indexicals. For example, the speaker of "I" represents himself or herself reflexively but as fitted in the semantical place held in the narrow linguistic sign by the word "I."

Past tense and future tense are both grammaticalized and hence indexical. If there were no past-tense or future-tense forms with which to contrast present-tense forms, present tense would not be grammaticalized. The bee dance, for example, tells when there is nectar—namely, now or today—as well as where there is nectar, but because Beemese has no contrasting forms indicating past or future affairs, its tense is not grammaticalized. Similarly, it is not lexicalized that it talks about nectar and not peanut butter. The bee dance does not contain indexical elements. Animal's signals to their conspecifics invariably concern the present, but this is not grammaticalized. Very likely there

was a time in the ancient history of human languages when the present tense was not grammaticalized. Now, however, it is.

Levinson remarks in his classic text *Pragmatics* (1983) that the "coordinate switching" of indexicals, say, from denoting the speaker to denoting the hearer to denoting some third person and so forth, "makes the acquisition of deictic terms seem a miracle" (p. 64). I have tried to show that the immediate precursors of indexical forms are among the most primitive of signs, and that we never depart very far from these primitives. A conventional language sign is merely a piece of the world, just as a natural sign is. It is different only in that it juts out from the rest of the world as designed to be a sign or part of a sign, and hence as requiring attention and interpretation. Adding lexical items that hold places showing where extralinguistic context is part of a sign surely simplifies matters for the interpreter rather than complicating matters.

I have argued that, when indicated by conventional forms of demonstration, the referents of demonstratives are reflexive signs that stand for themselves conventionally. Now the little word "the" with which definite descriptions are paradigmatically prefaced is etymologically derived from a demonstrative, the Old English masculine singular form "se." Suppose we inquire then into the relation between a demonstrative form and a definite description. Demonstratives often are accompanied by descriptions ("that book over there with the blue cover"), of course, as well as accompanied by indicating gestures. The role of the description in these forms is fairly obvious. In grammatically modifying the demonstrative, it functions as conventional means helping to direct the hearer's attention to the individual item described, which, once found by the hearer, stands for itself. Now compare "*that* book over there with the blue cover" with "*the* book over there with the blue cover," placed in the same context. These two do not seem to differ in meaning or way of functioning. At most, the use of the demonstrative rather than the definite article may hint that the hearer should look for a somewhat obvious form of conventional or intentional demonstration, a pointed look or a waved hand or some such. But there are many contexts in which demonstratives and the definite article are entirely interchangeable.

In other contexts, however, it may seem obvious that the referent of a definite description does not help to stand for itself. In "The first man on the moon," for example, the referent is not around to be examined by the hearer; indeed, it is unlikely either hearer or speaker was ever

acquainted with him. So the idea that "the" grew out of a demonstrative, conjoined with the idea that demonstratives require a referent that helps to stand for itself, seems puzzling.

Notice, however, that there are contexts in which demonstratives, like definite descriptions, refer to things that are not in the currently perceived context. Suppose I have just picked up a book and taken it in the next room to look at. I may call out "Have you ever met the author of this book?" expecting, of course, that you will know which book I am demonstrating—looking at, holding in my hand. You cannot see my conventional or intentional demonstration but you know what it is demonstrating. You do not see the book as I call, but you *know* which book I am looking at and that is enough. The book does not have to be currently observed by you in order to help stand for itself any more than I have to be currently looking at the map key in order for the word "hostel" to help determine the meaning of the x on the map.

But if something not currently observed but merely known about can stand for itself, then the puzzle about definite descriptions is easily resolved. Consider a fairly simple transformation of Russell's claim that definite descriptions claim uniqueness for what they describe. Definite descriptions are used when speakers take their descriptions to describe one and only one in the domain on which they are focusing, and either expect the hearer already to know this about the description or at least not to be surprised by it. That one thing is to stand for itself. The domains speakers are focusing on are tracked by hearers in the ways described in chapter 10. Hearers may or may not be concerned that their hearers know what this one thing is independently of the description; that is, their usage may be either referential or attributive in Donellan's sense (chapter 4).

Notice, for comparison, that demonstratives also are sometimes used attributively. Suppose, for example, that I am looking through a pile of old photographs in the next room taken by your grandfather and I call out "This must have been taken in Denmark." You reply "Very likely. Grandad spent most of 1945 in Denmark." Here I may have no expectation or concern that you will know just which photograph I am talking about. It is enough for my purposes that you know it is whatever one in the pile I am looking at.

IV Inner Intentional Signs

13 Inner Pushmi-pullyus

The remainder of this book is about inner intentional signs or "inner representations." The central question that I want to address is how and why, during the evolution of perception and cognition, organisms have acquired inner representations that are more sophisticated than pushmi-pullyu signs. How and why did perception-action cycles, which seem fully to characterize the cognitive character of the simplest animals, slowly give way to or become supplemented with more articulate and differentiated representations such as human beliefs, which are merely descriptive, and human desires, which are merely directive? Only quite speculative answers are possible here, but evidence from contemporary neurology and experimental psychology allow us to discern a few of the turning points pretty clearly, and may help us to tell a good story about some of the rest.

Much the most basic kind of intentional signs are the ones I called "pushmi-pullyu" signs in chapter 6. I will often call these "P-P signs" or just "P-Ps." P-Ps are signs that are undifferentiated between presenting facts and directing activities appropriate to those facts. They represent facts and give directions or represent goals, both at once. As mentioned in chapter 8, we find some P-Ps even in public languages: "No Johnny, we don't eat peas with our fingers!" So far as I know, all intentional signals used between nonhuman animals are P-Ps. On the one hand rabbit danger-thumps mean rabbit danger, but on the other they direct nearby relatives to take cover. Bee dances tell where the nectar is and at the same time tell where the watching bees are to go. The famous leopard, snake, and flying predator calls of the vervet monkeys both tell what kind of predator has been spotted and, simultaneously, direct behaviors appropriate to avoiding that kind of predator. Human smiles and frowns that are not yet overlaid with conscious intentions are simple P-Ps, telling that something potentially

rewarding has just been done and to keep doing it or do it again, or telling that something potentially damaging has just been done and to stop doing it or not to do it again. Similarly, the snarls of animals, or their tail lashings or waggings, their mating displays, their signaled invitations to play (the dog's "play bow"), and so forth, are all P-Ps.

P-Ps are also much the most common intentional signs occurring *inside* organisms. The bottom-most level of inner P-P signs is ubiquitously exemplified, not merely in neural matter, but in the many chemical messengers found in the body tissues and circulatory systems of animals. These are signals that are secreted, sent out, by one part of the animal's body to other parts, usually telling of the condition of the one part, and telling other parts how to respond. These P-Ps are the basic regulators of bodily homeostasis, coordinating the contributions of cells and organs so as to effect routine maintenance of bodily integrity. To suggest that genuine intentionality, genuine aboutness, with the possibility of misrepresentation, actually occurs at this level may at first seem far-fetched. But the idea is that there is intentionality here in the sort of way that zero is a number. These are the most humble sorts of limiting cases of intentionality. By treating such simple signals as intentional signs, just as by treating zero as a number, we will be able to examine their relations to various successors, and see the continuity between them and their more sophisticated relatives.

Simple reflexes, such as the reflex that withdraws the hand from something unexpectedly hot, are mediated by P-Ps. The neural signal that reaches the spinal cord tells what part of the body is exposed to something too hot and directs withdrawal of that part. The numerous neural mechanisms that work by negative feedback, such as the internal mechanisms that control tropistic behaviors in primitive animals, and portions of the mechanisms that control walking behavior in insects and, indeed, also in mammals, employ P-Ps. Negative feedback is a representation of the discrepancy between the value of a perceived variable and a set target value for that variable, telling what the discrepancy is and directly controlling the strength of the response needed to correct it. The primitive ability to follow a temperature gradient or a light gradient is mediated through P-Ps, as is the ability to follow a moving target with the eyes.

Moving higher up in the nervous system, the instinctive fear of snakes and heights that is built into many mammals including some humans (all babies instinctively shrink from precipices) are inner P-Ps that, when working as designed, are perceptions on the one hand and

directives on the other. Inner states that serve as reinforcers, either positive or negative, are P-Ps. Sweet tastes tell of nutritive value on one hand and direct continued eating or seeking more on the other, although, of course, there may be other systems capable of producing states whose functions are to override these P-Ps. Pain tells of damage to body tissues and directs present and future pain-avoidance, hence damage-avoidance, behaviors. That is how sweet tastes and pain work when they work as designed by natural selection. Similarly, perceptions of smiles and frowns, certainly in infants, are P-Ps.

According to J. J. Gibson and contemporary ecological psychologists (Gibson 1969, 1977, 1979; Michaels and Carello 1981; Reed 1982, 1993, 1996), basic perception consists in "picking up" or extracting certain abstract patterns in the ambient energies arriving at the organism's sensory surfaces, which patterns then guide various activities of the organism directly. No inference or calculation is required, but merely sensitivity to certain variants and invariants in the energies impinging on the active organism that, on the one hand, carry information about the relations of significant distal affairs to it and, on the other, directly guide its motions to take account or make use of these distal affairs. Basic perception is thus interpreted as perception of what Gibson called "affordances." Affordances are aspects of the environment that afford the possibility of various activities for the animal, such as walking on, climbing up on, going through or into, chasing or fleeing from (prey or predators), ducking away from (approaching objects), throwing, and so forth.

Gibsonians have generally assumed that if there were such things as inner representations they would have to be things calculated over, vehicles of inference, and hence, that the perception of affordances does not involve inner representations. But inner processes mediating the perception of and responses to Gibsonian affordances would certainly involve P-P representations, these being far more primitive than the representations Gibsonians reject. Information "picked up" by the organism would have to involve alterations to its inner states, which would be intentional signs of environmental affairs, these signs in turn guiding the organism's responses. Further, Gibson's claim that basic perception is perception of affordances is separable from his claim that perception is "direct." Basic perception might involve inference and still be perception of affordances. But in any event, chapter 9 argued that sensible use of the notion "direct perception" will include perceptions derived by translation from inner representations of prior

external aspects, such as edges, corners, oriented surfaces, and so forth. These translations should not be assimilated to inferences. Any response directly guided by a perception of relative size, distance, shape for picking up, angle of incline for climbing, momentum for throwing, and so forth, where dimensions of what is perceived directly guide dimensions of the response, would seem to involve inner P-P representations, whether or not the perceptual representation was derived, as Gibson proposed, as a direct function of invariances in ambient energy inputs. Any such perception could surely be considered perception of an affordance. It should be sufficient that there was a direct mapping between perceived variations in the environment and directly guided variations in behavioral response. At least that is the way I propose to use the phrase "perception of an affordance."

All in all we can conclude, I believe, that P-Ps occur on many different levels within organisms and that they vary greatly in sophistication.

P-Ps may be more articulated or less articulated. Some are merely intentional signals, their only significant variables being time or place. But even the pull of the magnetosome in the bacterium (chapters 3 and 6) is a bit more articulate than that, for it indicates also a direction, which can vary over a solid angle, directing the movement of the bacterium to correspond. The bee dance has variables indicating not only time and place but also direction and distance of nectar, hence direction and distance of the place the bees are to go. Perhaps the neural representation produced in the bee watching the dance is similarly articulated. The image projected onto the retina of the male hoverfly by its eye lens that causes it to dart toward and intercept an approaching female is a P-P (Millikan 1990). It guides the direction of the male's flight according to the direction of the female's flight, determined not as a linear function but as a certain trigonometric function of the retinal pattern she causes (to be exact, 180 degrees away from the target minus 1/10 the vector angular velocity measured in degrees per second of the target's image across the male's retina [Collet and Land 1978]). There can be no variable articulation of the pushmi face of a P-P sign (the descriptive side), of course, unless this articulation directs coordinate variation on the pullyu face (the directive side). P-P representations are intentional signs; their significant variables are the ones used to guide their interpreting mechanisms.

P-P representations can be highly abstract. In contrast, the empiricist tradition in philosophy has quite consistently maintained that what is

originally presented to the senses or in perception is concrete. A classical problem for the empiricists was then to explain how mental abstraction was accomplished, so as to separate out representations of kinds and properties from representations of their concrete instances. Given the description of intentional signs used here, this classic problem does not arise. An intentional representation is produced by a system designed to cooperate with an interpreting system in turn designed to use that representation in specific ways. The intentional content of the intentional sign is restricted to what the interpreter can read, that is, can make use of. No matter how rich and nested the *natural* information carried by an intentional representation is, only the part designed for use by the interpreting or consuming part of the cooperative system is represented intentionally (chapter 6). Let me quote here a philosopher who, though he does not espouse any form of teleosemantics, makes this particular point very clearly (with a little help from John Locke):

> The doctrine that picturelike representations won't do for general or adult or primate concepts involves a conceptual error. . . . Obviously you can't tell how a certain representation functions by confining your attention to the representation alone, or its "resemblances" to things in the world. You must know how the processors that act on it treat it. Thus a pictorial representation can express quite an abstract property, so long as the processors that act on it ignore the right specifications. To take a venerable example, a picture of an equilateral triangle can serve to represent triangles in general so long as the processors that act on it ignore the equality of the sides and angles. Similarly, a picture of a set of twins could represent or express the concept of a pair whose members are identical. (Block 1986)

An interesting result is that even very primitive P-Ps can represent very abstractly. For example, the statolith in the statoreceptor of the fish represents just one very abstract relation, namely, which way is down, hence which way to move to remain right side up. On the other hand, animals that recognize their individual conspecifics by smell also employ simple perceptual means, but in this case the intentional information represented is entirely concrete, for example, it might say "here's Mama." In general, there is no correlation between proximity to sensory input and the abstractness or concreteness of an inner representation.

P-P representations can represent either proximal or distal affairs. For example, even extremely primitive P-Ps, such as the magnetosome that represents the direction of lesser oxygen, can represent quite distal

affairs. Perceptual P-Ps represent things whose placement relative to the animal matters to the animal, things that it needs to take account of directly in action. What the pushmi or descriptive face of such a P-P represents is whatever environmental conditions it needs to vary with in order to guide its consumers properly.[1] Sometimes these conditions are absolutely proximal. Whether the skin is rapidly increasing in temperature may make quite a lot of difference to an animal, which is why our heat and cold receptors are designed to perceive just this very proximal affair and not, for example, the objective temperatures of the objects that touch the skin (Akins 1996). Pain and bad tastes also represent affairs absolutely proximal to the animal. But if the energies impinging on the organism, say, the kind and pattern of light or sound impinging on it, makes no particular difference to its well-being when within normal intensity ranges, these proximal patterns of light or sound will not themselves be perceived. Rather, any P-P signs derived from these energy patterns will concern more distal matters about which these patterns carry natural information.

Nor is there, in general, a definite distance at which a given sensory modality, such as sight or hearing, is designed to perceive. Contrary to much of the philosophical tradition, there is no single level of the outer world, such as physical objects versus the mere surfaces of physical objects, or such as the presence of certain phyical objects or of events versus mere sounds, of which the eyes or the ears are designed exclusively to produce direct representations. Depending on the animal's needs, various levels of distality of direct perception may be mediated by the same sensory end organs. The affairs naturally signified by retinal patterns, vibrating ear drums, stimulated odor sensors, and so forth, are at various distances and mediated in diverse ways.

Whatever affair a sight or a sound, or a scent is a natural sign of, that sight, sound, or scent can, in principle, be used to produce an intentional sign of that affair directly, without any intervening intentional signs (chapters 5, 6, and 9). Simple examples of this may be the intentional neural signs that mediate between many environmental signs that are "behavior releasers" and the "fixed action patterns" thereby released in many animals (Lorenz and Tinbergen 1939; Tinbergen 1951; McFarland 1987, pp. 1990ff.; Gould 1982). The proximal stimulus that

1. Granted, that is, that this correspondence is one the representation producer has a means of bringing about through its normal mechanisms of operation (see chapter 5, and also 6 on fertile female hoverfies).

guides the mother bird to drop food into the baby bird's open mouth immediately creates a neural pushmi that represents not a patch of red but a hungry baby's mouth, for only if the patch of red is indeed a local sign of a hungry baby's mouth will the behavior thereby released serve its purpose. The pushmi face of the neuronal P-P caused by a certain kind of dark shadow crossing the retina of a male hoverfly that causes him to fly off in a certain direction represents a female hoverfly, not a black moving thing or pattern of light or a moving image on the retina. In each of these cases it is likely that the representation is formed quite directly from retinal stimulations without passing through intermediate stages of representation. On the other hand, a sight or a sound or a scent may produce a direct perception of a distal object by passing through intermediate stages involving translation from prior representations ("direct perception" in the sense defined in chapter 9). In this case it is even clearer that there can be direct perception mediated by the same sensory end organ at various levels of distality, and hence that pushmi-pullyu representations can represent affordances at various levels as well.

P-Ps can represent affairs that are distal in time as well as distal in space. Probably few if any animals besides humans have developed uses for representations of past affairs, but many have need to represent future affairs. In chapters 3 and 4, local signs of such things as coming rain or approaching winter were discussed. Quite simple animals may need to translate natural signs of such future events into inner P-Ps that stimulate preparatory behaviors. Similarly, no matter how simple, most animals have an obvious need to recognize signs of approaching predators and to translate these into appropriate behaviors. There is nothing the least bit exotic about the production of inner representations of affairs distal in time any more than representations of affairs distal in space. Seeing into the future is exactly like seeing into the distance. The animal for whom a frosty night or the low angle of the sun serves to release winter preparation behaviors is being governed by inner P-Ps whose pushmi faces say that winter is on the way and whose pullyu faces direct what to do about it. For only if winter really is on the way will the behaviors that result serve the functions for which they were selected, and only if the P-P succeeds in producing these behaviors will it fulfill through its normal mechanisms the function for which it was selected.

Inner P-P representations may or may not require to be joined together with other P-P representations in order to do their work. The

P-P neural impulse produced in the frog's optic nerve by a passing fly reports when and at what angle the fly passes and provokes a corresponding response from the frog's tongue. This impulse forms part of a simple reflex arc that cannot be inhibited, even if the frog is completely sated. It reports a fact and issues an unconditional command. Similarly, during the first few days of its life, a rat pup whose snout comes in contact with a saliva-coated nipple grasps the nipple and continues to suck whether or not it is hungry. A few days later, however, this response is inhibited unless the pup is hungry (Hall, Cramer, and Blass 1975, 1977). Thinking of this in intentional terms, the pup's system is now sensitive to a new P-P signal indicating a current state of nutritional depletion and directing a response. The response is potentiation of the grasping and sucking reflex. The hunger signal says (roughly: see chapter 7), "Nutrients are depleted; if there is a nipple handy, suck on it!" Similarly, many small animals instinctively take cover if they see a small shadow gliding over the ground, such as would be cast by a flying predator. The shadow produces a P-P that means "predator overhead; if a cover-taking affordance appears, exploit it."

There seems no reason to suppose that affordances irrelevant to current needs are always, or even ever, perceived by most animals. There are, after all, lots of things we humans are capable of perceiving but generally don't perceive unless currently interested. Why would an animal feverishly translate every readable natural sign it encounters immediately into perceptions? Much of the currently perceptible world stays right where it is to explore perceptually later should it then become relevant. There is no need to think of simple animals as perceiving everywhere about them *mere* possibilities for action. More likely they perceive only what they have motivation, at the moment, to exploit.

To be in a position such that a primary goal, such as having a fly in the stomach, can be achieved by utilizing just one perceived affordance, such as a fly currently passing within reach of the tongue, is a *blissful* condition. Call such an affordance a "B-affordance." Call a negative condition that threatens immediate *disaster* a "D-condition." D-conditions are perceived by primitive animals as *negative* affordances, directing immediate avoidance or escape techniques. Call these "ND-affordances." (Similarly, so-called negative reinforcement is reinforcement of behavior that has afforded avoidance of or escape from negative consequences.) Organisms such as Venus flytraps and sea

anemones that do not move about may merely wait for B-affordances to pass by and then seize the moment. Similarly for ND-affordances. More sophisticated organisms make an effort to maneuver themselves into B-affordance conditions. The simplest way may be just to wander about directionlessly hoping to bump into one. This seems to be what clams do, for example. Other animals use more systematic techniques. The newborn baby's response to a touch on the cheek is to turn toward it, thus raising the probability of feeling a nipple on the mouth, which will afford immediate nourishment. Very simple animals show various kinds of taxis likely to take them into conditions where B-affordances are more likely to be encountered and D-conditions less likely. The frog recognizes places to approach and sit that are likely to attract flies. Perhaps it also recognizes places to avoid that are likely to attract snakes. One way to view the story of the evolution of perception and cognition is as a story about the acquisition of more and more sophisticated search techniques for maneuvering oneself into B-affordance conditions while staying out of D-conditions. Call an animal all of whose search techniques exploit only chains of perceived affordances so that its behaviors are entirely governed by inner P-P representations a "pushmi-pullyu animal." The central question that I wish to address in these last chapters concerns what the disadvantages of being a purely pushmi-pullyu animal might be and what remedies for these disadvantages may have been supplied during evolution of the higher species.

The cardinal principle involved for any pushmi-pullyu animal in raising the probability of encountering B-affordances is very elementary: Be constructed such that you can perceive affordances that will afford your probable placement in new positions from which you are likely to perceive new affordances that will afford your probable placement in newer positions from which . . . and so forth . . . finally placing you in B-affordance conditions. The trick is that this series of probabilities should have a product greater than the probability of B-affordances just happening along without any action on your part, the higher the probability the better. Thus the search domain is narrowed and then narrowed again.

Ecological psychologists speak of "perception-action cycles" during which input from the environment is said directly to guide motor output without additional input from the central nervous system. Perception produces action that results in new perception producing further action, and so forth. Robots that work entirely on the

perception-action principle have been constructed that perform simple tasks, for example, following walls, avoiding obstacles, picking up soda cans, recharging their batteries at the right times, and so forth (Brooks 1999). The various cycles that govern the behaviors of these robots are arranged so that the activation of some cycles will inhibit the activation of others, thus establishing an order of importance in activities.[2] The activities of insects may be largely or entirely governed in this way, by hierarchies of perception-action chains, or as ethologists call them, chains of "behavior releasers." Thus the digger wasp walks randomly this way and that until it encounters certain signs of a prey; the prey affords stinging and hence paralyzing, which affords being dragged to the entrance of the wasp's nest, which affords entering and circling and, if all is well, affords emerging again and dragging the prey within, and so forth. The cycle will be interrupted, of course, if at any point the wasp encounters ND-affordances, signs of danger.

In more complex animals, motivations such as hunger or fear may potentiate perception not just of one but of any of numerous alternative affordances, calling on a large set of alternative behaviors to be used contingent on the animal's situation. Extremely complicated long and branching chains of affordances leading to the probability of finding one or another further affordances, leading to the probability of finding one or another . . . and so forth, may be grasped by some animals, resulting in highly flexible behaviors. And it may be that correctly quantified increases and decreases in potentiations of response dispositions that change the ease with which these can be activated, produced by other relevant stimuli encountered along the way, help account for the tendency of the animal to perceive, from among equally available and relevant affordances, those objectively associated, in the animal's particular circumstances, with higher probabilities of eventual success. The result would be an animal whose behavior is highly flexibly governed by what Gallistel (1980) calls a "lattice-hierarchy."

Gallistel points out that an animal whose behavior is completely governed by a lattice-hierarchy may have acquired that lattice in large part

2. Because there is nothing in these robots that corresponds to a central processor or central nervous system orchestrating the whole by performing inferences, it is commonly assumed that these robots do not employ inner representations at all. But they do employ inner P-Ps in the sense defined here. These robots are not, of course, direct products of natural selection. But they are designed to function as they do by people who are themselves designed by natural selection and by learning processes that natural selection has designed capacities for. The robots have proper functions in the way all designed human artifacts do (chapter 1; Millikan 1984, chapter 2. For more details, see Millikan 2002).

by learning. Suppose that instrumental (operant) conditioning works in the way that classical American behaviorists claimed. The analogy with natural selection is quite strict. Responses to stimuli are strengthened whenever they are followed by "positive reinforcement" or reward and weakened when followed by punishment. Primary reinforcements, those that are not learned, are associated with B-affordances or ND-affordances. Reinforcing a behavioral response to a stimulation thus conditions the animal to perceive an intermediate affordance leading either toward a B-affordance or away from D-conditions. Perceptions of intermediate affordances then become "secondary reinforcers"—say, the introduction into the cage of a bar-for-pressing when a trained rat is hungry, or the lighting of a light that signifies that bar-pressing at this time will now afford food. Thus the animal can be trained to perceive affordances that afford further affordances in quite a long chain.[3] But this instrumental conditioning merely designs another lattice-hierarchy, a lattice-hierarchy constructed during ontogeny rather than phylogeny. The result might still be merely a pushmi-pullyu animal.

It is clear, then, that a pushmi-pullyu animal might be capable of navigating in the space-time-causal order from a great variety of starting positions relative to its goals so as to reach them with high probability. On the other hand, such an animal might also be subject to failures that strike us as rather ridiculous. Gallistel notes that even in quite flexible animals, available behaviors are by no means always chained so as to apply to relevant situations, even when the increment is very small. As an example, he cites Dilger's (1960, 1962) work describing a hybrid species of lovebirds that, although quite capable of safely carrying the strips of bark used for weaving their nests in their beaks, nearly always carried these strips by tucking them into their tail feathers, losing most of them on the flight back to the nest. The tucking behavior was a left-over from ancestor species that lined their nests with small chips, which more easily stay put in the tail feathers. Although these lovebirds were capable of perceiving and acting on each relevant affordance, these capacities were not chained in an efficient way (Gallistel 1980, pp.

3. According to Anthony Dickenson, "perhaps the most bizarre demonstration of the power of positive reinforcement comes from a legendary laboratory rat who was prepared to climb to the top of a spiral staircase, 'bow' to the audience, push down and cross a drawbridge, climb a ladder, use a chain to pull a model railway car, pedal the car through a tunnel, climb a flight of stairs, run through a tube, and descend in a lift, all to receive a single pellet of food" (McFarland 1987, p. 472).

306–308). A more homely example of this is the house cat that washes its ears by rubbing them with its paws and licking them off, and that may pull food scraps out of its dish with its paws to eat on the floor, but doesn't know to use its paw to clean out the yummies at the very bottom of an emptied can of cream of chicken soup. (An occasional cat does figure this out.)[4] Similarly, a purely pushmi-pullyu animal would lack the ability to recombine various segments of behaviors in its repertoire in new ways so as to achieve new goals. It could achieve new linkages of behavior chains only by reinforcement of accidental connections after the fact, never by inventively looking ahead.

One step toward inventive recombination of behaviors, I will argue, is the articulation of P-P representations that govern behaviors into segments that can be recombined to make novel P-P representations. I will discuss this kind of segmentation in chapter 14. Also, with the development of more complex articulation in P-P representations, some representations of pure facts become detached, producing "pushmis" that are ready for recombination with various alternative "pullyus." And some representations of facts are not only detached, but stored away for use on other occasions (chapter 15). The purely pushmi-pullyu animal, on the other hand, represents only facts that it already knows how to use, and represents them only in the context of their use. This means that it never has any extra information lying around, as it were, to employ in situations it has not already been genetically programmed, or trained by conditioning, to deal with. All of its facts are devoted to specific uses. Further, the purely pushmi-pullyu animal doesn't represent, and hence doesn't know about, anything it is not currently perceiving. It may have a memory for procedures but not for facts.

A correlate is that the purely pushmi-pullyu animal always represents affairs in its world as bearing certain relations to itself. An animal's action has, of course, to be initiated from the animal's own present location. To serve as an unmediated guide to immediate action, the descriptive face of an inner P-P representation has to represent the relation of the affording situation or object to the perceiving animal. This doesn't require that the animal represent itself explicitly, any more than the bee dance represents nectar, hive, and sun explicitly (chapter 7), but the self has to be represented at least implicitly. In the simplest

4. I don't mean to imply that the domestic cat is a purely pushmi-pullyu animal.

cases, the relevant relation may consist merely in the affording situation occurring in roughly the same location and at the same time as the animal's perception and consequent action. In less simple cases it will include more specific relations to affording objects, such as spatial relation to the animal, size relative to the animal's size, weight relative to the animal's weight or strength, and so forth. We humans, at the opposite extreme, are capable of forming beliefs not only about things and affairs very distant from us, but about things whose spatial and temporal relations to ourselves are completely unknown. Certainly, our descriptive representations do not, in general, represent relations of situations and objects merely to our current selves.

More striking than the failure to recombine behaviors in relevant ways is another possible failure of the pushmi-pullyu animal which was described by Lorenz and Tinbergen (1938). Although the greylag goose apparently reacts intelligently to an egg that has rolled out of the nest by bringing the bill behind the egg and rolling it back into the nest, if the egg slips sideways out of control of the bill, this movement may still be carefully completed, "as if it were a vacuum activity" (Tinbergen 1951). It seems that the goose does not understand the purpose of its own behavior. Dennett (1984) has popularized the example of sphex, the digger wasp, that can be sent into a behavioral loop from which it never emerges by removing its paralyzed prey a few inches away from the door of its nest every time it goes inside to inspect, preparatory to dragging the prey inside. The wasp seems not to understand the purpose of its own activity so as to know when that purpose has been accomplished. I once watched a pair of hamsters repeatedly stumbling over one another as each returned a large cracker to its own corner again and again from the other one's corner just opposite. Neither seemed to notice that its own corner continued to remain empty.

This sort of failure is exactly what we should expect of a purely pushmi-pullyu animal, for such an animal does not represent its goals in a format that enables it to know whether or when it has reached them. The language in which the directive side of a P-P representation is expressed is not the same as the language in which the descriptive side is expressed, so that when acting on the directive produces or fails to produce a corresponding state of affairs, this is not automatically recognizable to the animal. The purely pushmi-pullyu animal does not, as it were, *project* its goals. Its behaviors are controlled completely from

behind by emerging environmental contingencies. It does not represent its goals as purposed future occurrences or states to which actual accomplishments will be compared. Chapter 16 will concern the transition from purely pushmi-pullyu animals to animals that represent their goals in the same representational system in which they represent their facts.

14 Detaching Representations of Objects

Beginning with minimally articulate P-P representations, the evolution of inner representations seems likely to have paralleled evolution writ large. First, representations have become more articulate, so that more and more of what they represent is represented explicitly. More complex functions are then built up of out of more specialized functions of the articulated parts. Then ways to perfect these more specialized functions somewhat independently have developed, sometimes by the development of new generate and test procedures. These articulated specialized functions are then recombined and reintegrated in new ways. The general strategy involved—disassemble, tune the parts separately and recombine—is typical of evolutionary developments more generally. We encountered it before in chapter 2 when discussing the evolution of evolvability. The following chapters outline some of the results of this strategy during the evolution of inner representations, keeping a special eye out for developments that begin to separate the pushmi from the pullyu sides of representations, that is, the descriptive from the directive.

Primitive behavior releasers are often activated by quite crude patterns of proximal stimulation. Anything that is red and gaping of about the right size on the retina will serve as a stimulus to drop in a worm for a parent bird at the edge of its nest or for a bird engaged in courtship feeding. Jackdaws that had imprinted on Konrad Lorenz as if a member of their own species tried to drop worms in his ears (Lorenz 1952). Imprinting itself tends to occur on any object at all that is moving about near the animal during the earliest moments of its life. Anything decorated with two blobs above and a sort of horizontal line below will attract the gaze of a human infant in the first weeks of life, thus focusing its attention on human faces, which are important things for it to study. The greylag goose will try to pull into its nest just about any

proximate object with rounded contours that is not bigger than the goose itself. If other objects similar to the targeted object are unlikely to appear in the contexts in which perception of a certain affordance is potentiated, these crude perceptions of affordances may serve quite well. But for other tasks, more discriminating methods are needed.

The central problem for an animal that needs to discriminate more carefully among various affording affairs involving distal objects is that the same kind of distal affair is likely to have myriad alternative effects on the sensory surfaces of the animal, depending both on its spatial relation to the animal and on mediating or intervening conditions such as lighting conditions, atmospheric conditions, sound absorption and reflectance properties of surrounding objects, obscuring conditions such as intervening objects, masking sounds and odors, and so forth. This doesn't always matter. If all the animal needs to know is that winter will be along within a month or so, then there may be quite a number of telltale proximal stimulations that it is likely to encounter sometime within the necessary time frame, such as frosty nights, certain smells, the look of the sun appearing lower in the sky, and so forth. But most distal affairs that matter to an animal need to be registered more immediately. This means that the animal needs to be able to recognize the affording affair from a wide variety of perspectives and under a wide variety of conditions. To be as useful as possible, the apparatus that recognizes the affording distal object or situation must recognize it over as wide a range of relations to the animal as possible, near, far, overhead, underfoot, left, right, partly occluded, and under a variety of mediating conditions such as lighting conditions, sound transfer conditions, when the animal is moving or still, when the object is moving or still, despite "static" such as fog, wind, and other extraneous noises, dappled shadows, other entwined smells, and so forth.

Similarly, affording distal affairs typically need to be registered not merely as somewhere within the animal's vicinity but as currently bearing some quite definite relation to the animal, for the appropriate reaction on the part of the animal will vary as a function of this relation. For example, the animal needs to know not just that a certain predator or prey is present but in what direction and at what distance. Thus the animal may also need to command a variety of ways to recognize the same relation to the affording object or affair, depending on conditions. It may need to hear and feel direction as well as to see direction, for example, and it may need to employ a variety of alternative

means of depth and distance perception, using several sensory modalities as well as several methods within in a single modality.

This kind of need is graphically illustrated by a story that has been circulating fairly widely, and though it is probably apocryphal, it makes its point very well.[1] The story is that certain venomous snakes perceive mice for purposes of striking by sight, trace the path of the dying mouse by smell, and find its head so as to swallow that part first by feel, and that none of these jobs can be accomplished using any other sensory modality. A snake that was wired up this way would merely perceive first a "strike me," then a "chase me," and finally a "swallow me," having no grasp at all that what it struck, followed, and swallowed was the same thing. The story is probably apocryphal, but it has a certain plausibility. Both objects themselves and their perceived relations to oneself show up differently through different sensory modalities. Consolidating one's inner representational system so that the same object and the same relation to that object are always recognized *as* being the same is not a trivial task. But surely it would be more efficient for the snake to be able to use each of its various ways of recognizing a mouse in some relation to itself for whatever purposes it might have for a mouse at the moment. Moreover, it would be good if its ways of perceiving its relation to the mouse when striking, following, and swallowing could be detached for use in perception of similar relations to other kinds of affording objects.

In a similar vein, recall the connectionist face-recognizer, VisNet, from chapter 9 that recognizes seven individual faces from each of nine angles but would have to start all over again to learn to recognize an eighth face. Clearly this sort of holistic approach to face recognition is not very efficient. If you have need to recognize many faces, especially if you need to learn to recognize new faces quickly, it is more efficient if you first put in place the more abstract but multipurpose capacity to recognize any shape, as such, from any arbitrary angle and at any arbitrary distance. It is better if you have already in place the capacity that psychologists call "shape constancy." Generalizing this, suppose that a great many different objects need to be recognized by an organism, each under numerous different conditions. It will be efficient to divide or disassemble the function of the organism's object-recognizers into

1. The original source seems to be the Swedish zoologist Sverre Solander, who gives no references, however, and despite requests from myself and others, has offered no data yet, to my knowledge.

parts or stages, each of which recognizes members of a certain class of properties such as shape, color, size, texture, quality and direction of movement, identity of sound at origin (e.g., voice constancy), solidity or malleability, weight or angular inertia, and so forth, under many different conditions. Then these various recognition capacities can be combined as necessary, using each over again in as many new configurations as possible. That is, you build in a set of prior apparatuses that register simple objective physical properties and relations as constant through changing perspectives and intervening media. In different combinations, these properties are likely to be locally recurring signs of a wide variety of affording objects, kinds, stuffs, events, and other world affairs, each of which it will now be easy to learn to recognize.

Also, just as the same property such as shape, color, or texture may affect the sensory surfaces in different ways on different occasions, which of an object's various properties are currently observable to an animal also varies from occasion to occasion. It is best then if the animal can recognize affording objects by way of many different alternative sets of diagnostic properties. This is a theme that I have developed at length elsewhere,[2] however, and I will not say much more about it here.

Representations of properties used for the purpose of detecting affording objects are not dedicated to particular practical purposes settled in advance. They are not pushmi-pullyu representations. They tell of the disposition of properties among various objects in the environment without yet saying what is to be done as a consequence. They are steps on the way to perception of affordances, but they are not themselves perceptions of affordances. The same is true, of course, for any representations that are prior to representations of properties of objects, such as representations of lines or edges with a particular orientation, right angles, ocular disparity, directional movement, color edges, and so forth. These represent detached facts. So we have discovered one small way in which pushmi-pullyu representations may begin to come apart, namely, descriptive representations detaching from directive ones. It would be a mistake, however, to assume that use of these descriptive signs required the animal to have concepts of properties. The capacity to discriminate a property is not, as such, a concept of a property. (Many one-celled organisms discriminate dark from light; it does not follow that they employ any concepts.) The concept of a property, presumably, is the kind of thing that can play a

2. In Millikan (2000).

role in propositional judgment and mediate inference, and there is no reason to suppose that simple property-detectors are employed toward any such ends.

The identification of affording objects and situations is, however, only half of what is needed for guidance of effective action. Recognition of an affording object tells the animal what to retreat from or what to approach, what to pick up or to eat or to climb up on and so forth. But the animal will not perceive *how* to perform any of these maneuvers unless it also perceives its own relation to the affording object or situation. Perceiving that there is an apple that affords eating is one thing; perceiving exactly how I would need to reach from here to obtain it is another. These two perceptual aspects need to be combined in a single articulate pushmi-pullyu representation to guide the procurement and eating of a suitable object. Correlatively, two general-purpose skills must be combined here to make up a single-purpose skill. First is the general skill that allows perception of an apple and hence perception of an eating affordance, through any of a variety of media from any of many perspectives. Second is the set of skills that allows perception of the current relation of the perceiving animal to the affording object so as to guide activities of approaching, picking up, and conveying to the mouth. Skills of this latter sort can be practiced, of course, in many contexts other than that of apple eating. We have here another example of disassemble, tune the parts, and recombine.

The distinction between skills involved in identification of affording situations and objects versus skills involved in identification of relations of objects to the animal for purposes of interaction or manipulation seems to have relatively clear neurological correlates. Vision has been the most closely studied of our perceptual capacities, and current neurological data on vision indicate these different aspects of perception quite clearly.[3] As mentioned in chapter 9, the origin of vision involves, first, the translation of gradients of luminance across larger or smaller areas of the retina to detect various rudiments of visual form (lines, edges, orientations, angles, movement, direction, etc.), which are later processed to yield information about properties of objects. But even at the level of the ganglion cells in the retina, a division already occurs between what are to become two relatively separate neural

3. Most of the following information on vision can be found in Jeannerod (1997) or in Norman (2002). The latter is a paper with a thesis, but it contains a balanced review of the literature on the two visual channels discussed below.

pathways of visual perceptual analysis, usually referred to as the *dorsal* and the *ventral* pathways or channels. Roughly speaking, the dorsal pathway is concerned with guidance of the organism's movements in relation to perceived objects, while the ventral pathway is concerned with identification of objects.

Ganglion cells that feed information into the dorsal channel process information from all over the retina including all peripheral areas, feed into channels that process information at high temporal frequencies, and help to produce special sensitivity to the larger patterns on the retina and special sensitivity to ocular disparity and to motion. As the object or the organism moves, ocular disparity and motion displacement are fundamental, for example, to the ability of the dorsal channel to detect direction, distance, angle, location, and size relative to the organism. In sum, the dorsal channel processes information relevant to performing motor movements in relation to perceived objects and events, walking or running toward or away from things, between or through them, climbing, pointing, reaching, grasping, and so forth: "[H]ow large should the gap between the thumb and forefinger be in order to pick up that block?" (Norman 2002, 3.4.6). It processes information about affordances at the level of immediate movement and contact with objects-in-general. Thus Jeannerod calls the representations processed by the dorsal channel "pragmatic representations" and says of them that they "refer to rapid transformation of sensory input into motor commands" (1997, p. 77). A crucial fact about these representations may be that they are always used only for this purpose. For example, they cannot be used as the basis of expressed perceptual judgments or discrimination tasks of other kinds. They cannot be "brought to consciousness."[4]

Those retinal ganglion cells that feed mainly into the ventral channel lead to slower processing of information but at higher spatial frequencies. The ventral channel processes information from more central areas of the retina and it yields more detailed form, pattern, and color analysis. It helps to produce special sensitivity to features necessary for object identification, such as more exact shape and size. It detects relations between objects better than relations to the animal. In general, it detects *what* objects the animal confronts so that it will know whether to go toward or away from them or what to do with them, separating this information from the animal's current accidental relations to these objects. Jeannerod (1997) calls representations processed by the ventral

4. For a contrary opinion, see Neisser (2002).

channel "semantic representations" and says of them that they "refer to the use of cognitive cues for generating actions." A dramatic illustration of the separation of these two systems is the experience of an experimental subject who is required, over a period of days, to adjust to wearing glasses that reverse right and left so that the world appears as a mirror image of itself. After a time he may be perfectly capable, for example, of riding a bicycle through traffic, everything seeming to him perfectly normal again, except that the license plate numbers on the cars are still backwards.

The dorsal and ventral visual channels are, of course, but two faces of one system. One's peculiar momentary relation to an object of interest needs to be canceled out in order to recognize the object, but of course it must be figured back in again if one is to manipulate the object or alter one's relation to it. "Thus, when one picks up a hammer, the control and monitoring of the actual movements is by the dorsal system but there also occurs intervention of the ventral system that recognizes the hammer as such and directs the movement towards picking up the hammer by the handle and not by the head" (Norman 2002, 3.4.8). A differentiation between these two channels is apparently very old, being found in the visual systems of mammals from hamsters through monkeys to humans. A similar division between two information channels is found within the auditory system as well.

The ventral channel is often called the "what" channel while the dorsal channel is called the "where" or the "how" channel. These names are not very helpful, I believe. The ventral channel may typically represent "what it's for" rather than merely "what," showing not detached facts, primarily, but affordances. Thus Gibson told us that we perceive apples as affording eating and postboxes as affording letter-mailing (1979, p. 139). This may well be so most of the time for people, and perhaps all of the time for most animals.[5] In the first instance, at

5. In a classic experiment, Sperling (1960) briefly presented human subjects with twelve letters arranged in three rows. They could usually report no more than four of the letters. But when prompted in advance to report on a specific row, they could report any row on demand even if the prompting cue was presented up to 250 milliseconds *after* the offset of the target stimulus. It seems that the basic information needed for the construction of all twelve rows was represented in early processing, but not brought to completion in full identification of the letters unless this information was specifically called for. In general, it is reasonable to suppose that there is a huge amount of natural information readily available to a human person through perception at any given time which that person is capable of translating into inner intentional representations, but that only a very small proportion immediately relevant to current concerns ever actually gets translated.

least, perception of objects immediately serves practical, not theoretical purposes. The dorsal system, on the other hand, is not alone in representing either where objects are or other attributes relevant to performing motor movements in relation to objects. "A large set of attributes are in fact relevant to both the semantic and the pragmatic processing: . . . shape, size, volume, compliance, texture, etc . . . An essential aspect of object-oriented behavior is therefore that the same object has to be simultaneously represented in multiple ways . . ." (Jeannerod 1997, pp. 78–79). Nor does the dorsal system merely represent "how." To direct how to move appropriately it has to represent *what* the relevant relations are between the animal and what it would act with respect to. The difficulty with all three of these designations, "what," "where," and "how," is that each fails to recognize the double aspect of the representations generated by these systems. Each either fails to recognize the pushmi or the pullyu aspect. For dorsal and ventral systems each produce P-P representations, though not complete ones. These representations need to be joined to each other to make, as it were, a complete P-P sentence. On the other hand, it is not very good to refer to these two kinds of representations just as "dorsal" and "ventral" either, for the important distinction to be drawn is one of function, the degree to which the anatomical division between dorsal and ventral channels in humans or other animals accurately reflects the distinction being, of course, a matter for empirical investigation. Indeed, it is known that the ventral and the dorsal systems are able to exchange some information when necessary, although less accurately and at a delay. I will suggest what may be a more perspicuous terminology below.

It is commonly claimed that dorsal-system signs are "viewer centered" or "body centered" or "egocentric" whereas ventral-system signs must be "object centered" or "allocentric." Often this difference is interpreted as a difference is the coordinate systems used for the two types of representation. In this connection it is worth noting that it is not a logical requirement for spatial representations that they employ any coordinate system at all. A model ship is a detailed spatial representation of a ship, but no coordinate system is employed in this representation. The sentences "New York is south of Boston," "New York is larger than Boston," and "New York is twenty miles from here" are all spatial representations that employ no coordinates. The important difference between the two kinds of representations, I suggest, has nothing to do with coordinate systems but is as follows.

The dorsal part of the full P-P sign or the part that represents the relation of the affording object to the animal needs, of course, to represent the perceiving animal itself, but this representation need only be implicit. These signs are used in the context of showing relations of situations and objects to the animal for immediate guidance of the animal's motions. We can call these relations "enabling relations." Obviously the animal can only act for itself, only move its own limbs and so forth. It is perfectly clear, then, that there will be no intentionally significant transformations of these signs showing other animals bearing these same enabling relations to these objects. These signs will contain no variable parts or aspects representing the animal itself. They will represent the animal implicitly in exactly the same way that the bee dance represents the hive, the sun, and the nectar implicitly, there being no transformations of it that talk about the big oak tree, the moon, or peanut butter (chapter 7). This sort of sign is "ego centered" only in that the ego is so central to it that it doesn't even need to be mentioned. Only the enabling relation needs to be mentioned. I will say that these signs represent "enabling relations" and that they are "ego-implicit" signs.

On the other hand, the part of the full P-P sign that represents the affording object, configuration, or state of affairs but without representing its relation to the animal as needed for action may or may not represent the animal itself, but if it does represent the animal, it will represent the animal explicitly. It will permit significant transformations yielding intentional representations of things other than the animal in place of the animal's self. For example, if this kind of sign can represent an apple being about so far (a yard, say) from me, it can represent an apple being about so far from you in exactly the same way.[6] This seems a good reason to call this second kind of sign "objective," for if such a sign represents the self it represents it as one object among other objects. But, of course, an animal need not be capable of explicitly representing itself at all. It may well be that most animals do not have the capacity to represent themselves as objects, so that they harbor only ego-implicit representations and egoless representations, never ego-explicit representations. None of their objective representations includes themself as an object among other objects.

6. This is a point about what can be represented, not an epistemological point. I am not saying that it is always as easy to tell how far one thing is from another as it is to tell how far it is from me.

The ego-implicit versus objective distinction is an important one for the animal that needs to learn many new behaviors, for it partitions these tasks into two aspects, each of which can be learned and practiced separately, then combined. On the one hand, there is the capacity to recognize the same individual or kind of object or kind of objective situation again. This rests on recognizing the same objective properties again, a skill that is practiced whenever objects sharing properties in the same range need to be identified. On the other hand, the animal can develop general skills for manipulating arbitrary objects and its relations to these objects. It can learn how to move among objects, climb up on them, jump from one to another, move them at will, grasp and pick them up, turn them over, throw them, and so forth. Young mammals often seem to be practicing such general skills while playing, with no more distant goals in focus. The playing cat doesn't care whether it is a mouse or a leaf or its tail that it chases. It is practicing chasing just *things*.

The task for the animal designed to learn a great deal by operant conditioning, which requires appropriate generalization and discrimination following successes, would be staggering if the animal were working with no prior knowledge of what variety of proximal stimulations might be signs of the same sort of distal object or affair, or of what variety of proximal responses might produce the same distal effects. Further, the attainment of objective representation may allow the animal to analyze its various activities into distinct achievement stages, each stage being recognized as what-the-objective-situation-is-now, apart from the animal's momentary position within that situation. Activities can then be understood as series of transitions from one objective situation into another objective situation, as stages in an objective process. There is evidence, for example, that it is by recognizing the objective completion stages of a process that animals sometimes learn from one another by imitation. Despite folklore about monkeys and apes who "ape" others, the evidence is that what is "aped" in nature is not bodily motions but a completed series of project stages (Byrne 1999, 2002). Taking a homely example, many cats will make a try at opening a door by reaching for the doorknob, even taking it between their two paws in an attempt to turn it. The human model whose results in action they are attempting to replicate does not use two paws, however, but one hand. A few animals can learn to imitate observed bodily motions; for example, surprisingly, dolphins can (Herman 2002, forthcoming). But the natural focus of animals' capaci-

ties to represent objective processes in the world is not on their bodies but on the objects they manipulate (Tomasello 2000; Rumbaugh et al. 2000; Whiten et al. forthcoming). Their bodies are represented only implicitly by the systems that must perceive relations of their bodies to other objects for guiding manipulation of these objects.

I have emphasized that perceptual representations that represent objects objectively, representing them apart from their momentary enabling relations to the acting animal, are not representations of facts but of partial affordances. Prey are perceived as for chasing, predators as for escaping from, and so forth. On the other hand, the same object may be perceived by an animal as having different affordances on different occasions, depending on the animal's current projects and needs. The ability to recognize water, for example, probably has a considerable variety of uses for most land animals. For the snake, the mouse affords striking, then following, then swallowing. If the snake were able to recognize the mouse for each of these purposes through any of its three sensory modalities, would this general ability to recognize a mouse be an ability to recognize the detached *fact* of the presence of a mouse? The kitten sees the mother cat as a source of food, a source of warmth, as protection, as a friend to play with, and so forth. Can it represent the detached fact of the presence of its mother? If learning how better to recognize an object for one purpose is carried over and applied to recognizing the same object for many other purposes, doesn't this amount to the ability to perceive a pure fact, detached from any practical use it might have?

Representations of pure fact are representations that are not dedicated to any particular purposes. They stand ready to be combined in the production of actions with purposes that have not been determined in advance, and perhaps with other factual knowledge that has not been determined in advance. The animal that represents a variety of objects and objective situations, recognizing different affordances of these at different times, certainly might be said to represent facts. But we should distinguish such an animal from one that can represent facts that it is not interested in exploiting at the moment, or that it doesn't yet know any uses for at all. In chapters 18 and 19, I will argue that this is quite an important distinction, perhaps one that helps to demarcate the peculiar intellectual capacities of humans from those of other species.

15 Space and Time

Before turning to the problem of how pure representations of goal states develop from pushmi-pullyu representations, there are two more common kinds of factual or semifactual representations—pushmi representations each of which has a variety of pullyu uses—that should be mentioned. A great variety of animals seem to be capable of collecting over time and storing representations of both the spatial layouts of the territories in which they live and also the temporal layouts, the patterns of conditional probabilities of ordered occurrences that characterize their environments.

Perception normally requires the integration of information that flows in to the sensory surfaces over a period of time. This is true even of organisms having the most humble of nervous systems if we take into account habituation and sensitization. Gibsonian psychologists have been especially helpful in highlighting the importance of the animal's own movements in gathering information during perception. In humans, eye saccades are especially important, along with movements of the head and body, movements of the hands when exploring by touch, approaching and moving around objects, and so forth. These exploratory motions have been called "epistemic actions." A great many perceptual illusions that can be induced when freedom of movement is denied dissolve as soon as the subject is allowed to move freely again. Epistemic activity is important for the perception of objects and objective situations. It is equally important, of course, for the perception of enabling relations (chapter 14).

Epistemic activity helps the animal to separate aspects of the surrounding energy array that vary with its own movements from those that are invariant, thus helping to detach perception of objects from perception of enabling relations. It helps the animal to take into account the particular media through which perception is occurring, for

example, to separate lighting conditions from properties of seen objects, and to separate the sources of various sounds reaching the ears. It also helps in the construction, over a short period of time, of perceptions that represent, for example, the whole shape of an object, not merely the part that is seen at a given instant. It is this last benefit of epistemic activity on which I wish now to focus.

In chapter 14, I mentioned that one of the abilities that is important for an object-recognizing animal to have is the capacity to identify the same object or kind by a variety of alternative diagnostic properties or sets of properties. This is because the animal's encounters with an object or kind will reveal certain of its properties under some perceptual conditions and from some perspectives, but other properties only under other conditions or from other perspectives. Like most people, Helen Keller was able to identify a very large number of different objects and kinds under the right conditions. Her handicap was in the scarcity, for her, of right conditions.

One way to recognize a shape is to recognize a given object and to remember what shape that object has. For example, suppose that by prior epistemic activity I have discovered and now recall what shape my flashlight has, the relative position of its various parts, and so forth. If I now recognize my flashlight in the dark by feeling one end of it, I may be able to represent the enabling spatial relations of its various other parts to my groping hand, thus allowing me to adjust my hand so as to find the switch with my thumb, at the same time pointing it to shine in the right direction. Similarly, if by prior epistemic activity I have discovered what lies inside apples, that is, in what spatial configuration their various inner parts lie, then I may need to see or feel only the outside of an apple to be enabled to begin neatly to quarter and core it. Knowledge of the inner configuration of oranges, on the other hand, directs quite different initial motions in the process of procuring the meat. Think too of performing these tasks in the dark.

Consider another example. Free-living young gray squirrels soon acquire a lot of experience with trees. Standing trees generally have a similar sort of abstract shape, with a trunk leaving the ground roughly vertically, then branching into limbs that in turn branch into branches and then twigs. I don't imagine that a squirrel encountering a tree trunk as it races from a dog chasing after needs to look up in order to recognize it has come to a tree. And it seems likely that its starting up the trunk is guided in part by a representation of an enabling relation to higher parts of the trunk and to the branches assumed to be above. Its

movements are guided by a representation of the trunk stretching upward, and of branches overhead in a certain enabling relation to it, and the safety these are envisioned to afford.

It is an empirical matter, of course, whether this is really what happens in squirrels, or whether the squirrel is merely following out a chain of affordances as described for the purely pushmi-pullyu animals of chapter 13. But the point is general. The purely pushmi-pullyu animal is certainly capable of learning, but this learning is what psychologists call "procedural learning." It learns what to do after what, completion of each link in the chain producing perception of a new affordance, which guides production of the next link. But here I have in mind the capacity for quite a different kind of learning. This kind of learning is sometimes called "declarative" because it is tested for in humans by asking them to state or declare what they have learned. There is no reason to suppose that memories of this sort are representations that could be translated or portrayed by sentences, however (chapter 7). On the other hand, this kind of memory does seem to be a memory for facts. Such a memory is a descriptive representation of an affair in the world that can be recombined in perceptions of alternative affordances. A representation of the abstract way trees are laid out may be a detached fact representation for squirrels, useful in connection with chasing, fleeing, finding food, finding shelter, and so forth. The animal stores a representation of the spatial configuration constituting a certain kind of whole object, then later recognizes its current relation to some part of the object. Joining these two yields a representation of current affordances of unobserved parts and of their currently enabling relations.

Perceiving and remembering the spatial configuration of something you are outside of is continuous with perceiving and remembering the spatial configuration of something you are inside of, such as a sheltering structure or a spatial locale. The squirrel climbing and scurrying about in the tree is a good example of this. Is the squirrel inside or outside of the tree? Learning to manipulate objects and learning to navigate paths are in many ways similar. Recalling that the perception of objects often requires the integration of information gathered by epistemic activity over a shorter or longer period of time, consider an animal engaged in exploring its territory. Very many animals do this quite systematically. The animal seems to be constructing a representation of the territory it is exploring, possibly a sort of spatial map, or possibly a representation in some other form. Rats that have been

allowed to explore a maze for some time before having to run through it for a reward show clearly that they have already learned much about how to navigate the maze without the reward. Many animals and birds are able to remember very large numbers of places in which they have cashed food, the nutcrackers being the most celebrated. Clark's nutcracker, in particular, is able to remember hundreds of cashing places in some cases for tens of weeks (Balda and Kamil 1992). Chimpanzees that are carried around while a human experimenter hides food in a variety of places will later go directly to each those places to retrieve the food. This sort of learning cannot be explained by classical principles of conditioning. That many animals can head straight for home after being transported blind to any portion of their known territories has often been considered evidence that they have something like maps in their heads, and not merely topological maps, but representations of a metered geometrical space (see, for example, Olton and Samuelson 1976). In any event, it seems they can construct some kind of general-purpose representation of the spatial layout of a territory.

Presumably, such a representation usually includes a representation of navigable paths or corridors within the territory. It will represent these as generalized or detached affordances, that is, as paths or corridors for traveling from here to there, for whatever further reason. Such a representation might serve any of numerous purposes when attached on one side to the animal's perception of its current location within the territory and on the other to representations of where various affording objects and places are within the territory. Joining an ego-implicit representation of its own momentary position and orientation within the mapped terrain to its egoless representation of the terrain as a whole, it is now ready to see any affordances within the terrain, old or newly learned, from within any part of the terrain. Recognizing the same place or part of the terrain again from another orientation, in different lighting conditions, in different weather and so forth is, of course, like recognizing any other object again from another angle, under different mediating conditions and so forth. The animal's perceptual representation of its current environment can now be extended or filled out by memory to produce a representation of more distant affordances and of the animal's current enabling relations to these more distant affordances. The representation of the enabling relation that extends beyond what is currently directly perceived may be a very abstract representation, but still definite enough to set the animal on the right path or heading in the right direction. Enabling relations will

be represented more exactly as the animal proceeds on its way and as more concrete places along the path come to be perceived directly and recognized.

Many animals are capable of representing temporal layouts as well as spatial layouts. As emphasized in chapter 13, there is nothing more mysterious about a capacity to perceive future events than about a capacity to perceive spatially distant events. And future events, like current events, are often perceived by integrating information over time. You perceive where a ball is going to land by tracking its projectory for a time with your eyes.[1] Among natural signs of the future are obvious goal-directed actions of others. For example, you perceive that a person is about to pick up the salt by seeing the direction of her reaching movement or by seeing where she is looking as she reaches. And just as many animals construct representations of the spatial layouts in which they live, many construct representations of regularities in the temporal layouts in which they live. This does not mean that they represent historical time, a time line that has various past and future events on it. In order to represent historical time they would have to have some use for representations of dated past events, and it is unclear what use this would be.[2] But an animal need not understand time as linear in order represent it. Rather, time seems to be represented by animals as a set of conditional probabilities or temporal contingencies, probabilities that concern what is likely to accompany what in time, what is likely to occur after what but less likely to occur otherwise, and at what temporal intervals various events are likely to occur.

It used to be thought that classical or "Pavlovian" conditioning caused an animal to treat the conditioned stimulus as a substitute for the unconditioned stimulus. For example, if a bell is rung every time just before a puff of air is blown into your eye, your eye will soon close automatically at the sound of the bell. But closer analysis suggests that what is learned is better described as a contingency relationship. The conditioned response is merely one that attempts to take account of that learned relationship. For example, Pavlov's dogs, for which the ringing of a bell was paired with the arrival of food, did not attempt to eat the bell when there was no food, but rather salivated in

1. There is evidence that perception of the location of moving objects precedes the actual location of these objects (Nijhawan 1994). They are perceived as being further along than they are.

2. Chapter 19 will address the question of how humans are able to construct and use representations of the historical past.

appropriate response to the expectation of food. The rat conditioned by a buzzer that accompanies electric shock jumps and its heart beats faster when it is shocked, but it cowers and its heart beats slower when it hears the buzzer alone (Gleitman 1991, p. 110). Simultaneous pairing of the unconditioned and conditioned stimuli is much less effective in training than when the conditioned stimulus is presented just before the unconditioned one. And pairing of a conditioning stimulus with an unconditioned one is completely ineffective if the unconditioned stimulus is presented without the conditioning one as frequently as with it (Gleitman 1991, p. 127; Rescorla 1988).

Perhaps the most dramatic evidence that a representation of temporal order is what lies behind classical conditioning is the work of Matzel et al. (1988). They showed that if rats are conditioned to associate a tone with a flash of light that followed at an interval *i*, then conditioned to shock occurring without a tone but just *before* the light at an interval less than *i*, the effect is a fear reaction on hearing the tone again.[3] That is, although the rats have never experienced shock after the tone, they have experienced light after the tone, and they have experienced shock just before the light. The result is that they represent shock as sandwiched between tone and light. Apparently, just as an exploring animal may construct, over time and over separate episodes of exploration, something like a map of the spatial locale in which it lives, it may also put together something like a map of the temporal contingency locale in which it lives.

Similarly, although it used to be thought that instrumental conditioning acts directly on responses, increasing the probability of whatever the animal did just prior to positive reinforcement, there is now evidence that what instrumental conditioning directly affects is the animal's representation of temporal conditional probabilities or of cause-effect relations. A nice illustration is given by the experiments of Colwill and Rescorla (1985). They showed that rats who have learned to pull a chain for sweetened water and to push a bar for food, abruptly stop pulling the chain but do not stop pushing the bar if (in another situation) drinking sweetened water is followed by nausea. Much more simply, pigeons, if conditioned to peck a key for food, peck with slightly open beaks ready to eat; if conditioned to peck for water, they peck with beaks almost closed ready for sipping water; if conditioned to gain access to a mate, they coo as they peck the key. It is clear that

3. Standard controls were used, of course.

they do not just respond to a stimulus but anticipate the results of their pecking.[4]

These, then, are two common kinds of factual knowledge that many animals seem to pick up and store away, ready for uses that may be discovered only later.

4. For an excellent discussion of the evidence that all conditioning is basically learning about conditional probabilities in the animal's environment rather than just learning to react or behave in certain ways, see Gallistel (1990).

16 Detaching Goal State Representations

The animal that constructs intentional representations of temporal contingencies must have a use for these representations. That follows from the description of intentional signs in chapter 6. The obvious use of these representations is preparation for the future. The animal uses representations of the future as representations of *current* affordances, as guides to *current* action, enabling it to begin *now* to position itself to make use of upcoming events and upcoming situations, or to take steps *now* to avoid them. Depending, perhaps, on its current needs, or its current state of potentiation, representations of future events serve for it as pushmi sides of P-P signs, which direct activities such as ducking an approaching ball or blow, salivating to be ready for swallowing, preparing to execute a sharp turn coming up just ahead when running, and so forth. That is, perception of what is distal in time operates exactly as does perception of what is distal in space. Just as the animal is guided *here* by a perception of what is *there*, the animal is guided *now* by a perception of what will be *later*.

I have spelled this out rather carefully for the following reason. In the case of an animal that predicts and represents the result of its own purposive action, it is easy to slip into thinking that in representing the future it is guided, not just *by* the future it represents, but *toward* the future it represents. That is, it is easy to confuse anticipating future events for which it itself is purposively responsible with using representations of those future events as guiding goals. Let me first illustrate the difference between these two, and then attempt to analyze the relation between having a purpose that one also expects the fulfillment of, and having that fulfillment as a goal guiding one's behavior.

I once had a cat, named Sam, who learned (as cats will) to push open the screen door to let himself out. But the door had a very strong spring on it, and it would close just in time to catch the end of Sam's tail. Sam

tried to avoid this by running faster, but the faster he ran, the more he instinctively lowered his tail to streamline himself, and the further behind him the end of his tail would be. The irony was, of course, that if he had just walked slowly through the door with his tail straight up, it would never have been pinched at all. Now the lowering of his tail was instinctive and biologically purposeful. And he predicted the effect of this purposeful motion and tried to avoid it. But representing the possible future effects of his activity in lowering his tail so as to prepare for or ward off these effects clearly was not the same as controlling the purposive activity of his tail itself in accordance with a representation of its possible effects.

Similarly, for an animal whose behavior is controlled entirely by perceptions of affordances, every facet of activity would be controlled by perceptions of what the brute facts are or are going to be. That some of the brute facts about what is going to be result from the animal's own dispositions to follow certain affordances would not change this matter. Adjusting to the consequences of the manner in which it is already disposed to act is something such an animal might learn, over time, to do. And learning not to do, in future, what it is currently disposed to do is also something that, over time, it might learn. But the goals of conditioning mechanisms are not projected ahead. They govern changes in behavior only after the fact. For such an animal, then, perceptions of the future effects of its activities could not serve as guides to control those same activities *currently*, but could at best control concurrent compensating activities. That an animal anticipates the outcome of what it is doing does not imply that this anticipation is what controls or guides what it is doing. Something more must be added if we are to understand what it is for an animal to *project a goal* and adjust its activity, on the spot, in order to meet that goal, rather than merely predicting the outcome of its previously learned dispositions to follow affordances.

But if it is my purpose to do *A*, and I anticipate that what I am actually doing will not lead to doing *A*, isn't it obvious that this will cause me to alter what I am doing? First, recall the protective eye-blink reflex (chapter 1). When the eye doctor is trying to put drops in your eye, you anticipate your purposive blink; it does not follow that you can control it. In this case the reason, or at least a reason, seems clear. The purpose of the eye-blink reflex is not a represented purpose. It is not on the same level as your conscious represented purpose of keeping the eye open, and so it will not be canceled out by this represented purpose.

Similarly, recall the greylag goose that doesn't notice that its egg-retrieval routine has failed in the middle of the process. Again, a reasonable assumption is that this happens because the purpose of the goose's activity is not a represented purpose. It is the purpose of an automatic response. Can we also explain, in this manner, why the hamsters didn't notice that the purpose of their food-gathering and storing activities was not being accomplished as they stumbled over one another, retrieving the same crackers again and again from one another's corners (chapter 13)?

That the hamsters didn't know what they were doing might be harder to imagine, because their activity was clearly being guided by various perceived affordances in the environment. They had to perceive the crackers as things to be taken home and stored, and they had to perceive the paths to their corners as paths to be followed, and so forth. However, the crucial question is whether perceiving an affordance involves representing the end state to which the affordance leads. I have spoken of the perception of an affordance—the harboring of a P-P representation—as a perception at once of what is the case and of what to do about it. But, first, a representation that directively guides one's doing, one's motions, is not the same as a representation of where one will end up as a result of that motion. The movement of my hand and arm may be directly guided by the turns of the bannister on the way down the old-fashioned staircase without involving any representation of the place the bannister ends. Similarly, a path can guide my walking without my knowing where it, hence I, will end up. Second, notice that if the representation of an affordance does somehow represent the goal state at the end of the path it offers, it surely can't represent it in the same representational system in which it represents the current situation. If the very same sign says both what the current state is and what future state is directed to be, it must be saying these things in different languages.

There is a sense in which the perception of an affordance is always a directive representation of a goal state. But usually it is an extremely inarticulate representation, directing no more than *when* the goal state is to be achieved, namely, at some more or less definite interval after the time of the representation. This is what it directs because this is what it is its purpose to achieve and this is what it maps onto, time mapping onto time (chapters 4 and 13). Other significant variables of the directive aspect of the P-P perception direct motions or paths, not end states. Consider, for example, how a guided missile tracks a target.

The information that guides it concerns the angle between its direction of motion and the target. It responds by continuing to correct its direction of motion until it matches the direction of the target, so if all goes well it ends up where the target is. But it employs no representation of itself ending up at the target. To suppose that it does would be to confuse a representation of the target's relation to its direction of motion with a representation of hitting the target. It would be to confuse a representation of an enabling relation to the target with a representation of what guidance by that relation will enable.

Similarly, consider how your motion is guided when reaching to pick up an object. What directly guides your motion is a perception of the current relation, the distance and direction, of the object relative to your hand (and of your hand to the rest of your body, for leverage and balance). Your motion is a complicated function of that distance and direction. The function is such that the motion of the hand (and compensating motions of the body) will result in your hand arriving where the object is. Nothing has been said yet about a *representation* of your hand arriving or having arrived where the object is. That would be another kind of representation. Following an affordance that results in picking up the object is being guided by your current relation to the object so that your movement is a function of that relation. The purpose of following that affordance is to achieve a state in which your hand is on the object. But merely in your perceiving and following the affordance, that end state is not represented, or not represented articulately. A "goal," in the sense that a goal may be perceived or represented in thought in the perception of an affordance, is just a place, such as the inside of a net or a basket. The goal may be represented, but the goal state corresponding to the affordance, say, the ball's being in the basket, is not part of what mere perception of the affordance represents. It is not part of what the P-P representation represents.

This discussion brings into sharper focus a central question I have proposed to consider in part IV. How does it happen that the descriptive and the directive sides of primitive P-P representations eventually separate and become independent? It now becomes clear that the answer cannot be entirely straightforward. It cannot be just that the P-P representation breaks in two, the two aspects or sides coming apart so that each stands alone. First, a *projected* goal, one that guides the animal in planning and in knowing when it has reached its goal, has to be a representation of a *goal state.* And a projected goal state that will

allow the animal to know whether or when that goal has been reached needs to be represented in the same language as the animal's descriptive representations. It needs to be represented in pushmi language. The goal state that is projected during an *explicitly* purposeful act is, in general, arrival at a state of affairs that will afford something further. It will afford arrival at the perception of another affordance, in the final instance, arrival at perception of a B-affordance (chapter 13). But for arrival of an affordance to be recognized as such, the language in which this arrival is explicitly projected as a goal needs to match the language in which it is later to be perceived as a fact. It has to match the descriptive side of the P-P representation that it anticipates.

A point that may superficially appear to be parallel is recognized by many researchers when discussing the roles either of efference copy or of reafference in guiding motor activity. For example, Jeannerod (1997) says, "In order to be useful for the comparison process, the reafferent signals must be compatible with the efferent ones. In other words, the two must be coded in the same 'language' for being mutually understandable and for the matching process to be possible" (p. 178). However, this seems to presuppose that what is articulately represented in the language in which motor activity is directed is the end state to be achieved rather than a path to be followed, which would surely lead to a problem of individuating or counting these end states. Jeannerod says, "Goal-directed behavior implies that the action should continue until the goal has been satisfied. The description of the motor representation must account for this property, that is, it must involve, not only mechanisms for steering and directing the action, but also mechanisms for monitoring and eventually correcting its course, and for checking its completion" (ibid., pp. 173–174). This image collapses the representation of the changing relation of the organism to its target that dynamically guides its motor activity with a representation of the end product to be achieved. It confuses dynamic perception or prediction of a changing enabling relation to a target with representation of a *goal state*. Representations of a variety of goal states and sub–goal states may be divided or individuated during an activity, but surely representation of changing relations to a target during short movements is dynamic. Indeed, with practice, movements directly under the control of dynamic input are extended to have longer and longer duration. The experienced tennis player sees the ball coming and hits it to a targeted place. Typically, no representations of any sub–goal states

intervene from start to finish. That, I take it, is a part of the Gibsonian vision that we want to preserve.

That there is a level on which representations of intended goal states and perceptual representations speak the same language has recently been argued for from experimental evidence by Hommel et al. (2001). Their claim is that there is a level of common coding for perception and for plans of action, and that this coding is of distal events. In the "Authors' Response" section of the paper, they are explicit that by "action planning" they do not mean motor coding, and that the level of perception they have in mind is not the sort produced by the dorsal system. Thus their claim, as I understand it, is compatible with the general point I have been trying to make here. An animal that not only looks ahead, predicting the results of its actions, but plans ahead, adjusting the predicted results of its actions to represented goal states, and that knows when and whether it has reached those goal states, must code its action plans in the same representational system in which it codes perception of the results of its actions. Unlike sphex and my hamsters, it has to represent *goal states* in such a way that it can recognize when they have been achieved. Moreover, since this kind of planning ahead is a stage by stage sort of process, it has to be undergirded by more dynamical action processes, presumably by direct guidance in accordance with perceptions of affordances, governed most directly by representations of enabling relations supplied perhaps mainly by the dorsal system. This also fits, for example, with the suggestion from chapter 14 that the achievement of objective representation may allow the animal to analyze its various activities into distinct achievement stages, understanding its activities as composed of series of transformations of one objective situation into another objective situation, hence as a series of completion stages in an objective process.

Compare the way in which language forms having the same satisfaction conditions but different linguistic functions or purposes may employ what are basically the same semantic mapping functions. Thus "Close the door," "The door will be closed," and "The door is closed" may all have the same satisfaction conditions, though the latter two have to be said at different points in time for this to be so. And they all express their satisfaction conditions in the same basic representational system. There are systematic ways to change an imperative mood sentence into an indicative mood sentence with the same satisfaction conditions, and systematic ways to change a future-tense sentence said at one time into a present-tense sentence said at a later time while retain-

ing the same truth conditions. Compare also the Fregean idea that one can have different "propositional attitudes" toward the same proposition (though distinctions between the tenses have not generally been understood this way). But I don't want to suggest too close an analogy with sentences or with Fregean propositional attitudes either. Unlike sentences, mental representations with the same satisfaction conditions toward which different mental attitudes are taken might not need to display different syntax but merely to have different uses or functions. And as for propositional attitudes, we have been dealing with perceptual representations and with representations continuous with or joined to perceptual representations through recall, not with conceptual thought. No position on the relation of perceptual representation to conceptual representation, or even on whether there is such a distinction, or whether one fades into the other and so forth, is intended.

The idea to be explored is that beyond pushmi-pullyu representation, a *common* system of mental representation may develop in which projected goal states, objectively represented future states, and objectively represented present states can all be expressed. ("Objectively represented" means that the representing organism is not implicitly represented, and hence if represented at all it is explicitly represented, represented as an object. See chapter 13.) We suppose that the pushmi side of an inner P-P sign of an objectively represented affair is already expressed in the required format, for example, the pushmi side of your objective representation that coffee is now in the cup for drinking. Recall that the pushmi sides of some inner P-P signs represent coming events rather than present ones. Sometimes these coming events are not represented objectively, for example, it is doubtful that the nutcracker who responds to signs of winter coming by storing nuts represents winter coming as an objective future occurrence. The nutcracker's perception is a representation of winter coming in that it will serve its own or proper function in a normal way only if winter is indeed coming, but it is unlikely to be an articulate representation of that; for example, it is unlikely that any transformation of that representation says that summer or spring is coming or, as is more relevant here, that winter is already here. We are now postulating a new kind of representational capability, such that coming events and present events are coded in a common representational system, the difference between a representation of something future and a representation of the same thing as present perhaps lying only in the use rather than the coding of these representations. What this makes possible, in the first

instance, is that the representing organism may learn to represent certain coming events, on the basis of present local signs, by a trial and error or corrective process in which its predictions are remembered and compared with later outcomes. It can do some learning about what in the present is a sign of what is in the future without having to use the criterion that is the success or failure of immediate practical activities that depend on these predictions. It forms hypotheses about what objectively follows what and tests these by observation rather than practical action.

Besides present and future states and events, we postulate that projected goal states are also represented in this common representational system. Thus the organism is able to know when or whether it has reached these states. This allows cessation at appropriate times of activities designed to lead toward projected goals. It thus prevents the sphex effect (chapter 13). Does it also allow learning what kinds of motor activities do and do not accomplish certain goals?

The assumption of classical ideomotor theories of perception, as Hommel et al. (2001) put the matter, is that

> [t]hough the learning refers to linkages between movements and their effects, the result of this learning needs to be organized in a way that allows to use the linkages the other way round, that is, go from intended effects to movements suited to realize them. . . . If one takes for granted that the links between [representations of] movements and [representations of] effects can be used either way, a simple conceptual framework for the functional logic of voluntary action offers itself. This framework suggests that actions may be triggered and controlled by goal representations—that is, representations of events the system "knows" (on the basis of previous learning) to be produced by particular movements. (sec. 2.2.2)

Hommel et al.'s "goal representations" seem to be what I have been calling "goal state representations." But we are assuming here that motor activities are always directly guided by perceptions of affordances, where complete perceptions of affordances always include perception of enabling relations (chapter 14). Learning what kinds of motor activities do and don't accomplish certain goal states is thus the same as learning which affordances, which ways of being guided by perception, lead to which objective results. Actions will not be directly triggered by goal state representations unless suitable affordances, including enabling relations, are perceived. On this assumption, what will be learned by the animal is not what events will be produced by particular movements, but what affordances, if

followed or acted on, will lead to what objective results. That, it seems likely, is part of what young mammals and birds are learning when they are playing. In the case of human infants, experimentation begins in the first few hours of life as the infant moves its body parts in response to both internal and external stimulations, at first quite randomly, later with more and more purposive intent, to see and hear and feel what happens.

Hommel et al. emphasize that the level at which "common coding" occurs concerns events at a distal level. For example, this coding may represent the effect that is a red light flashing on to your left rather than representing the hand motion (pushing a key down or lowering your finger) that causes the red light to flash. Their studies also suggest that when attention is shifted slightly inward to concentrate on a more proximal effect, say, on the key pressing rather than the light flashing, then the common coding level shifts also to the more proximal level. These conclusions are reached mostly from the results of reaction time experiments that suggest coding interferences and coding facilitations among actions of various kinds under different instructions. These interference and facilitation relations may shift within what is exactly the same experiment looked at in terms of actual outcomes produced by the experimental subjects, but in which the subjects are given different instructions designed to encourage them to focus their interest on more or less distal aspects of the outcome. For example, they are asked to be sure to press the left or the right key or, alternatively, to quickly make the light on the left or on the right come on. It appears the common coding may shift with attention shifts from representing the position of the hand or finger to representing the position of the intended light (Hommel 1993).

This suggests that motor-perceptual learning, learning how to be guided by perceptual inputs so as to produce predicted outputs, may take place at many levels of distality, depending on the interests of the learner. The infant learns how to be guided by perception so that an object is transferred (by reaching and grasping) from any of various positions into its own hand, or from one hand into the other, or so that an object across the room is transferred (by the infant's crawling) to being within reach. It learns how to perform simple manipulations, such as turning things over, twisting, throwing, and so forth, such that the outcome is as anticipated. The tennis player learns how to be guided by perception so that the approaching ball is transferred to a designated position in the opposite court. Indeed, there is evidence that

motor-perceptual learning that concerns distal events of this kind is more primitive than motor-perceptual learning that concerns merely motions of the organism's own body. As mentioned in chapter 14, very few animals can even learn to imitate bodily motions, and perhaps none but humans[1] do so naturally.

The animal that has learned its lessons well concerning the objective results of following certain kinds of affordances is in a position to predict the fulfillment of its immediately projected goal states in situations perceived as enabling. This clarifies the status of explicit intentions which, on the one hand, are projections of goal states, but on the other, are beliefs about the future. To firmly intend to do a thing consists, in part, in believing that one will do it. Otherwise one intends only to try. A confident intention, somewhat like a pushmi-pullyu representation, tells both what to do and what will have been done, so that further plans that depend on that settled future can now be made. Perhaps we could call a confident intention a "pullyu-pushmi" representation, though the symmetry is not perfect. Unlike the pushmi-pullyu representations, the two faces of a confident intention are both written in the same code. The very same aspect of the representation has two functions, one to produce action, the other to represent a future state of affairs.

A fascinating piece of evidence that the representation of goal states is not part of the mere representation of affordances in the case of humans is the difference between patients suffering severe damage to the prefrontal lobes, and those suffering with an "anarchic hand," caused by damage to the supplementary motor cortex. In both cases it appears that the behaviors of the patient or of the patient's hand simply follow the most obvious current affordances exhibited in the environment, whether appropriate or not. Thus the patient with severe prefrontal damage will pour water into a glass and drink it as many times as water is presented in a pitcher, regardless of whether she is thirsty. The hand of a patient with an anarchic hand may button and unbutton available buttons, grasp a pencil and scribble with it, or grasp a doorknob and turn it in entirely inappropriate situations. A difference between these two kinds of patients, however, is that the patient with prefrontal damage is in no way disturbed by her inappropriate behavior, whereas the patient with an anarchic hand is very disturbed by it,

1. Actually, it looks as though dolphins may do this. See Herman (2000, forthcoming).

in some cases even tying the anarchic hand down to prevent its wanderings. This difference is interpreted by Frith et al. (2000) to be the difference between patients that have intentions that conflict with the affordances they are following and those that simply don't have any intentions, intentions being a product of the prefrontal lobes. As mentioned in chapter 1, automatic behaviors are accepted as one's own so long as one doesn't disagree with them.

17 Generating Goal State Representations

The picture that emerges from chapter 16 is apparently very neat. If the animal has trained itself well, given the perception of an appropriate enabling affordance, the representation of a goal state becomes a representation of a predicted future state, which, if no infelicities occur, is soon followed by a correlative perception of a current state. But these optimistic reflections leave an urgent question unanswered. The neat picture has a hole in the middle. Recall that the animal driven only by representations of affordances derives these representations, and hence derives its motivations, from current perceptions of its environment. This is so even when current perception is supplemented or extended by stored knowledge of the spatial and temporal layout of currently unperceived parts of the animal's home domain (chapter 15). But where do the representations of projected goal states come from? What is their origin; what prompts them? We have discovered some possible origins of detached representations of fact. But nothing has been said about the origins of detached representations of goal states.

Recall, for example, the squirrel studying how to get to a bird feeder (chapter 1). I watched one recently trying to reach a feeder hanging on a chain from the eaves overhanging the deck of our house. It studied the situation from under the feeder, then from one side of the deck, then from the other. It ran slowly along the deck railing, looking from one side, then from the other. It did this several times on several different days. Finally it took a run along the railing from one side, ricocheted off the screen of the door to the house, landed precariously with its front paws on the edge of the feeder and pulled itself up. Without doubt that squirrel had a goal in view the whole time, indeed, quite literally in view. It saw the bird feeder, which afforded approaching and feeding from. The squirrel's difficulty was that it did not yet perceive any enabling relation to guide it to utilize that affordance. A complete

affordance would have to include mediation by a path between the squirrel and the bird feeder. The squirrel's perception of a goal controlled its action as it moved from side to side trying to see a path. But this goal representation was probably the representation of a goal, not of a goal state. It was a representation of something present, not future. The origin of this representation was perfectly plain. It derived directly from the squirrel's perception of its then current environment.

On the other hand, consider what it was for the squirrel to be "looking for" a path. Recall that even the purely pushmi-pullyu animal, if advanced at all, does not just happen to perceive and hence act on affordances. Depending on its current needs and on its current environment, it will be ready to act on certain affordances and not others. Moreover, as suggested in chapter 13, likely it is disposed to perceive only certain affordances and not others. Thinking of this on a connectionist model, its nervous system is primed to register certain kinds of inputs easily whereas registration of other kinds may be inhibited. Priming a certain perception is effected by partially activating the neuronal patterns whose full activation would constitute that perception. Thus the animal might envision ahead or "imagine." It is looking for something quite definite, or for something within a definite range. It knows or envisions, though guite abstractly, what it is looking for, and it is ready to respond appropriately when it finds it. Similarly, the squirrel may know or abstractly envision what it is looking for. It is envisioning some sort of path. Its visual system is primed to register paths. Will this priming or envisioning actually help to guide or control the squirrel's search, or will it merely shorten the response time by a few milliseconds if a path happens to be picked up by its visual systems?

There is a way, I believe, in which what it envisions will actually help to control its search. It is primed to see paths; hence it will see partial paths too. It will concentrate on or visually explore paths from where it is to places nearer the feeder. And it will concentrate on or visually explore paths from places near the feeder to the feeder. Working both from where it is forward and from the feeder backward, by a trial and error process but one that is carefully directed or constrained, it may eventually discover a path all the way from where it is to the feeder.

If that is right, it is a process rather like practical reasoning. Practical reasoning is often described as reasoning in something like the form of a proof: *I desire A; doing B will probably lead to A; therefore I will do B*. But being more careful, that is not the way practical reasoning generally goes, but only the way practical conclusions are justified to other

people. The core of a practical reasoning processes is a *search* for a proof. Just as in mathematical reasoning you are likely to start with something you would like to prove, in practical reasoning you begin with something you would like to do or to have done and then attempt to construct something like a proof, a path from premises you have to a conclusion you would like to reach. And you do this largely by controlled trial and error. You start with what you would like to prove and work backward, trying to find plausible steps that might lead to that conclusion, and you start also with things you already know to be true and work forward to see where these things might lead. You try to fill in the gap between what you find going forward and what you find going backward. The squirrel is like the practical reasoner in that its search for a path is actively controlled by a vision of its goal as well as by its perception of where it now is in relation to that goal. It differs from the practical reasoner in that its vision is of a currently existing goal object rather than of a future goal state (chapter 16). Also, the path it is searching for will be discovered by perceiving a pattern in the actual current situation. This path is a configuration of objects, not a chain of possible future states of affairs leading to a goal state.

Compare the squirrel looking for a path to the bird feeder with one of Köhler's famed chimpanzees that is looking for a way to reach a banana. A chimpanzee that has been allowed to play with boxes that can be stacked one on another will sometimes see that the way to reach a banana that is high overhead is to stack the boxes and climb up on them. The chimpanzee does not just look until it perceives an actual path. It looks, or thinks, puts representations together, until it sees how to *construct* a path. In this case it seems clear that the animal must be representing objective situations or states of affairs that would result from following certain affordances. Recall that what was missing in the purely pushmi-pullyu animal was that it did not represent to itself where the affordances it looks for lead, nor, of course, does it follow perceived affordances because it knows where they lead. The disposition to look for and follow an affordance comes first, having resulted from natural selection or from conditioning. The representation, if any, of expected results follows after. It is not what controls the behavior.

The chimp, on the other hand, apparently not only sees the boxes as affording stacking, but knows from experience that the result of following this affordance is the creation of a path that can be climbed. Like the squirrel looking for a path from here to the food, it works from here to there and from there to here. From here it sees boxes that afford

stacking, which will result in a path that can be climbed. From there it is looking for a path to the bananas, a place from which it can reach the bananas. Following the stacking affordance will result in a complete path. Its representation of the result of following the box-stacking affordance now directs its activity rather than following after. The chimp follows that affordance because it knows where it leads. Apparently we have here an animal capable of projecting goal states and following affordances because it represents them as leading to these goal states. The result is what animal psychologists call "insight" or "reasoning." And it is quite a bit like explicit human practical reasoning. It is a form of trial and error, an important part of which takes place in the head using representations of mere possibilities (chapter 1).

Less dramatic than the chimp's performance is the performance of the rats mentioned in chapter 14, who were conditioned to pull a chain to obtain sweetened water but who ignored the chain after experiencing nausea from drinking sweetened water in a different context. Apparently what the rats learned was that following the chain-pulling affordance results in the presence of sweetened water. When they were interested in what sweetened water affords, namely drinking, they pulled the chain. When they were no longer interested in what sweetened water affords, they were no longer interested in producing the result of chain-pulling, and hence were no longer interested in pulling the chain. The anticipated result of pulling the chain was in control of the rat's behavior. Before the rat had experienced the nausea, it was motivated perhaps by thirst, which arose from the current state of its body. The thirst primed it for perceiving drinking-affordances, among them the perception of sweetened water. It was thus looking for or had an eye open for sweetened water. Its perception of the chain-pulling affordance and its memory of the outcome of following this affordance then produced the representation of a complete path from where it was to drinking.

The squirrel, the chimp, and the rat all differ from the purely pushmi-pullyu animal in this important way. Some of the affordances they perceive are perceived as mere possibilities. They are not motivated by every affordance they perceive, but only by what they see as part of a complete path to a goal or goal state they project. Though the squirrel may perceive many paths leading from here toward its goal, it does not follow most of these paths but rejects them when it cannot envisage their completion. Similarly, the chimp searching for a way to the bananas may perceive the boxes as affording things other than stack-

ing, such as turning over and climbing inside or climbing up on without stacking. Searching systematically for a path may require following various leads merely far enough to form representations of where they are leading, then rejecting them. Similarly, the once poisoned rat may perceive the chain-pulling affordance but reject it. Nor need we suppose that the perception of a variety of possibilities is always serial. In describing the lattice hierarchy, Gallistel emphasized that many alternative behaviors may be potentiated at once by the same stimulus. Similarly, the animal that envisages the results of following various perceived affordances may be capable of representing a variety of branching possible futures in parallel (though not necessarily consciously) and "deciding" among them.

In chapter 1 I noted that there is a gap between a certain stimulus or experience acting as a reinforcer and one's awareness of what it is about that stimulus or experience that makes it reinforcing. Presumably this gap is not filled by the purely pushmi-pullyu animal for any of its reinforcers. We humans are not always aware of what it is that conditions our behavior either, and sometimes we are aware of aversions or attractions without knowing exactly what it is that averts or attracts (chapter 1). But the ability to represent causes of reinforcement or to represent situations offering affordances leading to B-affordance conditions (chapter 13) as *objective* situations or occurrences is prerequisite to projecting these as goal state occurrences. You have to know what you want if you are to represent having it as a goal state.

In chapter 14 I remarked that the achievement of objective representation may allow the animal to analyze its various activities into distinct completion stages, its extended activities being grasped as a series of transitions from one objective situation into the next objective situation. The objective situations grasped were not merely factual situations but situations that afforded this or that. Thus the chimpanzee might grasp a transition from the box being on the floor to its being on top of another box, or you might grasp a transition from the tea's being in the pot to its being in the cup. The ability to carry out an extended activity from a starting point *A* through a series of transitions to *B*, then to *C*, and so forth, finally to *D*, where *D* is the presence of a B-affordance condition that was projected from the start as a goal state, may be most likely to result from prior experience of having progressed first from *C* to *D*, then from *B* to *C* to *D*, then from *A* to *B* to *C* to *D*. Certainly that is the easiest way to teach an animal to progress through such a series of stages to reach a goal state. But in considering Köhler's

"insightful" chimpanzees, we looked at another kind of possibility. Köhler's chimps had experienced progressing from *C* to *D*, from being in a place where food was in sight, to moving to a place where it was within reach, to procuring food. They had also experienced progressing from *A* to *B*, from being within reach of boxes, through stacking the boxes, to being in a place that afforded being within reach of things higher. Their accomplishment was to put these two partial paths together to make a full path to obtaining the banana. The difficult part, we can suppose, was not to see that if these two paths were joined they would lead to procuring a banana, but to happen to represent putting these two paths together.

The difficulty was that not only the boxes but many other things in their cages afforded a great many alternative activities. To think of stacking the boxes might seem like happening on the needle in the haystack. Nor am I prepared to speculate by what mechanism the intelligent animal's search for the right combination proceeds in such cases. What does seem evident, however, is that if the animal needs to put together for the first time more links in a chain, more sections of a path that has never been traveled before, the problem of finding a correct linkage increases in difficulty exponentially. It should get harder and harder to happen to think of a combination that will work. And indeed, putting together several links in such a chain does seem to be something that perhaps only people can do.

How do people do it? Again, I am not prepared to speculate much on mechanisms. But first, it is worth pointing out that putting together such chains does often take considerable time and considerable concentration. Planning a trip, for example, can take many hours, not just collecting information, but figuring out how to put that information together to produce the desired result. Or suppose that I wish to communicate a message to Paula. Working backward, I might think of doing this, say, by direct encounter, by phone, by e-mail, by letter, by messenger. Each of these possibilities may be considered. Do I know where Paula will be in the next day or two? Will I be in any of those places? Could I easily get to any of those places? Do I know anyone else who is going to any of those places? Does Paula have a phone? Do I have her number? Do I know someone who knows her number? Will she be in the phone book under her own name or under her husband's? Do I know where there is a phone book? Do I know where there is a phone? Does Paula have e-mail? Do I know her e-mail address? And so forth.

Perhaps more important, long-range activities that we plan are never planned in detail. Planning and execution have a hierarchial structure. The plan is first filled in in chunks, and then the details of the chunks are filled in as we proceed. When we carry out a long-range project or action consciously and deliberately, what we explicitly intend to do at the start is represented very abstractly. Given our past experience, we know what general sorts of ends we are generally capable of achieving from what sorts of starting points. But the details of exactly how we will fulfill a particular intention in a particular case are represented only by the confidence, "I will know how to do that part when I get there." In order to attend a meeting, I plan very definitely to go to Boston on November 16, knowing I am, in general, capable of getting to nearby cities, and from within them to designated hotels. But perhaps I do not know even what basic form of transportation I will take, let alone the thousand other details of my trip. These will depend on the details of circumstance that I encounter later or along the way, such as how much the university will reimburse me, whether the trains run there at reasonable hours, how close the train station is to the hotel, and whether it turns out to be raining that day. Similarly, how the chimp will move in proceeding to stack his boxes will depend on where they are currently placed in his cage, how heavy he finds them, how large they are compared to his arm length, and so forth. The chimp also projects ahead the result and the means of its planned labors only abstractly.[1]

1. Here is Gallistel's description of how action is planned within the motor system:

> As a rule of thumb, the higher the level receiving a sensory input, the more global and diverse will be the possible effect of that stimulus on the animal's action. As one ascends the hierarchy, stimuli play more and more of a role in determining the *general course* of action and less and less of a role in determining the *particular pattern* of muscular activity used to pursue that course at a given moment. A correlary of this principle is that the higher one goes in the hierarchy the more elaborate the sensory/perceptual analysis of sensory signals; or, what is not quite the same thing, the more global the sensory factors that serve as inputs. The generals determine where the armies are to be deployed. In doing so, they must respond to the geography of the country and the deployment of the opposing armies. The lieutenants determine where the trenches are to be dug. In doing so, they must respond to the local topography and the disposition of opposing forces in their locales. The sergeants determine where the latrines are to be dug. In doing so, they respond to the distribution of bushes in their immediate vicinities. (Gallistel 1980, p. 286)

18 Limitations on Nonhuman Thought

The causes of the wide separation between the abilities and accomplishments of humans and of other animals have been a topic of interest for thousands of years. I don't expect to add anything definitive to that ongoing discussion. But I will make a couple of suggestions about ways in which the representational capacities of humans may differ from those of nonhuman animals, certainly in degree and probably in kind. If I am correct, these differences could account for at least a portion of the separation.

Although under some circumstances some nonhuman animals are capable of putting somewhat novel chains of behavior together in pursuit of their goals, most animal behavior and also most animal learning is not so flexible. James Gould describes the "rigidly programmed plasticity" (1982, p. 268) characteristic of most animal learning as follows:

> [. . .] Learning is adaptively programmed so that specific context, recognized by an animal's neural circuitry on the basis of one or more specific cues, trigger specific learning programs. The programs themselves are constrained to a particular critical period . . . and to a particular subset of possible cues. Nothing is left to chance, yet all the behavioral flexibility which learning makes possible is preserved. (Gould 1982, p. 272)

> Learning, even in higher vertebrates, seems less a general quality of intelligence and more a specific, goal-oriented tool of instinct. Bouts of learning such as food avoidance conditioning, imprinting, song learning, and so on, are specialized so as to focus on specific cues—releasers—during well-defined critical periods in particular contexts. Releasers trigger and direct the learning, and in general the learned material is thereafter used to replace the releaser in directing behavior. As a result animals know what in their busy and confusing world to learn and when, and what to do with the information once it has been acquired. Most learning, then, is as innate and preordained as the most rigid piece of instinctive behavior. (Ibid., p. 276)

In this preordained way, many animals learn either by trial and error or from conspecifics what to eat and what not to eat. Some learn from others which local species are their predators. The European red squirrel laboriously learns how to open, specifically, hazel nuts. The oyster catcher laboriously learns how to open, specifically, oysters. The chimpanzee laboriously learns, specifically, how to open nuts by using a rock and an anvil. Despite the fact that some animals are capable of some "insight learning" in new, though carefully designed, situations, speaking generally, what animals are capable of learning, and hence, it is reasonable to suppose, what they are capable of developing representational systems to support, is closely tied to specific skills or specific ends found to be useful in the past history of the animals' species. Through rigorous and careful step-by-step training by humans, individuals of many higher species can laboriously be brought to recognize perceptual affordances of kinds quite remote from any they were specifically designed to learn, and recognizing these affordances may involve recognizing properties and kinds of objects with no history of relevance to the animal's species. But they seem to be recognized only as things that have proved useful in the individual's previous experience and only as affording those known uses. Nonhuman animals do not learn to recognize objects or kinds for which neither they nor their species has yet found any practical uses.

Similarly, that an animal can collect and later use specific kinds of purely factual information about the space it lives in, or about the temporal order it lives in, has no implications for whether it can represent any other detached facts. That it collects and remembers information about local spaces and temporal contingencies depends on the fact that this kind of information has been useful during its evolutionary history, and useful in entirely specific ways. For example, some species of birds can remember hundreds or even thousands of caching places in which they have left food. It does not follow that they are capable of collecting and remembering any other particular kinds of facts. Nor does it follow that they can use knowledge of these caching places for purposes other than finding food again when they are hungry. Likewise, although it appears that some nonhuman animals can learn something about causal chains, about what can turn into what, the chains that they recognize are not arbitrary chains. They are chains that have at the end something antecedently of interest to the animal. The representations of fact that animals collect always seem to be dedicated in advance to

very specific uses, to completing complex P-P representations of predetermined kinds.

Just as individual animals may collect specified kinds of factual information ahead of having uses for it, they may also collect specified kinds of skills out of the context of serious uses for them. Young mammals, in particular, do a lot of playing. But play in nonhuman animals seems always to be practice for well-defined species-typical adult activities. The detached skills they are learning are always closely related to future uses known, as it were, to the species, though not to the individual. The historical experience of the species tells the genes of the young animal what it should practice.

We humans, on the other hand, collect and remember facts of kinds for which neither we nor our ancestors have yet found any practical uses. We are capable of learning thousands of facts about what has occurred or is occurring at times and in places to which we have no potential access, let alone past or present practical acquaintance. The nonfiction sections of libraries are repositories, largely, for immense collections of such facts. Some people memorize baseball scores and batting averages, or time tables for railroads all over the country. We are curious about what will cause what and why, wholly apart from any envisioned practical applications for this knowledge. We may be curious about how things work, where they came from, what properties and dispositions they have, in a completely disinterested way. We notice and remember not just what we can cause, or what causes something we want, but what causes what, quite out of context. We interpret natural signs and also linguistic signs of world affairs that are distant from us in both time and space, quite outside of the realm of our powers of action. We are adept at learning to interpret new kinds of signs, not just human language signs, but signs produced by meters and scopes and a multitude of other instruments. And we make inferences from these various kinds of facts to further disinterested facts, reconstructing large portions of the layout of the world that are hugely distant from us in space, time, and magnitude, far removed from the level of perception for which evolution has specifically prepared us.

We also spend energy and time developing skills, both physical and intellectual, for which neither we nor our ancestors have had any practical uses. Children at play practice bouncing balls, juggling, standing on their heads, spinning hula hoops, solving Rubik's cubes, riding skate boards, cracking their knuckles, wiggling their ears, blowing

bubbles, whistling through their teeth, spinning around to make themselves dizzy, and so forth. Both children and adults become absorbed in games of all kinds, from sports games through board games to gambling games. Two-month-old infants, provided with a device that allows them to activate a mobile over their cribs by moving their heads, will keep smiling and cooing, whereas they very soon tire of a mobile that moves independently of their actions (Watson 1967). It seems that the development of any sort of skill, the discovery and mastery of affordances with any sort of determinate outcomes, may be of interest to a human child.

It is also interesting to contrast what motivates nonhuman animals with what motivates humans. The motivations described in chapters 13 through 16 all originated either with the perception of affordances in the immediate environment or with the animal's perception of its own current needs. The animal's goals arise out of past experience of having reached those goals in certain ways in the past plus awareness of present relations to things that were involved. Even our most respected and intensively studied relatives, the monkeys and apes, seem to derive their motivation entirely from perception of the current situation. Thus, for example, Merlin Donald summarizes the literature on signing in apes: "the 'meaning' of an ASL sign to an ape is simply the episodic representation of the events in which it has been rewarded . . ." (1991, p. 154), and "The use of signing in apes is restricted to situations in which the eliciting stimulus and the reward are clearly specified and present, or at least very close" (ibid., p. 152). No nonhuman animal, I suspect, wonders where its next meal is coming from unless it is already hungry, nor does it wonder how it will cope next winter. Of course, appropriate migrating behaviors are elicited, in certain species, by natural signs that current food sources are running out, various behaviors are elicited by natural signs connected with the immanent approach of winter, and so forth. The indicative facets of the inner P-P representations that are responses to these natural signs concern the future, for these representations will produce behaviors that are appropriate only if certain future events are indeed imminent. But these are present perceptions, and they derive directly from the past history of the species. We humans, on the other hand, ardently collect dreams of things we would like someday to do or have done, places we would like someday to go, things we would like someday to build or to have or to be, without necessarily having any notion yet of how to fulfill any of these dreams. Certainly

these dreams neither reflect currently perceived affordances nor originate from currently perceived needs. We store desires that we do not know how to fulfill just as we store facts that we do not know how to use.

Representing irrelevant facts, irrelevant affordances, irrelevant dreams! In short, we appear to be compulsive collectors of all kinds of representational junk. Moreover, we use these representations primarily for the purpose of making more representations of the same kind, moving from one representation to another via inference, filling out our knowledge of places, times, and magnitudes far removed from immediate practical experience and activity. But, of course, although most of the individual facts, skills, and dreams that we collect may never find uses, the general disposition to collect junk does find uses. If you have enough storage space and a good enough retrieval system, some pieces of that junk may well come in handy sometime, though there was perhaps no way to tell in advance *which* pieces. Having stored enough tools and materials in the attic over the years, eventually some of it is bound to come in handy, granted one is an inventive enough tinker with ideas. (The adage says you need only wait seven years.)

But I think that a difference between us and the other animals may not be just that we have bigger storage barns, bigger brains, than do neighboring species, although that may well be some part of the matter. It may be that we are also peculiar in having what Dennett (1996) likes to call "Popperian" minds, that is, minds that spend a good part of their time "generating and testing," making thought trials and errors, learning by experimenting with inner representations rather than by making false starts in outer behavior. And it may be that to be efficient, Popperian minds need to operate on representations coded in a different kind of representational style than that needed for direct perception of affordances.

Two demands that would be seem to be placed on the representational system used by a Popperian mind suggest this. First is a demand for free inferential interaction among representations regardless of content, for the Popperian mind cannot tell in advance what may need to be put with what in building a useful result. This may require a uniform notation not found on the level of perception, moving as it does directly from inputs to the various senses to guidance of action, variations in input mapping fairly directly to variations in output. The second concerns the development of representations that have not been tested for accuracy through practical experience.

I will attempt a sketch of how the distinctively cognitive systems, unlike the action-guiding perceptual systems, may employ representations of a different type than any we have discussed so far. These may be representations that are more like sentences than, say, bee dances, in that they are articulated into subject and predicate and are sensitive to an internal negation transformation. Call this the development of theoretical concepts and theoretical knowledge. This development, I will argue, would make it possible to represent time as dated or "historical" rather than as a mere set of conditional probabilities concerning temporal relations. Representation of historical time, in turn, makes it possible to conceive of, plan, and carry out projects that purposefully change the future in unprecedented ways, rather than merely repeating past successes. It allows representation of novel future possibilities that, in turn, uncap new motivational springs. I will take up these themes in order, the second in chapter 19.

The argument of chapter 16 suggests that the original code in which perceptions of detached facts and representations of projected goal states are represented is the same as the code in which the descriptive sides of pushmi-pullyu representations are coded. These codes are designed, in the first instance, to perform two functions at once. Transformations of P-P representations must, on the one hand, correspond systematically to transformations of the affairs in the world that they signify, but on the other, they must correspond to transformations of the responses they govern such that the responses are adapted to those affairs. The detached representations of facts we have considered, for example, representations of remembered spatial and temporal surrounds of the animal not currently perceived, must also be coded in such a way that they can be joined to representations of the animal's current position in that surround so as to direct the animal's motions immediately. But we are now considering an entirely new kind of representation. We are considering an animal that collects facts for which it has, as yet, no known uses, combining these facts with one another to produce, by inference, more facts with no known uses. The human animal is engaged, much of the time, in constructing large portions of a four-dimensional map of a whole dated world in progress, mapping not only things that endure or recur (individuals, places, natural kinds, repeated patterns of events) but also unique occurrences, both in its own locale and in other places. Many of the things represented, moreover, are not things that could possibly guide motions directly, being too small, or too large, or too amorphous or too abstract. All these facts

may be represented quite apart from any currently known relevance to the thinker's practical interests. The immediate, even though not the ultimate, use of these representations is merely the efficient production of representations of more and more of the structure of the disinterested objective world. What that project requires primarily is not representations suitable for guiding continuous motions, but representations that are able freely to interact with one another in inference.

Whether or not representations can freely interact in inference does not depend on satisfaction conditions. Nor does it depend, of course, just on their locations in the brain, on whether or not they are physically isolated. Encapsulation of information used for one purpose but unavailable for others will occur whenever incompatible notations are used. Putting things graphically, suppose that the first premise of a would-be inference is stored in a Venn diagram and the second in the notation of *Principia Mathematica*. No rule of inference could combine these premises directly, without some kind of translation. How the contents of mental representations are articulated and represented, as well as how they are stored and retrieved, is crucial to their interaction. But there seems no reason to suppose that a representational system tailored to safely guide an animal's continuous motions through its immediate environment would also be suitable for encoding and amplifying its theoretical knowledge.

There is a stronger reason to suppose that humans may need to employ a new kind of representational system in order to represent places, times, and magnitudes far removed from practical activity and experience. An animal attempting to construct maps of parts of the world that it is not currently using for anything clearly is at great risk of error. Compare generalization that connects experience to behaviors with generalization that connects experience to idle beliefs about facts. Practical generalization is naturally bridled. Unsuccessful behaviors do not always produce punishment but they do waste time and energy, naturally diverting the animal's responses into other channels. What kind of bridle is there on false generalization in the case of theoretical inference? Humans need somehow to collect evidence for the objective adequacy of their abilities to reidentify objects and properties through diverse appearances independently of using those objects and properties for practical purposes. They need to be able to test their empirical concepts independently of pragmatic successes and failures. When perception is used to guide immediate practical activity, the criterion of

correct recognition of affording objects or properties is easy. You are right that this is the same affording object, or kind, or property again if you can successfully deal with it in the same way again. The proof is in the eating. But how does one learn to recognize new objects, new kinds, and new properties that have for one, as yet, no practical significance?

Recall what is involved for an animal that needs to be adept at recognizing local signs of some object or objective property or relation. The difficulty is that local signs of the same thing are manifested through many diverse media, under a wide variety of conditions, and transmitted through a number of different sensory modalities. We are generally unaware of the enormous complexity of the task of interpretation required here. It is accomplished by a complex neural machinery of which we have no knowledge and less control. But the complexity is such that despite intensive study by neurologists and psychologists, only a sketchy knowledge of small parts of these mechanisms is yet available. The perceptual task of interpretation is enormously demanding. How does the organism know that it is doing it right? Or if some of these abilities have been built in by processes of selection during the evolution of the species, what life-supporting effects of their use were being selected for?

Start by thinking about how the organism tells that it is perceiving distances correctly. What *is* it for an organism to have represented a distance correctly in perception? Where is the perceptual dictionary written that tells what rule is the correct correspondence rule between a certain objective distance and some mental or neural trait that is required to represent that distance perceptually? The question has no sense. Perceiving spatial relations correctly just *is* knowing how to be guided by them during action. The correct representation is whatever one the action systems can read. Correctly perceiving *where* an object is simply *equals* knowing how to reach one's hand to it, how to kick it, how to walk to it or away from it, how to throw something to hit it, and so forth. For the animal engaged in practical activities, that is absolutely all there is to correctly representing spatial relations. This follows immediately from (1) the assumption that the primary kind of spatial perception is perception of spatially defined affordances, (2) the assumption that perceptions of affordances are pushmi-pullyu representations, and (3) the description that has been given of intentional pushmi-pullyu representations. There is no distinction to be drawn between wrong perceptual recognition of spatial relations and wrong

behavioral responses. When you put on a new pair of glasses so that the floor now looks like a small wavy hill in front of you, what is it that needs to be corrected, how your perceptual systems represent the floor or how your walking, and so forth, are guided by the new representation produced? Frith et al. (2000) remark that patients who suffer from optic ataxia such that the arm fails to extend properly in space, the wrist fails to rotate to match the orientation of the object to be grasped, and the hand fails to open properly in anticipation of grasping often attribute their difficulty to a problem with vision. But this is perfectly natural, given that in the normal case there is no distinction between seeing things wrong and reaching and grasping for them wrong. For normal persons, which faculty it is that requires recalibration is an empty issue.

Exactly similarly, for the animal whose only criteria of identity are practical, correctly recognizing the object that is just over there in such and such direction simply *equals* knowing how to respond to or use that object, given this or that context of practical concern. If what is seen is an affordance, the criterion of correctness for the descriptive side is that it match the directive side, and vice versa. By the criterion of practice, if two objects function the same way when used the same way, they are the same; if they function differently or must be used differently, then they are different. For example, I imagine that for our cat, all dogs are roughly the same thing again, with the exception of our own dog, Thistle, who is the cat's friend and interacts with the cat entirely differently than other dogs do. So far as the cat is concerned, Thistle is of a totally different practical kind from other dogs. Thistle is not a (practical) dog. Similarly, I imagine that although birds are not mice for the cat because these need to be chased in different ways, the differences among mice, voles, shrews, chipmunks, and other small ground mammals are not noticed, because these differences are of no practical concern. What is the same as what, when you have met the same thing again, is entirely a practical matter for the practical animal, perhaps for all animals except humans. For these animals perceive the world only as a subject of practical concern, not as a subject of theoretical judgment. Objects that offer the same affordances will count as the same object, despite wide diversity in all kinds of properties we humans find important for theoretical purposes.

More radical, for the animal that tells when it has encountered the same thing again merely by practical tests, there will be no need to distinguish clearly among representations of individuals, practical kinds,

practical stuffs, and practical properties. If all one needs to understand is how to treat something the same when one encounters it again, there will be no fundamental practical differences among learning how to treat Thistle again, how to treat mouse again, or water again, or cold outside again, or hot underfoot again, or being wet and cold again. For practical purposes, each of these things merely returns again, sometimes in one place, sometimes in another, sometimes in more than one place at once. What use, for example, would the cat have for distinguishing between the mouse it catches on Tuesday and the mouse if catches on Wednesday? Why should it be any more a different versus the same mouse for the cat than the sun that rose yesterday and again today? Whether it's individuals, kinds, or properties, all show some differences, of course, from occasion to occasion of meeting. But only the practical overall similarities matter.[1]

1. The similarities and dissimilarities among concepts of individuals, natural kinds, and natural stuffs are explored in detail in Millikan (2000).

19 Conjectures on Human Thought

To learn to represent all sorts of objects, stuffs, kinds, and properties with which you have no practical dealings, you need another way to tell when you are reidentifying these things correctly.[1] You need a representational system that shows *within it*, prior to action, when errors in identification occur. This is done, I suggest, with the introduction of subject–predicate structure into representations where the predicate is sensitive to a negation transformation. This form, the form of theoretical judgment, allows descriptive inconsistencies to emerge explicitly, right on the surface of the representational system. It allows thought openly to display coherence or incoherence in the ways it is representing the world prior to using those representations in practical activity. Inconsistencies show that corrections are needed in the ways being used to form judgments, that is, in the ways used to identify subjects and predicates of judgment—objects, events, and their properties. Contrast beaver-tail slaps and bee dances, for example, which are not sensitive to a negation transformation, indeed, cannot even display contrariety. Danger signals at different times and places are not contrary to one another, for there might really be that much danger around. Bee dances showing nectar at different locations are not contrary to one another, nor do the bees have any way of saying where there isn't any nectar. A subject–predicate sentence and its negation, on the other hand, are explicitly incompatible, incompatible right on the surface. Similarly, humans can think negative thoughts, and these thoughts contrast explicitly with possible positive thoughts. Whether the way

1. This chapter concerns theoretical knowledge and the use of negation and contradiction. These themes are developed in more detail in Millikan (1984), chapters 14 through 19. Further aspects are developed in Millikan (2000), especially chapter 7. Interested readers may wish to turn to those earlier discussions, for the account below is much abbreviated.

human thoughts are coded resembles the way language is coded in any other way, it is clear that our thoughts are sensitive to a negation transformation.

This feature of thought, I believe, explains how humans are able to gather for possible use an enormous variety of representations of world affairs that they do not use in practice, indeed, that are very distant from them in time, space, and magnitude. We are able to do this because we can test each method of gathering information about a subject matter against its use on other occasions and against alternative methods of gathering the same information, using agreement in judgments to confirm our abilities to reidentify objects and properties. Consistent agreement in results is evidence that these various methods of making the same judgment are all focusing on the same distal affair, bouncing off the same target, as it were. But, of course, agreement in judgments can be a test only because disagreement in judgments is possible.

If the same belief is confirmed by sight, by touch, by hearing, by testimony, and by various inductions one has made, this is a good test not only for the objectivity of the belief but for each of the methods employed in identifying and reidentifying the objects and properties the belief concerns. The same object that is square as perceived from here should be square as perceived from there and square by feel and square by checking with a carpenter's square and square by measuring its diagonals. Both one's general methods of reidentifying individual physical objects and one's methods of recognizing shapes are corroborated in this way. Similarly, if a person knows French as found today, that person should know French when found tomorrow and as inferred from the fact that he buys *Le Mond* every Saturday. If the same belief is confirmed by sight, by touch, by hearing, and by testimony and is also in accord with theories one holds, that helps to confirm the accuracy of one's visual, tactile, and auditory perception as well as the accuracy of one's theories. That the same substance is found to melt at the same temperature by checking with an alcohol thermometer, a mercury thermometer, a gas thermometer, and a bimetal expansion thermometer is evidence both that one is able to recognize the same substance again and that there is indeed some one real quantity that all of these instruments are measuring. In sum, that any method of collecting evidence is in fact a method of collecting evidence *for* something can be confirmed only by a record of agreement with other methods of collecting evidence for the same. This sort of agreement is evidence

both for the objective reality of the subject matter and at the same time for the reliability of the methods used in reidentifying both subject and predicate.[2]

Now a crucial point is that the possibility of agreement in judgments presupposes the possibility of disagreement. In its basic form, negation is a semantic operation on the logical predicate of a sentence (Millikan 1984, chapter 14; Horn 1989, chapter 6).[3] Logicians call this "internal negation." For example, the normal reading, say, of the classic negative sentence "The king of France is not bald" makes it equivalent to "The king of France is nonbald," so that the negative as well as the affirmative presupposes the existence of a king of France. More obviously, "John is not tall" is normally equivalent to "John is nontall" and "John does not know French" is equivalent to "John is ignorant of French," and so forth. There are also secondary uses of "not" to reject a sentence on non-truth-conditional grounds, as in "The slithy toves did not gyre and gimbal in the wabe" or "The square root of two is not blue" or "You didn't see two mongeese, dear, you saw two mongooses" or "The king of France is not bald, dear; France doesn't have a king." But the fundamental use of the negative is not to prohibit assertion of a sentence, but to make a positive, though indefinite, statement to the contrary. The standard negative sentence says something about its subject, namely, that it is characterized by some contrary or other of the predicate of the sentence. If John is not tall it is because he is short or of medium height. If John does not understand French it is because French sentences either leave his mind blank or produce in it thoughts different than for a Frenchman.

This point about negation assumes importance when we turn to epistemology and consider how evidence is gathered for a negative judgment. Begin with the obvious: The absence of a representation of a

2. In Millikan (1984, 2000) I defend a strong realism or objectivism about sameness or identity at considerable length.

Wittgenstein and Davidson hold that the only way to corroborate one's ways of recognizing the selfsame thing again is through agreement with others using the same language. Wittgenstein believed this because he was a linguistic idealist. He didn't believe there were objective identities prior to language and the thought that rests on language. Why Davidson believes this is less clear. (When questioned, he just says that if you can't see the point, there really is no point in talking further.)

3. External negation, which operates on the sentence as a whole, is called "immunizing" negation in Millikan (1984). Horn (1989) gives a parallel analysis, calling it "metalinguistic" negation as opposed to "descriptive" negation. The claim is that immunizing or metalinguistic negation is not a semantic operator.

certain fact is not equivalent to the presence of a representation showing the negative of that fact. Absence of a belief is not a negative belief. Similarly, absence of perceptual evidence leading one to form or confirm a belief is not perceptual evidence that leads one to form or confirm the negative of that belief. If you look again from another angle at what you took to be a square object but fail this time to see that the object is square, or reach out with your hand but fail to feel that the object is square, this by itself is not evidence against the object's being square. Perhaps the trouble is that you can no longer see the object at all, or although you see it, you can't make out its shape against the light. Perhaps the trouble is that the object is not where it appeared to be so that reaching out your hand to feel it you encounter nothing at all. To gather evidence against the object's being square, you must first see or feel the object, and then you must see or feel that its shape is some contrary of square, perhaps round or oblong. Gathering evidence for the negative of a proposition is always gathering positive evidence, evidence for some contrary of that proposition.

It follows that the ability to recognize contraries of a property through the variety of their diverse manifestations and to recognize them as being contraries, as being incompatible, is required in order to test one's abilities to identify subjects of theoretical judgment, and vice versa. The result is not an epistemological regress or circle. But both of these abilities do have to be in place before stability of theoretical judgment over time and over perspectives can emerge with regard to any particular kind of subject matter. Both these abilities have to be in place before steady evidence can accumulate that any successful identifications at all are being made. The first leg up is undoubtedly practical. Many of the things recognized as the same again for purposes of practical use do turn out to be pretty good subjects for theoretical judgment as well. The second leg up, as I will explain in a minute, is public language.

On the other hand, the bootstrapping into theoretical judgment is made more difficult by the fact that different categories of things suitable to be subjects of judgment have properties from different contrary ranges. For example, although each person and each building has some definite height, silver and milk have no height any more than the square root of two has a color, and although the leopard frog, as studied by the zoologist, is cold blooded and has a heart and lungs and also spots, it has no definite number of spots. If one finds that the leopard frog has twenty-seven spots on one occasion and twenty-nine on another, this does not cast doubt either on one's ability to recognize

leopard frogs or on one's ability to count spots, for only individual frogs have a definite numbers of spots. Use of the law of contradiction to test your abilities to identify subjects of judgment and their properties thus presupposes a grasp of what kinds of contrary spaces are coordinate to what kinds of subjects. It assumes, that is, some understanding of the structure of various ontological categories involved such as the categories *individual object*, *species*, *functional kind*, *organic substance*, *chemical kind*, and so forth.[4]

Language provides a leg up, indeed, takes us most of the way up, in the enormously difficult task of learning to identify suitable subjects for theoretical judgment and the predicate contrary spaces that complement them. It does this, initially, by a very simple means. Every language has a small number of phonemes which, in various arrangements, account for all of the words in the language. Having learned the phonological structure of a language (which infants do in the first few months) makes it possible to tell when the same word is being said again and when a different word.[5] If the thesis of chapter 9 is correct and listening to language is just one more form of direct perception of the world, then objects and properties that have been discovered and named in one's language community are made immediately available to one through language. A denoting word is a tracer for whatever it denotes. It evidences the existence of an objective subject matter or of an objective property already recognized by others in the community. Learning to agree with others in making judgments is learning to identify what others already know how to identify. Grammar, as well as the judgments others make, serves as a guide to ontological category. In this way, accurate abilities to locate and reidentify various objects, stuffs, events, and their kinds and properties, abilities it may have taken the historical community hundreds of years to achieve, are acquired nearly effortlessly by later generations (Millikan 2000, chapter 6).

The additional perspective on the world that understanding a public language affords adds much more than just another sensory modality through which objective identities are easier to perceive.[6] Or perhaps I

4. These cryptic remarks are expanded in Millikan (2000).

5. Not quite. This ignores that different words can have the same sound, so that their domains often need to be tracked as well (see chapter 10).

6. That it is a mistake to count sensory modalities by counting the end organs through which information is received was strongly, and I believe correctly, argued by Alvin Liberman over a period of many years. In particular, he argued that the perception of speech sounds uses different neural channels than the perception of other sounds, the separation occurring very close to the periphery (Liberman 1996).

should put it the other way around—that much more is required of one who understands a public language than that they learn to perceive through a new sensory modality. An important difference between ordinary perception and perception through language is that perception through language does not routinely yield information about the relation of what is perceived to the perceiver (chapter 9). But information about something that you do not understand your own relation to is useless for immediate practical ends. In order to utilize a perceived affordance, you have to know where it is in relation to you. But the one who picks up information through ordinary perception and subsequently transmits it through language knows only his own relation to the subject matter, not the relation that hearers will have, and certainly not the relation that hearers of hearers will have.

Language, then, is a vehicle primarily for transmission of theoretical judgment and thought, and also for the representation of goal states expressed using theoretical concepts. It is reasonable to suppose, then, that the development of human language and the development of theoretical thought were coupled. A developed language would be of no use without the ability to engage in theoretical judgment and thought, and the ability to engage in theoretical judgment and thought would be difficult or impossible to sustain without language, since each generation would have to start fresh in the project of developing the sophisticated capacities to reidentify that are needed to support theoretical concepts.

Possibly the most important achievement of theoretical thought, resting directly on the capacity for language, is the capacity to represent historical time. By historical time I mean dated time, that is, time represented as a straight path receding into the past in one direction and continuing indefinitely into the future in the other without repetition. Contrast this with representing time merely as a set of unchanging conditional probabilities of temporal sequence, one kind of event following after another. The latter understanding of time makes it exactly like space. For the most part a space remains the same no matter where you move within it. You can leave a part of the space, go through a sequence of neighboring parts, later return to that same part again, finding it the same as you left it. Then you can go through the same sequence again, or through another sequence, then back to the first, and so forth. Similarly, a representation of time as a set of conditional probabilities of sequence allows one to come back to the same position in the sequence again, or to go through another sequence, then come

back to the first. A representation of the space one lives in has to be updated occasionally. The same is so for this kind of representation of time. Sometimes a recurring sequence changes to a somewhat different recurring sequence. This happens, for example, when a young mammal is weaned by its mother. But updating a representation is merely changing it. As Kant admonished Hume in the parallogisms, a change in one's representation of something is not a representation of change. To change one's representation of the space or the temporal order one lives in is merely to correct current errors. It is not to represent that anything has changed. Temporal contingency sequences, even though frequently updated, need not be understood as happening within a dated time order. They need not be understood as happening within historical time.

An understanding of historical time implies the capacity to understand particular events as occurring in just one position in a linear time sequence, never to be directly encountered again. It implies the capacity to think of an individual object as having a property at one particular time but not necessarily at any other. In order to develop and test one's abilities to identify single events of this kind as valid abilities to recognize truly objective events, one has to perceive the very same time-bound events in more than one way. How do we do this? What evidence do we have, for example, that our representations of the past are of anything real? The primary evidence is that others often remember the same events. It is through other people's perceptions that we obtain more than one perspective on the same dated occurrence. But other people's perceptions are made available to us only through language. Grasp of historical time depends, in the first instance, on language.

Chapter 16 left us with an animal that was able to represent objective goal states, able to know when it had reached these goal states, and able, even, to make trials and errors in thought so as to invent new ways of reaching its goals. But this animal still lived in an entirely present-centered world, its goals derived only from present perceptions of nearby affordances and perceptions of present needs. Its world of perceived possibilities unfolded entirely from within its momentary perceptual experience coupled with its own and its species' past history of successes and failures. Planning for the future was entirely instinct controlled in these animals, not thought about, not figured out. In chapter 18, however, we discovered ourselves as creatures that collect desires, dreams, and ambitions concerning all sorts of things we

ourselves have never experienced, sometimes even things humankind has never experienced. We found ourselves explicitly representing possible future affairs that, if brought about, would be quite new in our own experience and perhaps in others' as well. And we found ourselves explicitly worrying about how to accommodate future needs of ourselves and others, sometimes needs in the distant future. Where did these extra desires and concerns come from?

An animal that represents time as it represents space moves into the future as if navigating a terrain that is already there. Its job is to avoid the pitfalls and to seek out the rewards already laid out ahead, not to create anything new. We humans who represent time as historical understand that we are *constructing* a sequence, not finding one. Permanent changes can be made in the layout of the world, and on top of those changes further changes can later be made. New permanent structures can be constructed that will stay there for use tomorrow and the day after. Permanent changes can be made in the dispositions of things, for example, by altering a tool or a machine, so that it will behave more as one wishes in future. The capacity to represent historical time gives rise to our ability to conceive of, to plan, and to carry out long-term projects that significantly change our environments. We quite purposefully and knowingly make what will exist in the future quite different from what has existed in the past.

References

Akins, K. 1996. "Of sensory systems and the 'aboutness' of mental states." *Journal of Philosophy* 93.7: 337–372.

Balda, R. P., and A. C. Kamil. 1992. "Long-term spatial memory in Clark's nutcracker, *Nucifraga columbiana*." *Animal Behaviour* 44: 761–769.

Bargh, J. 1997. "Reply to the commentaries." In R. S. Wyer, Jr., ed., *The Automaticity of Everyday Life, Advances in Social Cognition* 10, 231–246. Mahweh, N.J.: Erlbaum.

Blackmore, S. 1999. *The Meme Machine*. Oxford: Oxford University Press.

Black, N. 1986. "Advertisement for a semantics for psychology." In P. French, T. Uehling, and H. Wettstein, eds., *Studies in the Philosophy of Mind*, Midwest Studies in Philosophy 10, 615–678. Minneapolis: University of Minnesota Press. Reprinted in S. P. Stich and T. A. Warfield, eds., *Mental Representation: A Reader*. Oxford and Cambridge, Mass.: Blackwell.

Böer, S. E., and W. G. Lycan. 1986. *Knowing Who*. Cambridge, Mass.: The MIT Press.

Brooks, R. A. 1999. *Cambrian Intelligence: The Early History of The New AI*. Cambridge, Mass.: The MIT Press.

Byrne, B. 1996. "The learnability of the alphabet principle: Children's initial hypotheses about how print represents spoken language." *Applied Psycholinguistics* 17: 401–426.

Byrne, R. W. 1999. "Imitation without intentionality: Using string parsing to copy the organization of behavior." *Animal Cognition* 2: 63–72.

Byrne, R. W. 2002. "Imitation of complex novel actions: What does the evidence from animals mean?" *Advances in the Study of Behavior* 31: 77–105.

Byrnes, J. P., and S. A. Gelman. 1991. "Perspectives on thought and language: Traditional and contemporary views." In Gelman and Byrnes (1991): 3–27.

Chisholm, R. M. 1967. "Brentano, Franz." In Paul Edwards, ed., *The Encyclopedia of Philosophy*, vol. 1, 365–368. New York: Collier Macmillan.

Chomsky, N. 1995. "Language and nature." *Mind* 104: 1–61.

Clark, E. V. 1991. "Acquisitional principles in lexical development." In Gelman and Byrnes (1991): 31–71.

Collet, T. S., and M. F. Land. 1978. "How hoverflies compute interception courses." *Journal of Comparative Physiology* 125: 191–204.

Colwill, R. M., and R. A. Rescorla. 1985. "Postconditioning devaluation of a reinforcer affects instrumental responding." *Journal of Experimental Psychology: Animal Behavior Processes* 11: 120–132.

Cummins, R. 1996. *Representations, Targets, and Attitudes*. Cambridge, Mass.: The MIT Press.

Cziko, G. 1995. *Without Miracles*. Cambridge, Mass.: The MIT Press.

Damasio, A. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Avon Books.

Davidson, D. 1968–1969. "On Saying That." *Sythese* 19: 130–146. Reprinted in A. P. Martinich, ed., *The Philosophy of Language*, fourth edition. Oxford: Oxford University Press, 2001.

Dawkins, R. 1976. *The Selfish Gene*. Oxford: Oxford University Press.

Dawkins, R. 1983. *The Extended Phenotype*. Oxford: Oxford University Press.

Dennett, D. C. 1984. *Elbow Room*. Cambridge, Mass.: The MIT Press.

Dennett, D. C. 1987. *The Intentional Stance*. Cambridge, Mass.: The MIT Press.

Dennett, D. C. 1996. *Kinds of Minds*. New York: Harper Collins.

Dennett, D. C. Forthcoming. "From typo to thinko: When evolution graduated to semantic norms." In S. Levinson and P. Jaisson, eds., *Culture and Evolution*. Cambridge, Mass.: The MIT Press.

Dilger, W. C. 1960. "The comparative ethology of the african parrot genus agapornus." *Zeitschrift für Tierpsychologie* 17: 649–685.

Dilger, W. C. 1962. "The behavior of lovebirds." *Scientific American* 206.1: 88–99.

Donald, M. 1991. *Origins of the Modern Mind*. Cambridge, Mass.: Harvard University Press.

Donnellan, K. 1966. "Reference and definite description." *Philosophical Review* 75: 281–304.

Dretske, F. 1981. *Knowledge and the Flow of Information*. Cambridge, Mass.: The MIT Press.

Dretske, F. 1986. "Misrepresentation." In Radu Bogdan, ed., *Belief: Form, Content, and Function*, 17–36. New York: Oxford University Press.

Dretske, F. 1995. *Naturalizing the Mind*. Cambridge, Mass.: The MIT Press.

Dretske, F. 1998. *Explaining Behavior*. Cambridge, Mass.: The MIT Press.

Ekman, P. 1980. *The Face of Man: Expressions of Universal Emotions in a New Guinea Village*. New York and London: Garland STMP Press.

Epstein, R., R. P. Lanza, and B. F. Skinner. 1981. "'Self-awareness' in the pigeon." *Science* 212.4495: 695–696.

Evans, G. 1982. *The Varieties of Reference*. Oxford: Clarendon Press.

Fodor, J. A. 2001. *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge, Mass.: The MIT Press.

Fodor, J. A., and E. Lepore. 1992. *Holism: A Shopper's Guide*. Oxford and Cambridge, Mass.: Blackwell.

Fodor, J. A., and E. Lepore. 1994. "Why meaning (probably) isn't conceptual role." In S. Stich and A. Warfield, eds., *Mental Representation: A Reader*, 142–156. Oxford: Blackwell. Reprinted from *Mind and Language* 6.4 (1991).

Frith, C., S. J. Blakemore, and D. Wolpert. 2000. "Abnormalities in the awareness and control of action." *Philosophical Transactions of the Royal Society London* B 355: 1771–1788.

Gallistel, C. R. 1980. *The Organization of Behavior*. Hillsdale, N.J.: Erlbaum.

Gallistel, C. R. 1990. *The Organization of Learning*. Cambridge, Mass.: The MIT Press.

Gelman, S. A., and J. P. Byrnes, eds. 1991. *Perspectives on Language and Thought*. Cambridge: Cambridge University Press.

Gibson, E. J. 1969. *Principles of Perceptual Learning and Development*. Englewood Cliffs, N.J.: Prentice-Hall.

Gibson, E. J. 1977. "The theory of affordances." In R. E. Shaw and J. Bransford, eds., *Perceiving, Acting, and Knowing*. Hillsdale, N.J.: Erlbaum.

Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gilbert, D. 1993. "The assent of man: Mental representation and the control of belief." In D. M. Wegner and J. W. Pennebaker, eds., *Handbook of Mental Control*. Englewood Cliffs, N.J.: Prentice-Hall.

Gleitman, H. 1991. *Psychology*, third edition. New York: W. W. Norton.

Gomez, J. C. 1991. "Visual behavior as a window for reading the mind of others in primates." In A. Whiten, ed., *Natural Theories of Mind: Development and Simulation of Everyday Mindreading*. Oxford: Blackwell.

Gould, J. 1982. *Ethology: The Mechanisms and Evolution of Behavior*. New York: W. W. Norton.

Gould, S. J. 1991. "Exaptation, a crucial tool for evolutionary psychology." *Journal of Social Issues* 47.3: 43–65.

Gould, S. J., and E. S. Vrba. 1982. "Exaptation—A Missing Term in the Science of Form." *Paleobiology* 8.1: 4–15.

Gould, S. J., and R. C. Lewontin. 1979. "The Spandrels of San Marco and the Panglossian program." *Proceedings of the Royal Society of London* 205: 281–288.

Hall, W. G., C. P. Cramer, and E. M. Blass. 1975. "Developmental changes in suckling of rat pups." *Nature* 258: 318–319.

Hall, W. G., C. P. Cramer, and E. M. Blass. 1977. "Ontogeny of suckling in rats: Transition toward adult ingestion." *Journal of Comparative and Physiological Psychology* 91: 1141–1155.

Herman, L. M. 2002. "Vocal, social, and self-imitation by bottlenosed dolphins." In C. Nehaniv and K. Dantenhahn, eds., *Imitation in Animals and Artifacts*. Cambridge, Mass.: The MIT Press.

Herman, L. M. Forthcoming. "Intelligence and rational behavior in the bottlenosed dolphin." In M. Nudds and S. Hurley, eds., *Rational Animals?* Oxford: Oxford University Press.

Herrnstein, R. J., D. H. Loveland, and C. Cable. 1976. "Natural concepts in pigeons." *Journal of Experimental Psychology: Animal Behavior Processes* 2: 285–302.

Hommel, B. 1993. "Inverting the Simon effect by intention: Determinants of direction and extent of effects of irrelevant spatial information." *Psychological Research/Psychologische Forschung* 55: 270–279.

Hommel, B., J. Müssler, G. Aschersleben, and W. Prince. 2001. "The theory of event coding (TEC): A framework for perception and action planning." *Behavioral and Brain Sciences* 24.5: 849–878. "Replies to Commentaries": 910–937.

Horn, L. R. 1989. *A Natural History of Negation*. Chicago and London: University of Chicago Press.

Hull, D. L., R. E. Langman, and S. S. Glenn. 2001. "A general account of selection: Biology, immunology and behavior." *Behavioral and Brain Sciences* 24.2: 511–569.

Jeannerod, M. 1997. *The Cognitive Neuroscience of Action*. Oxford: Blackwell.

Kegl, J., A. Senghas, and M. Coppola. 1999. "Creations through contact: Sign language emergence and sign language change in Nicaragua." In M. Degraff, ed., *Language Creation and Language Change: Creolization, Diachrony, and Development*, 179–237. Cambridge, Mass.: The MIT Press.

Levinson, S. C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lewontin, R. C. 1978. "Adaptation." *Scientific American* 239.3: 212–230.

Liberman, A. M. 1996. "Speech perception takes precidence over nonspeech perception." In A. M. Liberman, *Speech: A Special Code*. Cambridge, Mass.: The MIT Press.

Liberman, I. Y., D. Shankweiler, F. W. Fisher, and B. Carter. 1974. "Explicit syllable and phoneme segmentation in the young child." *Journal of Experimental Child Psychology* 18: 201–212.

Lorenz, K., and N. Tinbergen. 1939. "Taxis und Instinkthandlung in der Eirollbewegung der Graugans." *Teirpsychologie* 2: 1–29.

Lorenz, K. 1952. *King Solomon's Ring*. Marjorie Kerr Wilson, tr. London: Methuen.

Matzel, L. D., F. P. Held, and R. R. Miller. 1988. "Information and expression of simultaneous and backwards associations: Implications for contiguity theory." *Learning Motivation* 19: 317–344.

McFarland, D. 1987. *The Oxford Companion to Animal Behaviour*. Oxford: Oxford University Press.

McLeod, P., K. Plunkett, and T. Rolls. 1998. *Introduction to Connectionist Modeling of Cognitive Processes*. Oxford: Oxford University Press.

Michaels, C. F., and C. Carello. 1981. *Direct Perception*. Englewood Cliffs, N.J.: Prentice-Hall.

Milius, S. 2001. "Social cats." *Science News Online* 160.11.

Millikan, R. G. 1984. *Language, Thought, and Other Biological Categories*. Cambridge, Mass.: The MIT Press.

Millikan, R. G. 1990. "Truth rules, hoverflies, and the Kripke–Wittgenstein paradox." *Philosophical Review* 99.3: 323–353.

Millikan, R. G. 1993. *White Queen Psychology and Other Essays for Alice*. Cambridge, Mass.: The MIT Press.

Millikan, R. G. 1996. "Pushmi-pullyu representations." In James Tomberlin, ed., *Philosophical Perspectives* 9, 185–200. Atascadero, Calif.: Ridgeview. Reprinted in L. May and M. Friedman, eds., *Mind and Morals*, 145–161 (Cambridge, Mass.: The MIT Press, 1996).

Millikan, R. G. 1998. "Language conventions made simple." *Journal of Philosophy* 95.4: 161–180.

Millikan, R. G. 1999. "Historical kinds and the special sciences." *Philosophical Studies* 95.1–2 (The 1997 Oberlin Colloquium): 45–65.

Millikan, R. G. 2000. *On Clear and Confused Ideas*. Cambridge: Cambridge University Press.

Millikan, R. G. 2001a. "Purposes and cross-purposes: On the evolution of language and languages." *Monist* 84.3 (Special issue: *The Epidemiology of Ideas*, ed. Dan Sperber): 392–416. Reprinted in Dan Sperber, ed., *The Epidemiology of Ideas* (La Salle, Ill.: Open Court, 2003).

Millikan, R. G. 2001b. "The language-thought partnership: A bird's eye view." In Hans Johan Glock, ed., *Language and Communication* 21: 157–166.

Millikan, R. G. 2002. "Biofunctions: Two paradigms." In R. Cummins, A. Ariew, and M. Perlman, eds., *Functions: New Readings in the Philosophy of Psychology and Biology*, 113–143. Oxford: Oxford University Press.

Millikan, R. G. 2003. "In defense of public language." In L. Antony and N. Hornstein, eds., *Chomsky and His Critics*, 215–237. Oxford: Blackwell.

Millikan, R. G. Forthcoming. *Language: A Biological Model*. Oxford: Oxford University Press.

Neander, K. 1995. "Misrepresenting and malfunctioning." *Philosophical Studies* 79: 109–141.

Neisser, U. 2002. "The dorsal system and the ecological self." *Behavioral and Brain Sciences* 25.1: 114.

Nijhawan, R. 1994. "Motion exploitation in catching." *Nature* 370.6487: 256–257.

Norman, J. 2002. "Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches." *Behavioral and Brain Sciences* 25.1: 73–144.

Olton, D. S., and R. J. Samuelson. 1976. "Remembrance of places passed: Spatial memory in rats." *Journal of Experimental Psychology: Animal Behavior Processes* 2: 97–116.

Origgi, G., and D. Sperber. 2002. "Evolution, communication, and the proper function of language." In P. Carruthers and A. Chamberlain, eds., *Evolution and the Human Mind: Language, Modularity, and Social Cognition*. Cambridge: Cambridge University Press.

Quine, W. V. O. 1960. *Word and Object*. Cambridge, Mass.: The MIT Press.

Rao, S. C., G. Rainer, and E. K. Miller. 1997. "Integration of what and where in the primate prefrontal cortex." *Science* 276.5313: 821–824.

Reed, E. S. 1993. "The intention to use a specific affordance." In R. H. Wozniak and K. W. Fischer, eds., *Development in Context: Acting and Thinking in Specific Environments*, 45–76. Hillsdale, N.J.: Erlbaum.

Reed, E. S. 1982. "An outline of a theory of action systems." *Journal of Motor Behavior* 14: 98–134.

Reed, E. S. 1996. *Encountering the World: Toward an Ecological Psychology*. New York: Oxford University Press.

Rescorla, R. A. 1988. "Behavioral studies of Pavlovian conditioning." *Annual Review of Neuroscience* 11: 329–352.

Roberts, I. 1985. "Agreement parameters in the development of English modal auxiliaries." *Natural Language and Linguistic Theory* 3: 21–58.

Roberts, I., and A. Roussou. 2003. *A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.

Rumbaugh, D. M., M. J. Beran, and W. A. Hillix. 2000. "Cause effect reasoning in animals." In C. Heyes and L. Huber, eds., *The Evolution of Cognition*, 221–238. Cambridge, Mass.: The MIT Press.

Ryle, G. 1949. *The Concept of Mind*. London, New York: Hutchison's University Library.

Sellars, W. 1963. *Science, Perception, and Reality*. New York: The Humanities Press.

Sherzer, J. 1973. "Verbal and nonverbal deixis: The pointed lip gesture among the San Blas Cuna." *Language and Society* 2: 117–131.

Snell, W. 1964. "Kinship relations in Machiguenga." Master's thesis, Hartford Seminary, Hartford, Connecticut. Cited in J. R. Hurford, "The neural basis of predicate-argument structure." *Behavioral and Brain Sciences*, forthcoming.

Sober, E. 1984. *The Nature of Selection*. Cambridge, Mass.: The MIT Press.

Sperber, D., and D. Wilson. 1986. *Relevance: Communication and Cognition*. Cambridge, Mass.: The MIT Press.

Sperling, G. 1960. "The information available in brief visual presentations." *Psychological Monographs* 74 (11, whole no. 498).

Sterelney, K. 1995. "Basic Minds." *Philosophical Perspectives* 9: 251–270.

Sterelney, K. 2001. *The Evolution of Agency and Other Essays*. Cambridge: Cambridge University Press.

Swinney, D. 1979. "Lexical access during sentence comprehension: (Re)consideration of context effects." *Journal of Verbal Learning and Verbal Behavior* 18: 645–659.

Tinbergen, N. 1951. *The Study of Instinct*. Oxford: Oxford University Press.

Tomasello, M. 2000. "Two hypotheses about primate cognition." In C. Heyes and L. Huber, eds., *The Evolution of Cognition*, 165–183. Cambridge, Mass.: The MIT Press.

Watson, J. S. 1967. "Memory and 'contingency analysis' in infant learning." *Merrill-Palmer Quarterly* 13: 55–76.

Waxman, S. R. 1991. "Semantic and conceptual organization in preschoolers." In Gelman and Byrnes (1991): 107–145.

Wheeler, S. 1972. "Attributives and their modifiers." *Noûs* 6.4: 310–334.

Whiten, A., V. Homer, and S. Marshall-Pescini. Forthcoming. "Selective imitation in child and chimpanzee: A window on the construal of others' actions." In S. Hurley and N. Chater, eds., *Perspectives on Imitation: From Cognitive Neuroscience to Social Science*, vol 1. Cambridge, Mass.: The MIT Press.

Zipser, D., and R. A. Anderson. 1988. "A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons." *Nature* 331.6158: 679–684.

Index