Introduction

What is a mental representation? It is convenient to divide classical answers into three types: picture theories, causal and/or informational theories, and what I call "PMese theories" ("PMese" is Wilfrid Sellars's term for symbolic logic, from, I suppose, Principia mathematica). Each of these three theories of representation, if taken bare, runs into exactly the same problem, namely the problem of accounting fcr cognitive errors: misperceptions, false beliefs, confused concepts, bad inferences, unrealized intentions, and so forth. Each founders over the distinction between the facts of cognition and the norms of cognition. Call this the normativity problem. The underlying theme of the essays in this volume is that all of the basic norms applying to cognition are biological norms. Correct cognition, like healthy digestion and successful reproduction, is not what happens always or on the average. Correct cognition is what has accounted for the survival in the species of the complex mechanisms that are, at least occasionally and perhaps under ideal conditions, responsible for it. Let me begin by characterizing each of the three classical approaches to understanding mental representation in order to show at what point each runs into the normativity problem. I then say just a word to indicate how introducing biological norms helps with this problem. With this help, each of the three theories turns out to play an important role in the solution to the problem of what mental representation is.

1 Picture Theories

Picture theories of representation are based on the view that representations are like what they represent. More formally, a rule of projection

or isomorphism, roughly in the mathematical sense, is what relates a representation to what it represents. This rule of projection is a function that maps members from a whole domain of representations onto their respective representeds. Any projection rule defines at least a formal similarity between the items in its domain and its range. Picture theorists thus propose that the defining relation between representation and represented depends on likeness, either likeness of an ordinary sort or perhaps likeness of a very abstract, formal mathematical sort. For example, the classical British empiricists tended to hold crude picture theories of how "ideas" represent, taking the likeness to be quite an ordinary sort of likeness, whereas Wittgenstein in the *Tractatus* held a picture theory of how thoughts that or beliefs that represent, taking the likeness to be structural likeness of a very abstract, formal character.

According to what can be called "pure" picture theories, the fact that the represented is *like* the representation is what *constitutes* the relation between representation and represented; that is all it consists in. *Pure* picture theories are thus easily criticized by pointing out that there is an infinite number of kinds of likeness. Everything is like everything else in one respect or another, indeed, in many respects. Thus the question arises, Which of these kinds of likeness constitutes the relation between representation and represented, and why? Clearly, *something* more must be added to make a picture theory viable.

A second failing of pure picture theories not so often commented on is that they seem not to leave room for misrepresentation. If representing something is picturing it, then representing it wrongly should be picturing it wrongly. Presumably, picturing it wrongly would be picturing it with something that was not a likeness of it according to the given rule. But if representing it at all depends upon the designated likeness relation holding between it and its representation, representing it wrongly would then be indistinguishable from not representing it at all. For example, suppose that what makes my drawing a drawing of my cat is that it resembles my cat in accordance with a certain rule of projection onto two dimensions from the actual three-dimensional shape of my cat. Then if my drawing in fact fails to accord with my cat by this rule, it will turn out not to be picture of my cat. Perhaps it will be a picture of someone else's cat, perhaps of someone's in Tibet, that it does happen to resemble. But what it cannot be is a bad picture of my cat, be-

cause it won't be a picture relation is no more and when the likeness fails, be tion between nonreprese collapse. Nor can it be a senting at all and exact libad drawing of a cat will it roughly resembles. Son by the rule, but surely it adequate picture theory of ground the relation between way that only truth and upon whether there is sor the designated rule.

What, then, makes a surely not its likeness to cat in accordance with a picture of the cat is the intended it to be like a crayoned pictures, but thought onto its object require one to have a would lead to a regress of course, intend to present its designed to present of projection rule, be scratched or aberrar tions can be abnormal feven unrecognizable. T

1. Though it has many va problem, and it is not ever mediate correspondents of not flesh-and-blood repre Tractarian senses, Fregeat Each of these shades is nestations to map onto o If representing is picturing nothing real that the represental.

es a ction their rmal heor-1 and ort or or exicture ite an held a ng the er. ct that elation 1. Pure

re is an

! else in

luestion

etween

g more

ented on epresent-d be pic-cturing it wen rule. ness relawrongly t all. For my cat is projection spe of my / this rule, a picture does hapny cat, be-

cause it won't be a picture of my cat at all. If the "of" or "being-about" relation is no more and no less than a certain likeness relation, then when the likeness fails, being about should fail as well. Thus the distinction between nonrepresenting and wrongly representing threatens to collapse. Nor can it be repaired by requiring rough likeness for representing at all and exact likeness for correctly representing, for then my bad drawing of a cat will be about lots of people's cats at once, all those it roughly resembles. Some of these cats it may indeed resemble exactly by the rule, but surely it will not on that account be a good picture. An adequate picture theory of mental representation would have instead to ground the relation between representation and represented in such a way that only truth and falsity, and not representing itself, depends upon whether there is something onto which the representation maps by the designated rule.

What, then, makes a two-year-old's picture a picture of a cat? It is surely not its likeness to the cat, even though it may in fact be like the cat in accordance with some arcane projection rule. What makes it a picture of the cat is the two-year-old's intention in drawing it. She intended it to be like a cat. This works fine with two-year-olds and crayoned pictures, but an act of intending cannot be what hooks a thought onto its object. Intending requires thought, and thus would require one to have a prior thought of the intended object, which would lead to a regress. Now consider a camera. A camera does not, of course, intend to produce a likeness. Rather, it is an instrument that is designed to produce likenesses in accordance with a certain sort of projection rule. And a camera can fail in this: its lense can be scratched or aberrant, the film can be warped, the lighting conditions can be abnormal for its use, so that the picture comes out wrong, even unrecognizable. Take one more step now. We design a camera

^{1.} Though it has many variations, there is exactly one classic solution for this problem, and it is not even remotely satisfactory. That is to postulate as the *immediate* correspondents of all respresentations, or all complex representations, not flesh-and-blood representeds but shades: objectified concepts or ideas, *Tractarian* senses, Fregean thoughts, propositions, possibilities, and so forth. Each of these shades is nothing more nor less than just something for representations to map onto or picture when they don't map onto anything real. If representing is picturing, in the minimal sense of mapping onto, but there is nothing real that the representation maps onto, it must map onto something unreal.

to produce representations, in accordance with certain rules of projection, by a certain method of physical interaction with the represented. Might not evolution have designed our brains to produce representations, in accordance with certain highly abstract rules of projection, by certain methods of physical interaction with what is represented? Could it be that evolutionary design is what hooks a percept or a thought onto the state of affairs it represents? This is the point I will argue.

2 Causal/Informational Theories

Contemporary causal and/or informational theories of mental representation have at least two sources of inspiration. That thoughts are of their causes is a classical view, best known, perhaps, from Descartes and Locke, who wished to know which of our thoughts are like their causes, their causes being supposed to be what they are thoughts of. The second model for causal/informational theories is the natural sign, an object or event type that is a reliable indicator of another type of object or event, due perhaps to a causal connection between the two. Thus lion spoor is a sign of a lion having passed by because lion spoor are found only when lions have passed by. Similarly, on this view, the photograph of a cat represents a cat not because it is like a cat but because things of its type are never seen except when cats have been about to cause them. Sometimes proponents of causal/informational theories hold that a mental representation represents whatever it is that it never occurs without, just as lion spoor represents lions because it never occurs without lions. Sometimes they hold that it represents whatever it covaries with or "tracks," just as the "on" light on an appliance represents whether the appliance is on because it covaries with or tracks when the appliance is off and when on. Mixed causal/informational and picture theories of representation are, of course, also possible and have been held under the label "reliability theories." Here the "picturing" part of the theory is just "compositionality"; the projection rules are Tarskian. According to such theories, what a representation represents is what it maps onto and at the same time reliably indicates.

to the second second

Both Jerry Fodor and Fred Dretske, two committed causal/informational proponents, have emphasized that such theories have to face a normativity problem: the problem of "misrepresentation" (Dretske 1986) or "the disjunction problem" (Fodor 1987, 1990). Misrepresentation or false representation is, after all, exactly the occurrence of a representation without co-occurrence of its represented. Thus if the night is dark, skunks rather than cats can cause (mistaken) beliefs about cats. Does it follow that thoughts about cats are really thoughts about cats-or-skunks, since cat-or-skunk is all that they reliably indicate? A possible solution is to suppose that there are some special conditions under which there actually occurs the ideal dependency or covariation needed to define what represents what and that misrepresentation occurs only when these conditions are absent. Indeed, quite a lot of effort has gone into trying to delimit such a set of conditions under the label, for example, of "normal conditions."

Rather than attempting to specify a set of special content-determining conditions under which, by definition, whatever is perfectly indicated by a thought or percept is what is represented by it, suppose that we take seriously that indicating correctly is a biological norm for our cognitiveperceptual systems. Mental representations do not occur without their representeds so long as our perceptual-cognitive systems are operating correctly, i.e., in the way that has accounted for their survival in the species. Mental representations are accompanied by their appropriate representeds when these systems produce inner structures that map the environment in accordance with specified projection rules in the manner that accords with their design. One can determine which conditions are normal for the correct operation of these systems only by first specifying how these systems were designed to work, what projection rules they were designed to instantiate, what methods they were designed to use to effect the mappings. By itself this suggestion does not provide a definition either of representation or of representational content. But it explains how—through confusion of a biological norm with a simple disposition or fact—causal/informational theories achieve their initial plausibility.²

2. For a more accurate statement of the relation between causal/informational theories of representation and the one advocated here, see especially chapter 4.

3 PMese Theories

The PMese theory of mental representation is the classical theory of the twentieth century. It was inspired by the idea of a purely formal system, a system consisting of initially uninterpreted physical characters manipulated by rules operating over their physical forms. This idea was first applied in an attempt to understand, within a classical empiricist frame (with reference to theories of the origin of ideas), how "theoretical" terms in science, terms for things not observable, acquire meaning. The result was then generalized to apply to ordinary observation terms. These were taken to bear roughly the same relation to sense data or sensedata terms that theoretical terms had been taken to bear to observation terms.

The symbols in a purely formal system, it was thought, derive their meaning solely from the formal rules that govern their manipulation. For example, a formal symbol for "and," say '&," may derive its meaning from the following rules: (1) if one has written down the shape "p & q," one may then write down the shape "p," (2) one may also then write down the shape "q," and (3) if one has written down "p" and written down "q," one may then write down "p & q." That a symbol might derive its meaning from the rules of inference governing it, rather than the rules of valid inference being derived from prior meanings of the terms, was an entirely new idea, and all was swept before it. According to this new approach, the data of sense are not copied to produce mental representations, as supposed by British empiricists and phenomenalists. Rather, the data of sense are responded to with thoughts or sentences under the breath or inner symbols in a rulegoverned way, just as the axioms and hypotheses in a formal system are responded to in a rule-governed way. The inner representations that result are in turn responded to by rule, to produce more inner representations and then output responses, the meanings of the inner representations being created through this rule-governed process. For example, the meaning of one's mental representation "cat" is derived from such sources as that thinking "cat" is a result of certain kinds of retinal stimulations; that thinking "is a cat" produces thinking "is an animal," "is a mammal," "likes fish," and so forth; that thinking "I wish to pet a cat," when accompanied by catish retinal stimulations, produces

reaching gestures; and so forth. To be caught up in such a system of rules is what it is for something to be an inner representation with a meaning. Thus the mid-twentieth-century vision emerged of thought as a "theory" in the Quinean, Sellarsian sense. According to the PMese theory, inner representations are by their very nature things that are "calculated over" (Fodor) or that are "interanimated" (Quine).

PMese theories of mental representation are not necessarily at odds with picture theories. They are based on the claim that the role in a system of a certain kind—the most common idiom is "conceptual role" or "causal role"—determines representationhood and also representational content. But this might be so precisely because the role determines that there is a relevant picturing or mapping relation (likeness) between mental representings and the world. Sellars thought so, that the inference patterns executed in one's head are isomorphic to certain patterns in the world and that this is what produces mental "representing." Quine, on the other hand, famously said that, no, there is no determinate or unique mapping of any relevant kind from PMese-theory representations to the world, and hence that picturing reality or corresponding to reality cannot ground or be what reference and truth are about. Mixed causal/PMese theories are also possible and have been held under the label "dual aspect semantics" (e.g., Block 1986).

The normativity problem arises for PMese theories in a very direct way. If the actual inference dispositions that a thinker has are partial or full determinants of the meanings of the representations calculated over during inference, rather than the meanings of the representations determining which inferences are the valid ones, how can any invalid inferences occur? Or, to turn to input dispositions, since the dispositions that the thinker actually has to respond with inner symbols to the data of sense are determinants of the meaning of those very symbols, how can mistakes in recognizing occur?⁵

3. According to Sellars, "representing" is a relation between language and world, although "meaning" and "referring" are not.

world, although meaning and referring are not.

4. Besides, as Quine saw it, the picture theory would require introducing "creatures of darkness" into our ontology—propositions and possible men in the doorway—shadows for *misrepresentations* to map onto. See note 1.

5. Of course, if one is a Quinean holist, one believes that there is no such thing as a mistake that determinately concerns just a particular perception, belief, or inference. There are no particular correspondence relations to the world that

4 Solving the Normativity Problem

One popular contemporary way to try to solve the normativity problem is to hope that if we add up several loose constraints on what a person's inner representations mean, none of which taken singly could account for normativity, together they will do the trick. This is meant somehow to yield tighter constraints and a best fit all around, and thus *pretty* determinately to fix the meanings of mental representations. If that is the reader's inclination, then she will find it best to read the second half of the last chapter of this book first, for there I try hardest to show how vain such a hope is.

Sellars took an answer to the normativity problem from Wittgenstein. According to Sellars, inference norms come from community; inference rules are just internalized rules of the public language game. I do not think that thought is a social construction, nor do I believe that language use is governed by rules of the sort envisioned by Wittgenstein. Our language is in deep respects just like any other animal's communication system, its forms embodying analogues of evolutionarily stable solutions to coordination problems. Its functions derive from what its forms—words, syntactic structures, and so forth—do for speakers and hearers so that they continue to be reproduced by speakers and consistently understood by hearers. These things are discussed in LTOBC, chaps. 1–4, and are not touched on in this book.

My purpose in this book is to clarify, defend, and show some of the implications of a biological solution to the normativity problem. Picturing, indicating, and inference are equally involved in human representing, but as biological norms rather than as mere dispositions. It is not the facts about how the system *does* operate that make it a representing

make thinking a productive enterprise. Rather, error is supposed to consist just in the emergence here or there of contradictions. But it is not clear on such a view what contradictions are beyond something one finds unpleasant and tries to get rid of. Or if one takes the view that the content of thought is determined not just by input and inference dispositions but also by behavioral output dispositions, then error would seem to consist roughly in the whole system of thought failing to produce the good life. Without invoking correspondence of some sort, the problem of how to tell good reasoning from good digestion and good sex looms large.

system and determine what it represents. Rather, it is the facts about what it would be doing if it were operating according to biological norms. When functioning properly, a mental representation⁶ co-occurs with its represented, pictures what it represents, and (if it is of the right, rather sophisticated sort) participates in appropriate inferences.

Roughly indeed, the idea is this. Cognitive systems are designed by evolution to make abstract pictures of the organism's environment and to be guided by these pictures in the production of appropriate actions. This may involve deriving additional pictures by combining old ones in accordance with certain principles (inference, reasoning). It may also involve the formation of goal representations, pictures that guide the organism to produce or avoid what they picture (imagining, fearing that, desiring, intending). In the human case, biological design includes very general principles in accordance with which concepts are formed from appropriate sensory input, including input from perception through language. Forming adequate concepts is learning to represent or map in thought what is the same again as the same. On this ability are built the abilities to make judgments, then appropriate inferences, and also the abilities to acquire appropriate desires and get them fulfilled. Like anything else the organism does, this all takes place in a highly principled way when it works, but, of course, it does not always work, or if it does, not always smoothly.

The basic plan for this sort of solution to the normativity problem is laid out in LTOBC. This book does not in any way supplant that volume but rather rests upon it at numerous pivotal points. To put it straightforwardly, this book is designed to be read independently and, I hope, far more easily than LTOBC, but it is not designed to be criticized apart from the foundations laid in LTOBC. In the first two chapters I discuss and defend the theory of biological function needed for the project, and I defend the claim that our brains are adaptations for thinking, indeed, adaptations for thinking in the very ways in which we presently do think. I then clarify and defend a theory of mental representation. Implications for the science of psychology follow: that psychology is not at root a science involving laws, that explanations in psychology are unlike explanations in the physical sciences, that it in-

Burkey Commencer

^{6. &}quot;Indicative" ones, anyway.

escapably is a deeply ecological science dealing with how the organism interacts with its wider environment.

Next in chapters 10 and 11 comes discussion of some worries dear only to philosophers (thank goodness). I show that the "realism" involved in the flatfooted correspondence views of representation and truth put forth here withstands antirealist critiques. By substituting a theory of biological competence for the standard reliance on dispositions in the construction of theories of meaning and truth, I provide a theory of understanding language that undercuts the verificationist's moves against realism. In chapter 12, I open the possibility of a naturalist treatment of the nature of knowledge based on a biologically rooted theory of competence.

Chapter 13, "The Myth of the Essential Indexical," was originally part of chapter 14, "White Queen Psychology," but was torn out at the late Hector-Neri Castañeda's request to appear in (what he knew would be) his last special issue volume of Noûs, on reference. Who but Castañeda would have sought out and insisted on publishing in his own journal, moreover, on adding his own complimentary remarks to, a paper that attacked one of his best known and favorite ideas? (I do not mean to imply that he agreed with me.) "White Queen Psychology" is a diatribe against meaning rationalism, as was the Epilogue of LTOBC. My desire is to kill meaning rationalism dead, and then beat on it. Perhaps I will succeed in raising one or two more doubts about it. Reasoning, I insist, is done in the world, not in one's head. Logical possibility (known a priori) is impossible. And the only hope for intentional psychology is to embrace its biological roots.

Biosemantics

Causal or informational theories of the semantic content of mental states which have had an eye on the problem of false representations have characteristically begun with something like this intuition. There are some circumstances under which an inner representation has its represented as a necessary and/or sufficient cause or condition of production. That is how the content of the representation is fixed. False representations are to be explained as tokens that are produced under other circumstances. The challenge, then, is to tell what defines certain circumstances as the content-fixing ones.

1

Note that the answer cannot be just that these circumstances are *statistically* normal conditions. To gather such statistics, one would need to delimit a reference class of occasions, know how to count its members, and specify description categories. It would not do, for example, just to average over conditions-in-the-universe-any-place-any-time. Nor is it given how to carve out relevant description categories for conditions on occasions. Is it "average" in the summer for it to be (precisely) between 80 and 80.5 degrees Fahrenheit with humidity 87 percent? And are average conditions those which obtain on at least 50 percent of the occasions, or is it 90 percent? Depending on how one sets these parameters, radically different conditions are "statistically normal." But the notion of semantic content clearly is not relative, in this manner, to arbitrarily determined.

A number of recent writers have made an appeal to teleology here, specifically to conditions of normal function or well functioning of the systems that produce inner representations. Where the represented is R and its representation is "R," under conditions of well functioning, we might suppose, only Rs can or are likely to produce "R"s. Or perhaps "R" is a representation of R just in case the system was designed to react to Rs by producing "R"s. But this sort of move yields too many representations. Every state of every functional system has normal causes, things that it is a response to in accordance with design. These causes may be proximate or remote, and many are disjunctive. Thus, a proximate normal cause of dilation of the skin capillaries is certain substances in the blood, more remote causes include muscular effort, sunburn, and being in an overheated environment. To each of these causes the vascular system responds by design, yet the response (a red face), though it may be a natural sign of burn or exertion or overheating, certainly is not a representation of that. If not every state of a system represents its normal causes, which are the states that do?

ł

(

1

}

Jerry Fodor (1986a) has said that, whereas the content of an inner representation is determined by some sort of causal story, its status as a representation is determined by the functional organization of the part of the system which uses it. There is such a thing, it seems, as behaving like a representation without behaving like a representation of anything in particular. What the thing is a representation of is then determined by its cause under content-fixing conditions. It would be interesting to have the character of universal I-am-a-representation behavior spelled out for us. Yet, as Fodor well knows, there would still be the problem of demonstrating that there was only one normal cause per representation type.

A number of writers, including Dennis Stampe (1979), Fred Dretske (1986), and Mohan Matthen (1988), have suggested that what is different about effects that are representations is that their function is, precisely, to represent, "indicate," or "detect." For example, Matthen says of (fullfledged) perceptual states that they are "state[s] that [have] the function of detecting the presence of things of a certain type" (1988, 20). It does not help to be told that inner representations are things that have representing (indicating, detecting) as their function, however, unless we are also told what kind of activity representing (indicating,

Andrew Commencer and the second of the commencer of the c

detecting) is. Matthen does not tell us how to naturalize the notion "detecting." If "detecting" is a function of a representational state, it must be something that the state effects or produces. For example, it cannot be the function of a state to have been produced in response to something. Or does Matthen mean that it is not the representational states themselves but the part of the system which produces them which has the function of detecting? It has the function, say, of producing states that correspond to or covary with something in the outside world. But unfortunately, not every device whose job description includes producing items that vary with the world is a representation producer. The devices in me that produce calluses are supposed to vary their placement according to where the friction is, but calluses are not representations. The pigment arrangers in the skin of a chameleon, the function of which is to vary the chameleon's color with what it sits on, are not representation producers.

Stampe and Dretske do address the question what representing or (Dretske) "detecting" is. Each brings in his own description of what a natural sign or natural representation is, then assimilates having the function of representing R to being a natural sign or representer of R when the system functions normally. Now the production of natural signs is undoubtedly an accidental side effect of normal operation of many systems. From my red face you can tell that either I have been exerting myself or I have been in the heat or I am burned. But the production of an accidental side effect, no matter how regular, is not one of a system's functions; that goes by definition. More damaging, however, it simply is not true that representations must carry natural information. Consider the signals with which various animals signal danger. Nature knows that it is better to err on the side of caution, and it is likely that many of these signs occur more often in the absence than in the presence of any real danger. Certainly there is nothing incoherent in the idea that this might be so, hence that many of these signals do not carry natural information concerning the dangers they signal.

2

I fully agree, however, that an appeal to teleology, to function, is what is needed to fly a naturalist theory of content. Moreover, what makes a

thing into an inner representation is, near enough, that its function is to represent. But, I shall argue, the way to unpack this insight is to focus on representation consumption, rather than representation production. It is the devices that use representations that determine these to be representations and, at the same time (contra Fodor), determine their content. If it really is the function of an inner representation to indicate its represented, clearly it is not just a natural sign, a sign that you or I looking on might interpret. It must be one that functions as a sign or representation for the system itself. What is it, then, for a system to use a representation as a representation?

The conception of function on which I shall rely was defined in LTOBC and defended in chapter 1 under the label "proper function." Proper functions are determined by the histories of the items possessing them; functions that were "selected for" are paradigm cases. The notions "function" and "design" should not be read, however, as referring only to origin. Natural selection does not slack after the emergence of a structure but actively preserves it by acting against the later emergence of less fit structures. And structures can be preserved due to performance of new functions unrelated to the forces that originally shaped them. Such functions are "proper functions" too and are "performed in accordance with design."

The notion "design" should not be read—and this is very important—as a reference to innateness. A system may have been designed to be altered by its experience, perhaps to learn from its experience in a prescribed manner. Doing what it has learned to do in this manner is then "behaving in accordance with design" or "functioning properly" (LTOBC and chapters 3, 11 herein).

My term "normal" should be read normatively, historically, and relative to specific function. In the first instance, "normal" applies to explanations. A "normal explanation" explains the performance of a particular function, telling how it was (typically) historically performed on those (perhaps rare) occasions when it was properly performed. Normal explanations do not tell, say, why it has been common for a function to be performed; they are not statistical explanations. They

^{1.} An odd custom exists of identifying this sort of view with Larry Wright, who does not hold it. See chapter 1. Genetic selection is not the only source of proper functions. See *LTOBC*, chaps. 1 and 2.

cover only past times of actual performance, showing how these performances were entailed by natural law, given certain conditions, coupled with the dispositions and structures of the relevant functional devices.2 In the second instance, "normal" applies to conditions. A "normal condition for performance of a function" is a condition the presence of which must be mentioned in giving a full normal explanation for performance of that function. Other functions of the same organism or system may have other normal conditions. For example, normal conditions for discriminating colors are not the same as normal conditions for discriminating tastes, and normal conditions for seeing very large objects are not the same as for seeing very small ones. It follows that "normal conditions" must not be read as having anything to do with what is typical or average or even, in many cases, at all common. First, many functions are performed only rarely. For example, very few wild seeds land in conditions normal for their growth and development, and the protective colorings of caterpillars seldom actually succeed in preventing them from being eaten. Indeed, normal conditions might almost better be called "historically optimal" conditions. (If normal conditions for proper functioning, hence survival and proliferation, were a statistical norm, imagine how many rabbits there would be in the world.) Second, many proper functions only need to be performed under rare conditions. Consider, for example, the vomiting reflex, the function of which is to prevent (further) toxification of the body. A normal condition for performance of this function is presence, specifically, of poison in the stomach, for (I am guessing) it is only under that condition that this reflex has historically had beneficial effects. But poison in the stomach certainly is not an average condition. (Nor, of course, is it a normal condition for other functions of the digestive system.3)

^{2.} This last clarification is offered to aid Fodor ("On there not being an evolutionary theory of content," hereafter "NETC," unpublished), who uses my term "Normal" (here I am not capitalizing it, but the idea has not changed) in a multiply confused way, making a parody of my views on representation. In this connection, see also nn. 5 and 9.

^{3. &}quot;Normal explanation" and "normal condition for performance of a function," along with "proper function," are defined with considerable detail in LTOBC. The reader may wish, in particular, to consult the discussion of normal explanations for performance of "adapted and derived proper functions" in

If it is actually one of a system's functions to produce representations, as I have said, these representations must function as representations for the system itself. Let us view the system, then, as divided into two parts or two aspects, one of which produces representations for the other to consume. What we need to look at is the consumer part, at what it is to use a thing as a representation. Indeed, a good look at the consumer part of the system ought to be all that is needed to determine not only representational status but representational content. I argue this as follows. First, the part of the system which consumes representations must understand the representations proffered to it. Suppose, for example, that there were abundant "natural information" (in Dretske's, 1981, sense) contained in numerous natural signs all present in a certain state of a system. This information could still not serve the system as information, unless the signs were understood by the system, and, furthermore, understood as bearers of whatever specific information they in fact do bear. (Contrast Fodor's notion that something could function like a representation without functioning like a representation of anything in particular.) So there must be something about the consumer that constitutes its taking the signs to indicate, say, p, q, and rrather than s, t, and u. But if we know what constitutes the consumer's taking a sign to indicate p, what q, what r, etc., then, granted that the consumer's takings are in some way systematically derived from the structures of the signs so taken, we can construct a semantics for the consumer's language. Anything the signs may indicate qua natural signs or natural information carriers then drops out as entirely irrelevant; the representation-producing side of the system had better pay undivided attention to the language of its consumer. The sign producer's function will be to produce signs that are true as the consumer reads the language.

The problem for the naturalist bent on describing intentionality, then, does not concern representation production at all. Although a representation always is something that is produced by a system whose prop-

er funworld, function

For the fol accord use of, the co accord these 1 way.) sented functio Note t on the do. Th sentati tent ha for per respon functio for m tokens their t under corresp Just w

4. Strice explana a more under to 5. In the certain fulfillm function that a lerly be a different expression.

chap. 2 of that work, for these functions cover functions of states of the nervous system which result in part from learning, such as states of human belief and desire.

er function is to make that representation correspond by rule to the world, what the rule of correspondence is, what gives definition to this function, is determined entirely by the representation's consumers.

For a system to use an inner item as a representation, I propose, is for the following two conditions to be met. First, unless the representation accords so (by a certain rule) with a represented, the consumer's normal use of, or response to, the representation will not be able to fulfill all of the consumer's proper functions in so responding-not, at least, in accordance with a normal explanation. (Of course, it might still fulfill these functions by freak accident, but not in the historically normal way.) Putting this more formally, that the representation and the represented accord with one another so is a normal condition for proper functioning of the consumer device as it reacts to the representation.4 Note that the proposal is not that the content of the representation rests on the function of the representation or of the consumer, on what these do. The idea is not that there is such a thing as behaving like a representation of X or as being treated like a representation of X. The content hangs only on there being a certain condition that would be normal for performance of the consumer's functions, namely, that a certain correspondence relation hold between sign and world, whatever those functions may happen to be. For example, suppose the semantic rules for my belief representations are determined by the fact that belief tokens in me will aid the devices that use them to perform certain of their tasks in accordance with a normal explanation for success only under the condition that the forms or "shapes" of these belief tokens correspond, in accordance with said rules, to conditions in the world. Just what these user tasks are need not be mentioned.5

4. Strictly, this normal condition must derive from a "most proximate normal explanation" of the consumer's proper functioning. See *LTOBC*, chap. 6, where a more precise account of what I am here calling "representations" is given under the heading "intentional icons."

and with the control of the control

^{5.} In this particular case, one task is surely contributing, in conformity with certain general principles or rules, to practical inference processes, hence to the fulfillment of current desires. So, if you like, all beliefs have the *same* proper function. Or, since the rules or principles that govern practical inference dictate that a belief's "shape" determines what other inner representations it may properly be combined with to form what products, we could say that each belief has a different range of proper functions. Take your pick. Compare Fodor 1989 and "NETC."

Second, represented conditions are conditions that vary, depending on the form of the representation, in accordance with specifiable correspondence rules that give the semantics for the relevant system of representation. More precisely, representations always admit of significant transformations (in the mathematical sense), which accord with transformations of their corresponding representeds, thus displaying significant articulation into variant and invariant aspects. If an item considered as compounded of certain variant and invariant aspects can be said to be "composed" of these, then we can also say that every representation is, as such, a member of a representational system having a "compositional semantics." For it is not that the represented condition is itself a normal condition for proper operation of the representation consumer. A certain correspondence between the representation and the world is what is normal. Coordinately, there is no such thing as a representation consumer that can understand only one representation. There are always other representations, composed other ways, saying other things, which it could have understood as well, in accordance with the same principles of operation. A couple of very elementary examples should make this clear.6

tio

tin

otl

loc

da

of

arı

fu

nc

to

Tł

wi

re

fu

СC

ta

ea

Si

tr

aı

h٤

be

bŧ

di

a٥

fa

C

 \mathbf{p}

cı

li

tł b

fı

o

First, consider beavers, who splash the water smartly with their tails to signal danger. This instinctive behavior has the function of causing other beavers to take cover. The splash means danger, because only when it corresponds to danger does the instinctive response to the splash on the part of the interpreter beavers, the consumers, serve a purpose. If there is no danger present, the interpreter beavers interrupt their activities uselessly. Hence, that the splash corresponds to danger is a normal condition for proper functioning of the interpreter beavers' instinctive reaction to the splash. (It does not follow, of course, that it is a usual condition. Beavers being skittish, most beaver splashes possibly occur in response to things not in fact endangering the beaver.) In the beaver-splash semantic system, the time and place of the splash varies with, "corresponds to," the time and place of danger. The representa-

^{6.} These examples are of representations that are not "inner" but out in the open. As in the case of inner representations, however, they are produced and consumed by mechanisms designed to cooperate with one another; each such representation stands intermediate between two parts of a single biological system.

tion is articulate: properly speaking, it is not a splash but a splash-at-a-time-and-a-place. Other representations in the same system, splashes at other times and places, indicate other danger locations.

Second, consider honey bees, which perform "dances" to indicate the location of sources of nectar they have discovered. The tempo of the dance and the angle of its long axis vary with the distance and direction of the nectar. The interpreter mechanisms in the watching bees—these are the representation consumers—will not perform their full proper functions of aiding the process of nectar collection in accordance with a normal explanation unless the location of nectar corresponds correctly to the dance. So the dances are representations of the location of nectar. The full representation here is a dance-at-a-time-in-a-place-at-a-tempowith-an-orientation.

Notice that, on this account, it is not necessary to assume that most representations are true. Many biological devices perform their proper functions not on the average but just often enough. The protective coloring of the juveniles of many animal species, for example, is an adaptation passed on because occasionally it prevents a juvenile from being eaten, though most of the juveniles of these species get eaten anyway. Similarly, it is conceivable that the devices that fix human beliefs fix true ones not on the average, but just often enough. If the true beliefs are functional and the false beliefs are, for the most part, no worse than having an empty mind, then even very fallible belief-fixing devices might be better than no belief-fixing devices at all. These devices might even be, in a sense, "designed to deliver some falsehoods." Perhaps, given the difficulty of designing highly accurate belief-fixing mechanisms, it is actually advantageous to fix too many beliefs, letting some of these be false, rather than fix too few beliefs. Coordinately, perhaps our beliefconsuming mechanisms are carefully designed to tolerate a large proportion of false beliefs. It would not follow, of course, that the belief consumers are designed to use false beliefs, certainly not that false beliefs can serve all of the functions that true ones can. Indeed, if none of the mechanisms that used beliefs ever cared at all how or whether these beliefs corresponded to anything in the world, beliefs would surely be functioning not as representations but in some other capacity.

Shifting our focus from producing devices to consuming devices in our search for naturalized semantic content is important. But the shift from the function of consumers to normal conditions for proper operation is equally important. Matthen, for example, characterizes what he calls a "quasi-perceptual state" as, roughly, one whose job is to cause the system to do what it must do to perform its function, given that it is in certain circumstances, which are what it represents. Matthen is thus looking pretty squarely at the representation consumers, but at what it is the representation's job to get these consumers to do, rather than at normal conditions for their proper operation. As a result, Matthen now retreats. The description he has given of quasi-perceptual states, he says, cannot cover "real perception such as that which we humans experience. Quite simply, there is no such thing as the proper response, or even a range of functionally appropriate responses, to what perception tells us" (1988, 20).7 On the contrary, representational content rests not on univocity of consumer function but on sameness of normal conditions for those functions. The same percept of the world may be used to guide any of very many and diverse activities, practical or theoretical. What stays the same is that the percept must correspond to environmental configurations in accordance with the same correspondence rules for each of these activities. For example, if the position of the chair in the room does not correspond so to my visual representation of its position, that will hinder me equally in my attempts to avoid the chair when passing through the room, to move the chair, to sit in it, to remove the cat from it, to make judgments about it, etc. Similarly, my belief that New York is large may be turned to any of diverse purposes, but those which require it to be a representation require also that New York indeed be large if these purposes are to succeed in accordance with a normal explanation for functioning of my cognitive systems.

3

We have just cleanly bypassed the whole genre of causal-informational accounts of mental content. To illustrate this, we consider an example of Dretske's. Dretske tells of a certain species of Northern Hemisphere bacteria which orient themselve away from toxic oxygen-rich surface water by attending to their magnetosomes, tiny inner magnets,

7. Dretske (1986, 28) and David Papineau (1987, 67ff.) have similar concerns.

which pull toward the magnetic north pole, hence pull down (1986). (Southern Hemisphere bacteria have their magnetosomes reversed.) The function of the magnetosome thus appears to be to effect that the bacterium moves into oxygen-free water. Correlatively, intuition tells us that what the pull of the magnetosome represents is the whereabouts of oxygen-free water. The direction of oxygen-free water is not, however, a factor in causing the direction of pull of the magnetosome. And the most reliable natural information that the magnetosome carries is surely not about oxygen-free water but about distal and proximal causes of the pull, about the direction of geomagnetic or better, just plain magnetic, north. One can, after all, easily deflect the magnetosome away from the direction of lesser oxygen merely by holding a bar magnet overhead. Moreover, it is surely a function of the magnetosome to respond to that magnetic field, that is part of its normal mechanism of operation, whereas responding to oxygen density is not. None of this makes any sense on a causal or informational approach.

But on the biosemantic theory, it does make sense. What the magnetosome represents is only what its consumers require that it correspond to in order to perform their tasks. Ignore, then, how the representation (a pull-in-a-direction-at-a-time) is normally produced. Concentrate instead on how the systems that react to the representation work, on what these systems need in order to do their job. What they need is only that the pull be in the direction of oxygen-free water at the time. For example, they care not at all how it came about that the pull is in that direction; the magnetosome that points toward oxygen-free water quite by accident and not in accordance with any normal explanation will do just as well as one that points that way for the normal reasons. (As Socrates concedes in the Meno, true opinion is just as good as knowledge, so long as it stays put.) What the magnetosome represents, then, is univocal; it represents only the direction of oxygen-free water. For that is the only thing that corresponds (by a compositional rule) to it, the absence of which would matter, the absence of which would disrupt the function of those mechanisms that rely on the magnetosome for guidance.

It is worth noting that what is represented by the magnetosome is not proximal but distal; no proximal stimulus is represented at all. Nor, of course, does the bacterium perform an inference from the existence of

and the second contract of the second contrac

the proximal stimulus (the magnetic field) to the exister ce of the represented. These are good results for a theory of content to have, for otherwise one needs to introduce a derivative theory of content for mental representations that do not refer, say, to sensory stimulations, and also a foundationalist account of belief fixation. Note also that, on the present view, representations manufactured in identical ways by different species of animal might have different contents. Thus, a certain kind of small, swift image on the toad's retina, manufactured by his eye lens, represents a bug, for that is what it must correspond to if the reflex it (invariably) triggers is to perform its proper functions normally, while exactly the same kind of small swift image on the retina of a male hoverfly, manufactured, let us suppose, by a nearly identical lens, represents a passing female hoverfly, for that is what it must correspond to if the female-chasing reflex it (invariably) triggers is to perform its proper functions normally. Turning the coin over, representations with the same content may be normally manufactured in a diversity of ways, even in the same species. How many different ways do you have, for example, of telling a lemon or your spouse? Nor is it necessary that any of the ways one has of manufacturing a given representation be especially reliable ways in order for the representation to have determinate content. These various results cut the biosemantic approach off from all varieties of verificationism and foundationalism with a clean, sharp knife.

4

But perhaps it will be thought that belief fixation and consumption are not biologically proper activities, hence that there are no normal explanations, in my defined sense, for proper performances of human beliefs. Unlike bee dances, which are all variations on the same simple theme, beliefs in dinosaurs, in quarks, and in the instability of the dollar are recent, novel, and innumerably diverse, as are their possible uses. How could there be anything biologically normal cr abnormal about the details of the consumption of such beliefs?

But what an organism does in accordance with evolutionary design can be very novel and surprising, for the more complex of nature's creatures are designed to learn. Unlike evolutionary adaptation, learning is not accomplished by random generate-and-test procedures. Even when learning involves trial and error (probably the exception rather than the rule), there are principles in accordance with which responses are selected by the system to try, and there are specific principles of generalization and discrimination, etc., which have been built into the system by natural selection. How these principles normally work, that is, how they work, given normal (i.e., historically optimal) environments, to produce changes in the learner's nervous system which will further ends of the system has, of course, an explanation—the normal explanation for proper performance of the learning mechanism and of the states of the nervous system it produces.

To use a worn-out comparison, there is an infinity of functions which a modern computer mainframe is capable of performing, depending upon its input and on the program it is running. Each of these things it can do, so long as it is not damaged or broken, "in accordance with design," and to each of these capacities there corresponds an explanation of how it would be activated or fulfilled normally. The human's mainframe takes, roughly, stimulations of the afferent nerves as input both to program and to run it.8 It responds, in part, by developing concepts, by acquiring beliefs and desires in accordance with these concepts, by engaging in practical inference leading ultimately to action. Each of these activities may, of course, involve circumscribed sorts of trial and error learning. When conditions are optimal, all this aids survival and proliferation in accordance with a historically normal explanation, one of high generality, of course. When conditions are not optimal, it may yield, among other things, empty or confused concepts, biologically useless desires, and false beliefs. But even when the desires are biologically useless (though probably not when the concepts expressed in them are empty or confused), there are still biologically normal ways for them to get fulfilled, the most obvious of which require reliance on true beliefs.9

^{8.} This is a broad metaphor. I am not advocating computationalism.

^{9.} A word of caution. The normal conditions for a desire's fulfillment are not necessarily fulfillable conditions. In general, normal conditions for fulfillment of a function are not quite the same as conditions which, when you add them and stir, always effect proper function, because they may well be impossible conditions. For example, Fodor (1989 and "NETC") has questioned me about the normal conditions under which his desire that it should rain tomorrow will perform its proper function of getting it to rain. Now the biologically normal way

Yet how do we know that our contemporary ways of forming concepts, desires, and beliefs do occur in accordance with evolutionary design? Fodor, for example, is ready with the labels "pop Darwinism" and "naive adaptationism" to abuse anyone who supposes that our cognitive systems were actually selected for their belief- and desice-using capacities (1987, "NETC"). Clearly, to believe that every structure must have a function would be naive. Nor is it wise uncritically to adopt hypotheses about the functions of structures when these functions are obscure. It does not follow that we should balk at the sort of adaptationist who, having found a highly complex structure that quite evidently is currently and effectively performing a highly complex and obviously indispensable function, then concludes, ceteris paribus, that this function has been the most recent historical task stabilizing the structure. To suspect that the brain has not been preserved for thinking with or that the eye has not been preserved for seeing with—to suspect this, moreover, in the absence of any alternative hypotheses about causes of the stability of these structures—would be totally irresponsible. Consider: Nearly every human behavior is bound up with intentional act on. Are we really to suppose that the degree to which our behaviors help to fulfill intentions, and the degree to which intentions result from logically related desires plus beliefs, is a sheer coincidence, that these patterns are irrelevant to survival and proliferation or, though relevant, have had no stabilizing effect on the gene pool? But the only alternative to biological design, in my sense of "design," is sheer coincidence, freak accidentunless there is a ghost running the machine!10

for such a desire to be fulfilled is exactly the same as for any other desire: one has or acquires true beliefs about how to effect the fulfillment of the desire and acts on them. Biologically normal conditions for fulfillment of the desire for rain thus include the condition that one has true beliefs about how to make it rain. Clearly this is an example in which the biological norm fails to accord with the statistical norm: most desires about the weather are fulfilled, if at all, by biological accident. It may even be that the laws of nature, coupled with my situation, prohibit my having any true beliefs about how to make it rain; the needed general condition cannot be realized in the particular case. Similarly, normal conditions for proper function of beliefs in impossible things are, of course, impossible conditions: these beliefs are such that they cannot correspond, in accordance with the rules of Mentalese, to conditions in the world. For more on this theme, see chapter 8.

10. For more details, see chapter 2 herein.

Indeed, it is reasonable to suppose that the brain structures we have recently been using in developing space technology and elementaryparticle physics have been operating in accordance with the very same general principles as when prehistoric man used them for more primitive ventures. They are no more performing new and different functions or operating in accordance with new and different principles nowadays than are the eyes when what they see is television screens and space shuttles. Compare: the wheel was invented for the purpose of rolling ox carts, and did not come into its own (pulleys, gears, etc.) for several thousand years thereafter, during the industrial revolution. Similarly, it is reasonable that the cognitive structures with which man is endowed were originally nature's solution to some very simple demands made by man's evolutionary niche. But the solution nature stumbled on was elegant, supremely general, and powerful; indeed, I believe it was a solution that cut to the very bone of the ontological structure of the world. That solution involved the introduction of representations, inner and/or outer, having a subject/predicate structure and subject to a negation transformation. (Why I believe that that particular development was so radical and so powerful has been explained in depth in LTOBC, chaps. 14-19. But see also subsection "Negation and propositional content" below.)

5

One last worry about my sort of position is voiced by Daniel Dennett (1978a) and discussed at length by Fodor (1986b). Is it really plausible that bacteria and paramecia, or even birds and bees, have inner representations in the same sense that we do? Am I really prepared to say that these creatures too have mental states, that they think? I am not prepared to say that. On the contrary, the representations that they have must differ from human beliefs in at least six very fundamental ways.¹¹

11. Accordingly, in *LTOBC*, I did not call these primitive forms "representations" but "intentional signals" and, for items like bee dances, "intentional icons," reserving the term "representation" for those icons, whose representational values must be identified if their consumers are to function properly. See the subsection "Acts of identifying" below.

Self-representing elements

The representations that the magnetosome produces have three significant variables, each of which refers to itself. The time of the pull refers to the time of the oxygen-free water, the locale of the pull refers to the locale of the oxygen-free water, and the direction of pull refers to the direction of oxygen-free water. The beaver's splash has two self-referring variables: a splash at a certain time and place indicates that there is danger at that same time and place. (There is nothing necessary about this. It might have meant that there would be danger at the nearest beaver dam in five minutes.) Compare the standard color coding on the outsides of colored markers: each color stands for itself. True, it may be that sophisticated indexical representations such as percepts and indexical beliefs also have their time or place or both as significant self-representing elements, but they also have other significant variables that are not self-representing. The magnetosome does not.

Storing representations

Any representation the time or place of which is a significant variable obviously cannot be stored away, carried about with the organism for use on future occasions. Most beliefs are representations that can be stored away. Clearly, this is an important difference.

Indicative and imperative representations-

The theory I have sketched here of the content of inner representations applies only to indicative representations, representations which are supposed to be determined by the facts, which tell what is the case. It does not apply to imperative representations, representations which are supposed to determine the facts, which tell the interpreter what to do. Neither do causal-informational theories of content apply to the contents of imperative representations. True, some philosophers seem to have assumed that having defined the content of various mental symbols by reference to what causes them to enter the "belief box," then when one finds these same symbols in, say, the "desire box" or the "intention box," one already knows what they mear. But how do we know that the desire box or the intention box use the same representational system as the belief box? To answer that question we would have to

know what constitutes a desire box's or an intention box's using one representational system rather than another which, turned around, is the very question at issue. In LTOBC and chapters 3 and 8, I develop a parallel theory of the content of imperative representations. Very roughly, one of the proper functions of the consumer system for an imperative representation is to help *produce* a correspondence between the representation and the world. (Of course, this proper function often is not performed.) I also argue that desires and intentions are imperative representations.

Consider, then, the beaver's splash. It tells that there is danger here now. Or why not say instead that it tells other nearby beavers what to do now, namely, to seek cover? Consider the magnetosome. It tells which is the direction of oxygen-free water. Or why not say instead that it tells the bacterium which way to go? Simple animal signals are invariably both indicative and imperative. Even the dance of the honey bee, which is certainly no simple signal, is both indicative and imperative. It tells the worker bees where the nectar is; equally, it tells them where to go. The step from these primitive representations to human beliefs is an enormous one, for it involves the separation of indicative from imperative functions of the representational system. Representations that are undifferentiated between indicative and imperative connect states of affairs directly to actions, to specific things to be done in the face of those states of affairs. Human beliefs are not tied directly to actions. Unless combined with appropriate desires, human beliefs are impotent. And human desires are equally impotent unless combined with suitable beliefs.12

Inference

As indicative and imperative functions are separated in the central inner representational systems of humans, they need to be reintegrated. Thus, humans engage in practical inference, combining beliefs and desires in novel ways to yield first intentions and then action. Humans also combine beliefs with beliefs to yield new beliefs. Surely nothing remotely like this takes place inside the bacterium.

12. Possibly, human intentions are simultaneously in both indicative and imperative moods, however, functioning to represent settled facts about one's future and also to direct one's action.

Acts of identifying

Mediate inferences always turn on something like a middle term, which must have the same representational value in both premises for the inference to go through. Indeed, the representation consumers in us perform many functions that require them to use two or more overlapping representations together and in such a manner that, unless the representeds corresponding to these indeed have a common element, these functions will not be properly performed. Put informally, the consumer device takes these represented elements to be the same, thus identifying their representational values. Suppose, for example, that you intend to speak to Henry about something. In order to carry out this intention you must, when the time comes, be able to recognize Henry in perception as the person to whom you intend to speak. You must identify Henry as represented in perception with Henry as represented in your intention. Activities that involve the coordinated use of representations from different sensory modalities, as in the case of eye-hand coordination, visual-tactile coordination, also require that certain objects, contours, places, or directions, etc., be identified as the same through the two modalities. Now, the foundation upon which modern representational theories of thought are built depends upon a denial that what is thought of is ever placed before a naked minc. Clearly, we can never know what an inner representation represents by a direct comparison of representation to represented. Rather, acts of identifying are our ways of "knowing what our representations represent." The bacterium is quite incapable of knowing, in this sense, what its representations are about. This might be a reason to say that it does not understand its own representations, not really.

Negation and propositional content

The representational system to which the magnetosome pull belongs does not contain negation. Indeed, it does not even contain contrary representations, for the magnetosome cannot pull in two directions at once. Similarly, if two beavers splash at different times or places, or if two bees dance different dances at the same time, it may well be that there is indeed beaver danger two times or two places and that there is indeed nectar in two different locations.¹³ Without contrariety, no con-

13. On the other hand, the bees cannot go to two places at once.

101

flict, of course, and more specifically, no contradiction. If the law of noncontradiction plays as significant a role in the development of human concepts and knowledge as has traditionally been supposed, this is a large difference between us and the bacterium indeed.¹⁴ In LTOBC, I argued that negation, hence explicit contradiction, is dependent upon subject-predicate, that is, propositional, structure and vice versa. Thus, representations that are simpler also do not have propositional content.

In sum, these six differences between our representations and those of the bacterium, or Fodor's paramecia, ought to be enough amply to secure our superiority, to make us feel comfortably more endowed with mind.

14. In LTOBC, I defend the position that the law of noncontradiction plays a crucial role in allowing us to develop new methods of mapping the world with representations.