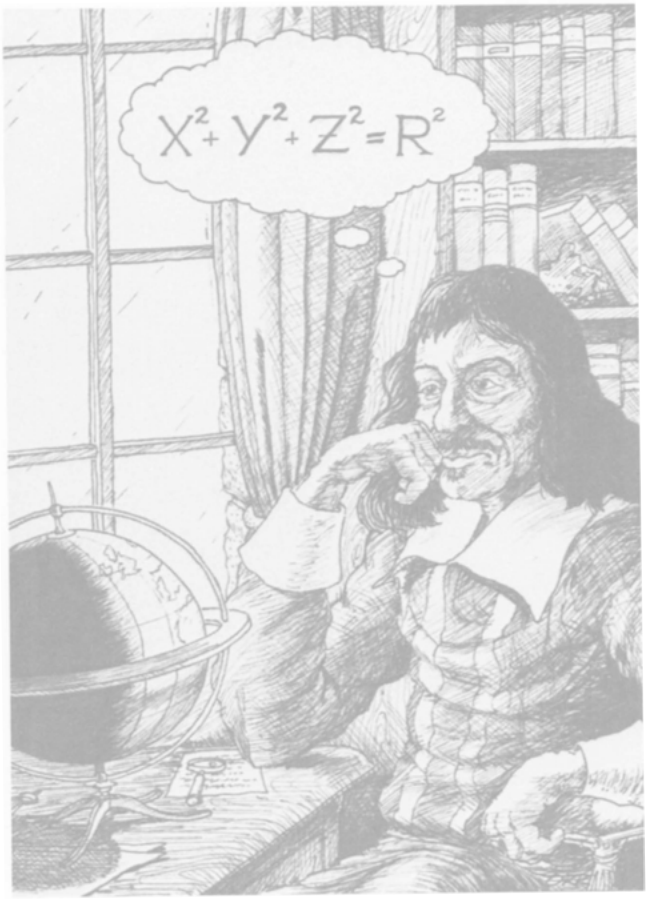
MEANING AND MENTAL REPRESENTATION

Robert Cummins

Meaning and Mental Representation



Meaning and Mental Representation

Robert Cummins

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

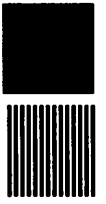
Third printing, 1995

First MIT Paperback edition, 1991

This book was set in Palatino and was printed and bound in the United States of America.

Library of Congress Cataloging in Publication Data

Cummins, Robert.
Meaning and mental representation/ Robert Cummins.
p. cm.
"A Bradford book."
Includes index.
ISBN 0-262-03139-6 hardcover, 0-262-53096-1 paperback
1. Representation (Philosophy) 2. Mind and body 3. Knowledge, Theory of. 4. Meaning. 5. Science — Philosophy. I. Title.
B105.R4C86 1988 88-856
153— dc19 CIP



Philosophy of science is philosophy enough.

W. V. O. Quine

Contents

Acknowledgments

Many students and colleagues have helped make this a better book than it would have been had I been left to my own devices. I would like especially to thank Georg Schwarz, Paul Smolensky, Georges Rey, David Kirsh, and Ned Block. My greatest debt is to Denise Dellarosa Cummins, who, through constant example and constructive criticism, prevented me from drifting into philosophical disputes that have lost touch with the practice of empirical research and explanation in cognitive science. The art work is by Brian Orr.

Meaning and Mental Representation

Chapter 1

Identifying the Problem and Other Preliminaries

Two Problems about Representation

We should be careful to distinguish two problems about mental representation. The first—the Problem of Representations (plural)—is a theoretical problem in empirical science. Although we know that states of and processes in the nervous system play the role of representations in biological systems, it is an open question just which states and processes are involved in which activities, and how. Moreover, it is an open question how these states or processes should be characterized. For example, orthodox computationalism holds that mental representations are realized as symbolic data structures, but there is considerable controversy among orthodox computationalists as to what kinds of data structures are involved in various processes. Connectionists (see, e.g., Rumelhart et al. 1986), on the other hand, hold that mental representations are realized as activation levels of ensembles of simple processors, and/or as the strengths of the connections among such processors. The problem to which these approaches offer competing responses is that of discovering a way of characterizing representations that will allow us to understand both their physical instantiations and their systematic roles in mental processes.

The second problem—the Problem of *Representation* (singular)—is, at least as I understand it, a paradigmatic problem in the philosophy of science. To a large extent, empirical theories of cognition can and do take the notion of mental content as an explanatory primitive. But this is a kind of explanatory loan

(Dennett 1978): If it turns out that the notion of mental representation cannot be given a satisfactory explication— if, in particular, no account of the nature of the (mental) representation relation can be given that is consistent with the empirical theory that assumes it—then, at least in this respect, that empirical theory must be regarded as ill founded, and hence as a less than adequate response to the drive for the kind of thorough intellectual understanding that motivates scientific theory in the first place.

We can get a better idea of what these two problems are, and how they are related, by surveying in very general terms the various answers that have been tendered to each of them.

The Problem of Representations

It is surprising that only four answers have been suggested concerning the sorts of things that can be mental representations. I am not certain that this list of ours is exhaustive, but every proposal I know of fits pretty clearly into one of these four. It doesn't really matter much; my topic is the nature of representation, not what sorts of things do the representational work of the mind. I survey the alternatives here mainly to help to put the main problem in some context.

Mind-stuff inFORMed An important scholastic theory holds that in perception the immaterial mind becomes inFORMed by the same FORMS that inFORM the thing perceived. The background metaphysics assumes that knowable or perceivable things are a combination of matter and FORM: the *stuff* and its properties. There are two basically different kinds of *stuff:* mental stuff and physical stuff. When physical stuff is inFORMed by redness and sphericity, the result is a physical red ball. When mental stuff is inFORMed by redness and sphericity, the result is an idea of a red ball—or, perhaps better, the result is a red ball *as mental object* (i.e., as idea) rather than a red ball *as material object.* According to this theory, when you perceive a red ball, the very same FORMS that make the physical object of your perception red and spherical make your idea red and spherical. But of course a red ball *in*

Figure 1.1
Aristotle mentally representing Graycat with a ball.

idea is a very different thing than red ball *in matter*. A red ball *in idea* doesn't take up physical space, though it does take up *mental* space.

The basic idea behind this theory is that to know something is, in a pretty straightforward sense, to *be* it. You know the red ball when you see it because *you* have what *it* has: redness and sphericity. Your mind literally *is* just what the physical stuff is, because to *be* red and spherical is just to be inFORMed by redness and sphericity. This doctrine seems to make the notion of mental representation perfectly transparent: The idea represents the red ball, and it represents it as red and as spherical because the idea *is* red and spherical and the redness and sphericity come from the physical ball. To represent the world is to have a model of it in (on?) your mind—a model made of different stuff, as models usually are, but a model just the same. It we draw a picture, we, as theorists, can just *see* what represents what—e.g., the thing on the left part of the thought represents the cat, and the thing on the right part of the thought represents the ball. According to this theory, representation is evidently founded on similarity (shared properties)—a similarity the theorist can just *see*. Of course, the thinker can't just see it, as Berkeley and Hume eventually pointed out, but that is an epistemological problem at most. The fact that we can't see the alleged similarity between our own mental representations and what they represent (or see the representations at all, for that matter) doesn't show that it isn't similarity that underwrites representation; it only emphasizes the trivial fact that we can't hope to infer the way the world is from prior knowledge of the fact that we have it represented correctly.

Images The favorite theory of Berkeley and Hume was that mental representations are images. Except for dropping the Aristotelian jargon, however, this is just the same theory over again; the "picture" in both cases is just the same. Images were frequently said to be red and spherical, though with some uneasiness. The scholastic metaphysics was gone, but the basic idea was the same: Images represent things in virtue of resembling them—i.e., in virtue of sharing properties with them (though, of course, a sphere in the mind—i.e., as it exists as an image—takes up no

Figure 1.2
Berkeley's mental representations look just like Aristotle's.

physical space, only mental space; it occupies a portion of the visual field, for example).

Symbols Haugeland (1985) credits Hobbes with being the first to have an inkling that mental representations might be language-like symbols. This is now the orthodox position, insofar as there is such a thing. The main thing to realize at this stage is just that if mental representations are symbols, then mental representation cannot be founded on similarity; symbols don't resemble the things they represent. The great advantage of symbols as representations is that they can be the inputs and outputs of *computations*. Putting these two things together gives us a quick account of the possibility of thought about abstractions. When you calculate, you think about numbers by manipulating symbols. The symbols don't resemble the numbers, of course (what would resemble a number?), but they are readily manipulated.

Connectionists also hold that mental representations are symbols, but they deny that these symbols are data structures (i.e., objects of computation). In orthodox computational theory the objects of computation are identical with the objects of semantic interpretation, but in connectionist models (at least in those using truly distributed representation) this is not the case.[1] Connectionists also typically deny that mental symbols are language-like. This is not surprising; given that the symbols are not the objects of computation, there would be no obvious way to exploit a language-like syntactic structure in the symbols anyway.

(Actual) neurophysiological states The crucial claim here is that mental representations cannot be identified at any level more abstract than actual neurophysiology. Mental representation, on this view, is a biological phenomenon essentially. Mental representations cannot be realized in, say, a digital computer, no matter how "brain-like" its architecture happens to be at some nonbiological level of description.

Like symbols, neurophysiological states cannot represent things in virtue of resembling them. Advocates of symbols or neurophysiological states must ground representation in something other than similarity.

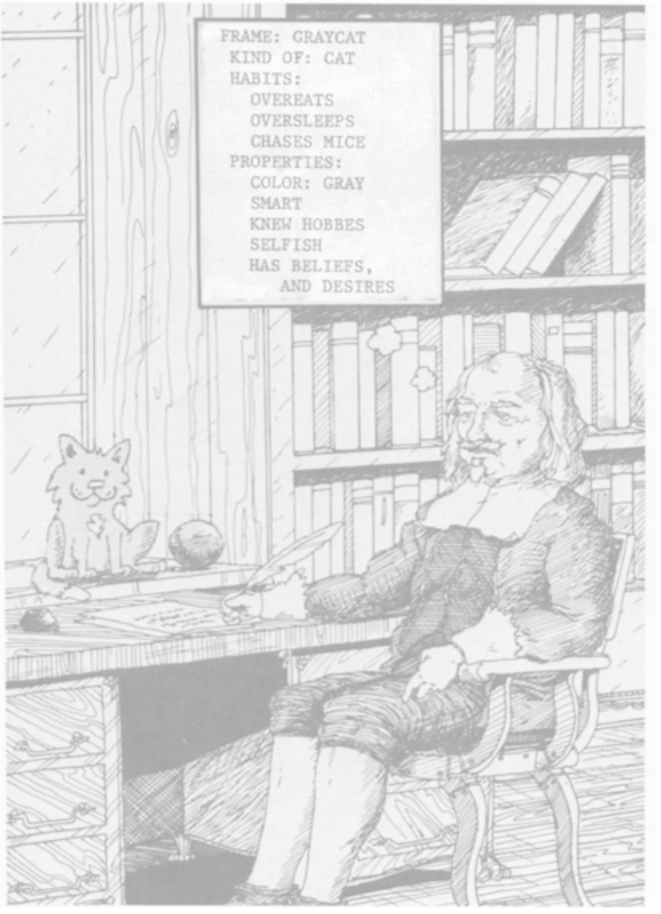


Figure 1.3
Hobbes representing Graycat.

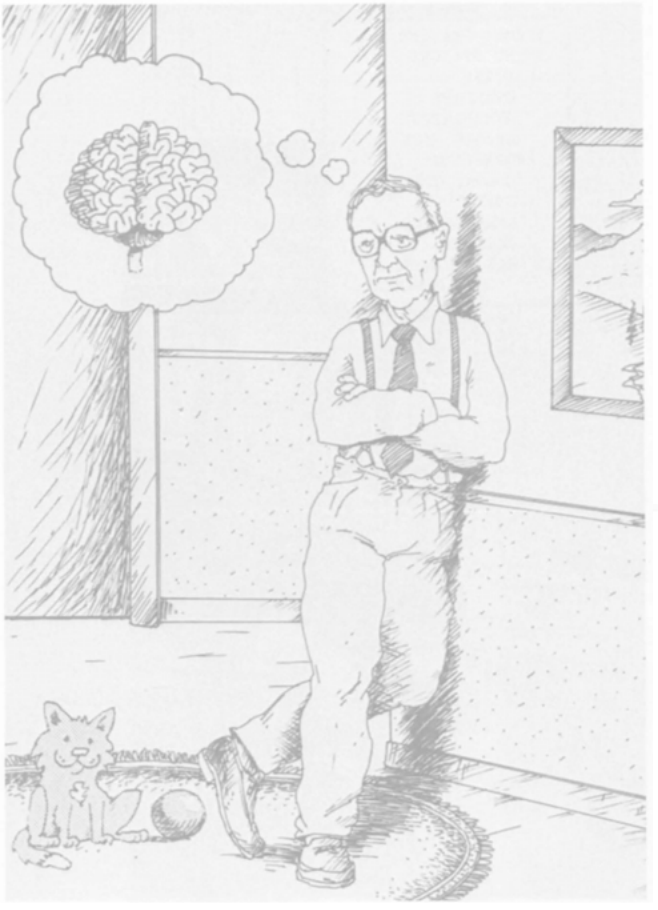


Figure 1.4
Hebb mentally representing Graycat.

The Problem of Representation

More surprising than the dearth of candidates to play the role of mental representations is the dearth of suggestions concerning the nature of representation itself. There are, I think, only four: similarity, covariance, adaptational role, and functional role. Each of these will be the subject of a chapter. For now, I will supply only brief intuitive sketches.

Similarity The thought that representation is grounded in similarity is what drives the idea that mental representations are in-FORMed mind stuff, or images. The crucial intuition, I think, is this: If you are going to think about things in the world, you need something to go proxy for those things in thought. You cannot, of course, literally turn over cats or the body politic in your mind; all you can turn over is ideas. But this, it seems, will be no help unless ideas are *like* the cats or the body politic: How could having an idea of a cat help you know about cats unless the idea is like the cat? I could say, "OK, this salt shaker represents the pitcher, and the pepper shaker represents the batter." But wouldn't *pictures* be much better—especially moving pictures, such as those in Rod Carew's batting instruction video?

Covariance The idea that representation is grounded in covariance or causation is most naturally motivated by reflecting on vision research.[2] How do we decide, for example, that a certain neural structure in the visual cortex of a frog is a motion detector? Roughly, we notice that a certain characteristic activity in the structure covaries with the presence of moving objects in the frog's field of vision. Given this fact, it seems natural to suppose that what *makes* that structure a motion detector is just the fact that it fires when there is motion in the frog's field of vision. What else could it be? So the fact that the firing of the structure in question represents the occurrence of motion in the frog's visual field is just constituted by the covariance between the firings and the motions represented. If you are attracted to covariance theories, you aren't going to think much of the idea that representations are images, because the similarities images promise to deliver are going to be irrelevant.

Adaptational role The idea that representation is grounded in adaptational role is most easily understood as a reaction to certain problems facing covariance theories. The orientation of a bee dance represents the location of flowers to spectator bees, but it doesn't covary with the location of flowers any better than it covaries with lots of things it doesn't represent, e.g., the absence of an insecticide cloud in the indicated direction. Millikan (1984) points out that we take "flowers over there" to be the content of the dance, even if flowers are not often "over there" (and hence there is no substantial covariance), because the cases in which spectators have found flowers (hence food) "over there" account for the continued replication of the dance and the characteristic response it evokes in spectators.

Functional or computational role This is just functionalism applied to mental representations. Functionalism says that a mental state is what it is in virtue of its functional role. It is functional roles that individuate mental states. But mental representations are, by definition, individuated by their contents. Hence, content must depend on functional role.[3]

Meanings and Meaningfulness

When we ask what it is in virtue of which something (a mental state, a stop sign, a linguistic utterance) has a meaning or has semantic content, there are two quite different things we may have in mind. We may be asking what it is in virtue of which things of the sort in question have any meaning at all, or we may be asking what it is in virtue of which some particular thing or type of thing has some particular meaning. Although it is rather obvious that a theory that answers the first sort of question (a theory of meaningfulness) needn't provide answers to question of the second sort, it is not so obvious that a theory that provides answers to questions of the second sort (a theory of meaning) must also be a theory of meaningfulness. All the theories I will examine in this book are intended primarily as theories of meaning, not as theories of meaningfulness; but each of them entails, in an obvious way, a theory of meaningfulness. I shall try to make

this explicit and, when appropriate, to be clear about whether the theory is being expressed and evaluated as a theory of meaning or as a theory of meaningfulness.

Theories of meaning, in the sense just staked out, should be sharply distinguished from theories that, as it were, distribute meanings (or some other semantic property) over the things that have them. For example, it is perfectly possible to articulate a theory that specifies a truth condition for every sentence in a language but that is entirely neutral concerning what it is in virtue of which a sentence has any truth condition at all, or in virtue of which it has the particular truth condition it happens to have. Tarski's theory of truth is, notoriously, just such a theory—truth is defined in terms of satisfaction, and satisfaction is defined recursively. The theory is completely silent about what satisfaction is. If we ask "In virtue of what is 'X_1 is a cat' satisfied by every sequence beginning with a cat?" the theory gives no answer (see Field 1971 and Cummins 1975a). Linguists and psychologists want to know which things have which meanings, and why. Philosophers want to know what it is to have a meaning. With any luck, good philosophy might help with the "why" part of the question asked by linguists and psychologists. By my lights, that is really the only thing that could make it good philosophy.[4]

"Content"

When we suppose a system to harbor cognitive representations, we are supposing that the system harbors states, or perhaps even objects, that are semantically individuated. Thus, the central question about mental representation is this: What is it for a mental state to have a semantic property? Equivalently, what makes a state (or an object) in a cognitive system a *representation?*

When we ask what it is for a cognitive state to have a semantic property, there are a number of different things on which we might choose to focus. What is it for a cognitive state to have a truth condition? What is it for a cognitive state to be about something, or to refer to something, or to be true of something?[5] What is it for a cognitive state to be an intentional state (i.e., to

have intentional properties)? The (very) recent tendency in philosophy has been to see all these questions as depending on two prior questions: What is it for a cognitive state to have a content? What is it for such a state to have some specified content, e.g., the content *that Brutus had flat feet* or the content *square?* This, I think, is a useful way to proceed—not because the notion of content is especially clear or simple, but because "content" can function in philosophical investigation as a kind of generic term for whatever it is that underwrites semantic and intentional properties generally. There is little to be gained, and there is a non-negligible risk of bias, if we begin by focusing in a fussy way on semantic or intentional concepts borrowed from theoretical or common-sense discourse about language and the attitudes—concepts that may not apply in any straightforward way to the problem of characterizing the representations assumed by contemporary cognitive theory. In what follows, when I write of the semantics of cognitive systems, or of representations, I mean to address these still poorly defined questions of "content." Since I shall be examining various "theories of content," there is no point in trying to say in advance what "content" means; let the theories speak for themselves. Meanwhile, our intuitive grasp of the thing will have to do.

Methodology

It is commonplace for philosophers to address the question of mental representation in abstraction from any particular scientific theory or theoretical framework. I regard this as a mistake. Mental representation is a theoretical assumption, not a commonplace of ordinary discourse. To suppose that "common-sense psychology" ("folk psychology"), orthodox computationalism, connectionism, neuroscience, and so on all make use of the same notion of representation seems naive. Moreover, to understand the notion of mental representation that grounds some particular theoretical framework, one must understand the explanatory role that framework assigns to mental representation. It is precisely because mental representation has different explanatory roles in "folk psychology," orthodox computational-

ism, connectionism, and neuroscience that it is naive to suppose that each makes use of the same notion of mental representation. We must not, then, ask simply (and naively) "What is the nature of mental representation?"; this is a hopelessly unconstrained question. Instead, we must pick a theoretical framework and ask what explanatory role mental representation plays in that framework and what the representation relation must be if that explanatory role is to be well grounded. Our question should be "What must we suppose about the nature of mental representation if orthodox computational theories (or connectionist theories, or whatever) of cognition are to turn out to be true and explanatory?" As I understand this question, it is a question in the philosophy of science exactly analogous to the following question in the philosophy of physics: What must we suppose the nature of space to be (substance? property? relation?) if General Relativity is to turn out to be true and explanatory?

The bulk of this book is an attempt to evaluate existing accounts of the nature of mental representation in the context of computational theories of cognition. By computational theories of cognition I mean *orthodox* computational theories—theories that assume that cognitive systems are automatic interpreted formal systems in the sense of Haugeland (1981, 1985), i.e., that cognition is disciplined symbol manipulation.[6] In the final chapter, I will consider briefly how things might look in a connectionist context.

Computational theories assume that mental representations are symbolic data structures as these are understood in computer science. This is the computationalist answer to the Problem of *Representations*. Although the instantiation of symbolic data structures in the brain is problematic, orthodox computationalism has demonstrated the physical instantiability of such structures and has made considerable progress toward demonstrating that at least some cognitive processes can be understood as symbol manipulation. But, like all theoretical frameworks in cognitive science, orthodox computationalism is silent about the nature of representation itself; it is entirely agnostic concerning what it is for a data structure to have semantic properties. Nevertheless, certain possibilities are ruled out by the empirical assumptions of the theory, as we will see.

I will need a short, convenient way to refer to what I have been calling orthodox computationalism; I'll call it the CTC, for the computational theory of cognition.

Representation and Intentionality

This preliminary issue of the explanatory role of mental representation in some particular theoretical framework would not be troubling if mental representation were a *commonplace* rather than a (variously) theoretically motivated *hypothesis*. Most philosophers aren't troubled; they think mental representation *is* a commonplace. They think this because they assume that the problem of mental representation is just the problem of intentionality—i.e., that representational content is intentional content. As I use the term, a system with intentionality is just a system with ordinary propositional attitudes (belief, desire, and so on). Thus construed, intentionality is a commonplace, and hence so is intentional content. So the assumption I want to scout is the assumption that the problem of mental representation is just the problem of what attaches beliefs and desires to their contents.[7]

Although it is evidently a mistake to identify intentionality with representation, there is a widely bruited philosophical theory, mainly due to Fodor, that forges a close connection between intentional contents and representational contents. I call this theory the *representational theory of intentionality* (RTI). The RTI holds that intentional states inherit their contents from representations that are their constituents. The familiar urtheory goes like this: To have a belief is to have a representation in one's belief box—a box distinguished from the desire box by its function, i.e., by which processes can put things in and take things out. (Belief-box contents are available as premises in inference; desire-box contents are available as goals, i.e., conditions whose satisfaction ends processing cycles.) My belief that U.S. policy in Central America is folly is *about* Central America because the relevant representation in my belief box represents Central America.[8] The RTI has some nice features. Most notably, it captures the two attributes of the propositional attitudes to which we allude when we call them by that name: that they have

propositional contents and that believing involves taking a different "attitude" toward a proposition than desiring. But in spite of its nice features, the RTI is no truism; it is a controversial and powerful empirical theory.

If you accept the RTI explicitly, you will, of course, want a theory of mental representation that attaches intentional contents—the contents of propositional attitudes—to representational states. You will also want a theory of mental representation like this if you are merely sloppy about the difference between mental representation and the attitudes. I think this particular bit of sloppiness is pretty common in a lot of recent philosophical discussion of mental representation, but it doesn't really matter; anyone who assumes, for whatever reason, that a theory of mental representation must give us intentional contents (e.g., objects of belief) is making a very large assumption, an assumption that isn't motivated by an examination of the role representation plays in any current empirical theories. After all, it isn't belief of any stripe that most theoretical appeals to mental representation are designed to capture. Just think of psycholinguistics, which got all this representation talk started. The data structures of your favorite parser are not even *prima facie* candidates for belief contents. This is nonaccidentally related to the fact that the CTC, as we will see in chapter 8, makes use of a notion of representation that is at home in computational systems generally, not just in cognitive systems and certainly not just in intentional systems. If we begin our investigation of mental representation by focusing on intentional states, we will miss what is most distinctive about representation as invoked by the CTC. We certainly do not want to assume, therefore, that the contents of beliefs as ordinarily attributed are the contents of any representations in a computational system. We need to keep open the possibility that, e.g., belief attribution, though a legitimate case of semantic characterization, is not a semantic characterization of any representation in the believer (Dennett 1978; Stalnaker 1984; Cummins 1987).

The fact that current philosophers who are interested in mental representation do not follow the methodological path that I recommended in the last section is explained to some extent by

the prevalence of the assumption (often bolstered by the RTI) that the problem of mental representation is to explain how intentional contents (the contents of belief, desire, etc.) get attached to mental states. This assumption puts very strong constraints on the theory of mental representation. In fact, the constraints are so strong—so hard to satisfy—that one is never tempted to look elsewhere for something to constrain the problem; the last thing one needs is another constraint. Thus, you will never be moved to ask after such things as the explanatory role of representation in, say, John Anderson's ACT* (1983). Conversely, once you abandon (or at least question) the idea that the theory of mental representation must yield contents for intentional states, you *need* a few constraints, and the explanatory structure of a theory that invokes the notion of representation is the natural place to look.

Inexplicit Content[9]

The attribution of intentional states (beliefs and desires) is not the only kind of semantic characterization of cognitive systems that must be distinguished from explanatory appeals to representational states. A computational system can also be semantically characterized in virtue of features of its structure. Here are some examples.

Content implicit in the state of control A word processor's search routine tries to match the character currently being read against the second character of the target only if the character read last matched the first character of the target. If it is now trying to match the second character, the current state of control carries the information that the first character matched the last character read; however, the system creates no data structure with this content. Nowhere is that information explicitly represented.

Content implicit in the domain I give you instructions for getting to my house from yours, all in such terms as "go left after three intersections" and "turn right at the first stop sign after the barn." Perhaps I even include things like "Make a left down the alley with the blue Chevy van parked in it," because I know you will

be coming after 5 o'clock and I know that the van is always parked there after that time. I rely on this in the same way I rely on the barn's staying put. Now, if you (or anything else) execute this program, you will get to my house. In the process, you never create a representation of the form "Cummins lives at location L"; yet, given the terrain, a system executing this program does "know where Cummins lives."

Content implicit in the form of representation Most of us don't know how to multiply (or even add) roman numerals. "XXII times LXIV" has the same *meaning* as "22 times 64," but the partial-products algorithm we all learned in school exploits information that is implicit in the second *form* but not present in the first—e.g., that shifting a column to the left amounts to multiplying by 10. This is the famous problem of knowledge representation in artificial intelligence: find a form that makes more efficient or psychologically realistic algorithms possible.

Content implicit in the medium of representation Are the two parts of figure 1.5 the same? If you had each one on a transparency, you could simply put one over the other and rotate them relative to each other to see if they would match. But this works only because of two properties of the *medium* (i.e., the transparencies): They are transparent, and they are rigid in the plane of the figures. When you rotate them, the information about the relative spatial relations of parts of a figure to other parts is implicit in the medium; its rigidity carries the information that these relations remain constant. A different medium might not carry this information, and you would then have to represent it explicitly.

I am sure these examples don't exhaust the cases in which content can be attributed to a computational system in the absence of any explicit representation having the content in question. I have listed them here only to emphasize the fact that represented content isn't all the content there is. There is also inexplicit content of various kinds, and if nothing like the RTI is true there is also intentional content.[10]

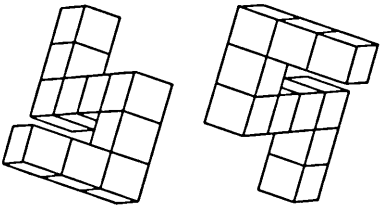


Figure 1.5
Are these the same figure?

Representation and the Language of Thought

Representation is often identified with what is really only one kind of representation: quasi-linguistic representation of the sort featured in Fodor's book *The Language of Thought* (1975). But it is at least possible that cognitive states might involve representations of some sort without involving quasi-linguistic formulas type-identified by their status in an internal code with a recursive syntax. In this book, when I mean language-like representations—sentences, or their constituents, written in a brain or in some other physical medium—I will make that explicit. In this connection, it is important to keep in mind that representations may well have propositional contents even though those representations are not language-like, for I take it that an essential feature of the Language-of-Thought Hypothesis—the hypothesis that mental representations are language-like—is that mental representations have a syntax comparable to that of a natural or an artificial language. But it is perfectly obvious that a symbol can have a propositional content—can have a proposition as its proper interpretation—even though it has no syntax and is not part of a language-like system of symbols. Paul Revere's lanterns are a simple case in point.

Cognition and the Mental

As is no doubt obvious by now, the use of the word "mental" in the title is misleading, for I will be talking about cognitive systems rather than minds. Some cognitive systems are not

minds (not, at least, as we know minds ostensively), and many aspects of mentality are not cognitive. Cognitive science is founded on the empirical assumption that cognition (hence the study of cognitive systems) is a natural and relatively autonomous domain of inquiry. I shall simply accept this assumption, but a few brief comments are in order.

When we run through mental phenomena as we know them from the human case, many seem inessential in that something could be a mind without exhibiting them. For example, it seems plausible to suppose that a creature could have a mind without having emotions, as is supposed to be the case with Star Trek's Mr. Spock. Descartes held that the essence of mind is thought, Locke that it is the capacity for thought. A system that could do nothing but think might be a rather colorless mind by human standards, but there seems to be something to the traditional idea that such a system would nevertheless *be* a mind. On the other hand, a system that could not think but could feel, have emotions, and so on does not seem to qualify as a mind. If this is right, then what cognitive science proposes is not, after all, very novel; it is just the idea that thinking (and/or the capacity for thought) is the essence of mind and can be studied independently of other mental phenomena.

It is important to be clear about what this hypothesis does and does not accomplish in the way of creating scientific elbow room. It *does* make it possible for the cognitive scientist to ignore (provisionally, at least) such mental phenomena as moods, emotions, sensations, and—most important—consciousness. The hypothesis that cognition is a relatively autonomous domain does not, however, entitle the cognitive scientist to ignore either human psychology or neuroscience. Human beings are the best and only uncontroversial example of cognitive systems we have to study. To try to study cognition without paying attention to how humans cognize would be like trying to study genetics without bothering about biochemistry; some progress is possible, but not a great deal.

Most objections to materialist theories of mind proceed by trying to establish either that a purely physical system could not

be a cognitive system or that a purely physical system could not be conscious. A materialist theory of *cognition* requires a response to the first sort of argument. But materialists, protected by the empirical hypothesis that cognition is separable from mentality generally, can afford to put off responding to the charge that a purely physical system could not be conscious. Perhaps consciousness isn't essential to mind in the way that cognition is.[11] This does not make the problem of consciousness go away, but it does make it, provisionally, someone else's problem.

Since my concern is with thought and not with mental processes generally, it would help to have a term that, unlike "mental representation," suggests only representations that play a role in thought or cognition. "Cognitive representation" isn't too bad; however, for stylistic reasons I will generally stick to the traditional "mental representations." Our questions will be "What is it for a mental whatnot to be a representation (i.e., to have a content)?" and "What is it for a mental representation, a whatnot with a content, to have some particular content rather than some other particular content?"

Bibliography

Anderson, J. 1983. *The Architecture of Cognition.* Hillsdale, N.J.: Erlbaum.

Ayer, A.J. 1951. *Language, Truth and Logic.* Second edition (first published in 1936). New York: Dover.

Bealer, G. 1984, "Mind and Anti-Mind: Why Thinking Has No Functional Definition." *Midwest Studies in Philosophy* 9:283-328.

Bennett, J. 1971. *Locke, Berkeley, Hume: Central Themes.* Oxford University Press.

Bennett, J. 1976. *Linguistic Behavior.* Cambridge University Press.

Block, N. 1978. "Troubles with Functionalism." In *Minnesota Studies in the Philosophy of Science, volume 9, Perception and Cognition: Issues in the Foundations of Psychology,* ed. C.W. Savage. Minneapolis: University of Minnesota Press.

Block, N. 1986. "Advertisement for a Semantics for Psychology." *Midwest Studies in Philosophy* 10: 615-678.

Block, N. 1987. "Functional Role and Truth Conditions." *Proceedings of the Aristotelian Society* suppl. 61: 157–181.

Burge, T. 1979. "Individualism and Psychology." *Philosophical Review* 95: 3-45.

Churchland, Paul. 1979. *Scientific Realism and the Plasticity of Mind.* Cambridge University Press.

Churchland, Patricia. 1986. *Neurophilosophy.* Cambridge, Mass: MIT Press. A Bradford Book.

Cummins, R. 1975a. "Functional Analysis." *Journal of Philosophy* 72: 741-760.

Cummins, R. 1975b. "The Philosophical Problem of Truth-of." *Canadian Journal of Philosophy* 5: 103-122.

Cummins, R. 1979. "Intention, Meaning and Truth Conditions." *Philosophical Studies* 35: 345-360.

Cummins, R. and Dietrich, E., 1982. "PATHFINDER: Investigating the Acquisition of Communicative Conventions." In *Proceedings of the Cognitive Science Society.*

Cummins, R. 1983. *The Nature of Psychological Explanation.* Cambridge, Mass.: MIT Press. A Bradford Book.

Cummins, R. 1986. "Inexplicit Information." In *The Representation of Knowledge and Belief,* ed., M. Brand and R.M. Harnish. Tucson: University of Arizona Press.

Cummins, R. 1987. "Why Adding Machines Are Better Examples than Thermostats: Comments on Dretske's 'The explanatory role of content.' " In *Contents of Thought: Proceedings of the 1985 Oberlin Colloquium in Philosophy*. Tucson: University of Arizona Press.

Cummins, R., and G. Schwarz. 1988. "Prospects for a Radical Connectionism." *Southern Journal of Philosophy 26 (Supplement, Spindel Conference, 1987: Connectionism and the Philosophy of Mind)*: 43-62.

Davidson, D. 1975. "Thought and Talk." In *Mind and Language*, ed. S. Guttenplan. Oxford University Press.

Dennett, D. 1978. *Brainstorms*. Cambridge, Mass.: MIT Press. A Bradford Book.

Dennett, D. 1987. The Intentional Stance. Cambridge, Mass.: MIT Press. A Bradford Book.

Dretske, F. 1981. Knowledge and the Flow of Information. Cambridge, Mass.: MIT Press. A Bradford Book.

Dretske, F. 1983. "Precis of Knowledge and the Flow of Information." *Behavioral and Brain Sciences* 6, no. 1: 55-63.

Dretske, F. 1986. "Misrepresentation." In *Belief*, ed. R. Bodgan. Oxford University Press.

Dretske, F. 1987. "The explanatory role of content." In *Contents of Thought: Proceedings of the 1985 Oberlin Colloquium in Philosophy*. Tucson: University of Arizona Press.

Dretske, F. 1988. *Explaining Behavior*. Cambridge, Mass.: MIT Press. A Bradford Book.

Dreyfus, H., and S. Dreyfus. 1985. *Mind Over Machine*. New York: Free Press.

Field, H. 1971. "Tarski's Theory of Truth." *Journal of Philosophy* 69: 347-375.

Field, H. 1977. "Logic, Meaning and Conceptual Role." *Journal of Philosophy* 74: 379-409.

Field, H. 1978. "Mental Representation." *Erkentniss* 13: 9-61.

Fodor, J. 1975. *The Language of Thought*. New York: Crowell.

Fodor, J. 1981. *Representations*. Cambridge, Mass.: MIT Press. A Bradford Book.

Fodor, J. 1983. *The Modularity of Mind*. Cambridge, Mass.: MIT Press. A Bradford Book.

Fodor, J. 1984a. "Why Paramecia Don't Have Mental Representations." *Midwest Studies in Philosophy* 10: 3-23.

Fodor, J. 1984b. "Semantics Wisconsin Style." *Synthese* 59: 231- 50.

Fodor, J. 1985. Psychosemantics. Unpublished manuscript.

Fodor, J. 1987. *Psychosemantics*. Cambridge, Mass.: MIT Press. A Bradford Book

Gregory, R. 1970. *The Intelligent Eye*. New York: McGraw-Hill.

Grice, H. P. 1957. "Meaning." *Philosophical Review* 66: 377-388.

Grice, H. P. 1968. "Utter's Meaning, Sentence-Meaning, and Word- Meaning." *Foundations of Language* 4: 225-242.

Grice, H. P. 1969. "Utter's Meaning and Intentions." *Philosophical Review* 78: 147-177.

Harman, G. 1982. "Conceptual Role Semantics." *Notre Dame Journal of Formal Logic* 23: 242-256.

Harman, G. 1987. *Change in View*. Cambridge, Mass.: MIT Press. A Bradford Book.

Haugeland, J. 1978. "The Nature and Plausibility of Cognitivism." *Behavioral and Brain Sciences* 2: 215-260.

Haugeland, J. 1981. "Semantic Engines." In *Mind Design*, ed. J. Haugeland. Cambridge, Mass.: MIT Press, 1981. A Bradford Book.

Haugeland, J. 1985. *Artificial Intelligence: The Very Idea*. Cambridge, Mass.: MIT Press. A Bradford Book.

Leeds, S. 1982. Causation and the Mental. Unpublished manuscript.

Lewis, D. 1969. *Convention*. Cambridge, Mass.: Harvard University Press.

Loar, B. 1982. "Conceptual Role and Truth Conditions." *Notre Dame Journal of Formal Logic* 23: 272-283.

Loar, B. 1985. "Social Content and Psychological Content." In *Contents of Thought: Proceedings of the 1985 Oberlin Colloquium in Philosophy*. Tucson: University of Arizona Press.

Lycan, W. 1981. "Form, Function and Feel." *Journal of Philosophy* 78: 24-49.

Lycan, W. 1987. *Consciousness*. Cambridge, Mass.: MIT Press. A Bradford Book.

Marr, D. 1982. *Vision*. New York: Freeman.

McGinn, C. 1982. "The Structure of Content." In *Thought and Content*, ed. A. Woodfield. Oxford University Press.

Millikan, R. 1984. *Language, Thought and Other Biological Categories*. Cambridge, Mass.: MIT Press. A Bradford Book.

Millikan, R. 1986. "Thoughts Without Laws: Cognitive Science Without Content." *Philosophical Review* 95: 47-80.

Papineau, D. 1985. "Representation and Explanation." *Philosophy of Science* 51: 550-572.

Parfitt, D. 1984. *Reasons and Persons*. Oxford University Press.

Pollock, J. 1987a. *Contemporary Theories of Knowledge*. Totowa, N.J.: Rowman and Littlefield.

Pollock, J. 1987b. "How to Build a Person." *Philosophical Perspectives* 1: 109–154.

Putnam, H. 1975. "The Meaning of Meaning." In *Language, Mind and Knowledge*, ed. K. Gunderson. Minneapolis: University of Minnesota Press.

Putnam, H. 1983. "Computational Psychology and Interpretation Theory." In *Realism and Reason*, volume 3 of Philosophical Papers. Cambridge University Press.

Pylyshyn, Z. 1984. *Computation and Cognition*. Cambridge, Mass.: MIT Press. A Bradford Book.

Quine, W. V. O. 1960. *Word and Object*. Cambridge, Mass.: MIT Press.

Rumelhart, D., and J. McClelland. 1982. "An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement of effect and some tests and extension of the model." *Psychology Review* 89: 60-94.

Rumelhart, D., J. McClelland, and the PDP Research Group. 1986. *Parallel Distributed Processing* (two volumes). Cambridge, Mass.: MIT Press. A Bradford Book.

Schiffer, S. 1981. "Truth and the Theory of Content." In *Meaning and Understanding*, ed. H. Parret and J. Bouraresse. Berlin: Walter de Gruyter.

Schiffer, S. 1982. "Intention Based Semantics." *Notre Dame Journal of Formal Logic* 23: 119-159.

Schwarz, G. 1987. Explaining Cognition as Computation. Master's Thesis, University of Colorado.

Searle, J. 1983. *Intentionality*. Cambridge University Press.

Smolensky, P. 1988. "The Proper Treatment of Connectionism." *Behavioral and Brain Sciences* 11, no. 1: 1–74.

Smolensky, P. Forthcoming. *Lectures on Connectionist Modeling*. Hillsdale, N.J. Erlbaum.

Stalnaker, R. 1984. *Inquiry*. Cambridge, Mass.: MIT Press. A Bradford Book.

Stich, S. 1983. *From Folk Psychology to Cognitive Science*. Cambridge, Mass.: MIT Press. A Bradford Book.

Tarski, A. 1956. "The Concept of Truth in Formalized Languages." In *Logic, Semantics and Metamathematics*, ed. A. Tarski. Oxford University Press.

Turing, A. 1950. "Computing Machinery and Intelligence." *Mind* 59: 434-460.

Ullman, S. 1979. *The Interpretation of Visual Motion*. Cambridge, Mass.: MIT Press.