Mental Pictures and Cognitive Science

Ned Block

The Philosophical Review, Volume 92, Issue 4 (Oct., 1983), 499-541.

The Philosophical Review, XCII, No. 4 (October 1983)

MENTAL PICTURES AND COGNITIVE SCIENCE

Ned Block

Cognitive scientists have made some amazing claims about mental images. We are told that people *rotate* mental images, and at measurable speeds. Three-dimensional shapes were rotated in one experiment at 60 degrees per second, while letters of the alphabet have been clocked at 800 degrees per second. We are told that mental images can be scanned, just as are scenes and pictures—and again at measurable speeds. Smaller mental images are supposed to be harder to see than larger ones, and when mental images are made larger, they eventually overflow. Further, what they overflow *from* has a determinable shape—roughly elliptical. There's more. Images subtend a determinate visual angle, which, at the point of overflow, is described as the angle of vision of the mind's eye (roughly 20 degrees, as it happens).[1]

Such claims are part of a viewpoint according to which mental images represent in the manner of pictures. It is very natural to think that such claims are confused or nonsensical. One of my purposes here is a limited defense of this supposedly confused doctrine, especially against its chief cognitive science rival. But this paper has another quite distinct topic: the concept of representation in general (not just pictorial representation) as it functions in cognitive science. I put the two together by arguing that it *matters* to cognitive science whether it has to encompass pictorial representations as well as the standard strings of zeros and ones (and similar language-like or descriptive representations). If cognitive science must postulate pictorial representations in the head, then cognitive science may be in serious trouble, for much of what it hopes to explain will probably be in the domain of a different discipline: neurophysiology.

[1]Work by Roger Shepard and his colleagues is the basis for the claim about rotation; the others stem from Stephen Kosslyn's work. See R. N. Shepard and L. Cooper, *Mental Images and Their Transformations* (Cambridge: MIT Press, 1982). See also the works mentioned in the next three footnotes. Kosslyn's work is summarized in his *Image and Mind* (Cambridge: Harvard University Press, 1980).

It will be useful for those who know nothing of the experiments that motivate the picture-in-the-head view for me to sketch a few of them. Readers who are familiar with these experiments can skip the next three paragraphs.

Block Figure Rotation

The three pairs of block figures in Figure 1 are from an experiment by Shepard and Metzler.[2] Your task is to look at pair A and say whether the objects pictured are congruent or not. Do the same for B, and then for C. The two drawings in A are the same except that they are pasted on the page in different orientations. In B we have two-dimensional projections of objects that differ not by a rotation in the plane of the page as in A, but rather by a rotation out of the page (i.e., around an axis in the page). In C, the figures are noncongruent mirror images. Subjects in this experiment were presented with 1600 such pairs and asked to pull one lever if the objects were congruent, and another if not. Result: the greater the angular separation of the congruent figures, the longer it took subjects to respond. Indeed, the time to respond was a *linear* function of angle of separation, independently of whether the rotation was in the plane of the page or perpendicular to it. (So if this experiment suggests pictures in the head, they will have to be 3-D pictures.) Subjects reported imagining one object rotated so as to be superimposable on the other. This experiment is often taken to indicate that people can rotate mental images, and at measurable speed. Metzler later chose a group of these subjects whose rotation speeds were relatively constant (not variable), and calculated their individual speeds. Then she gave subjects single figures (not pairs), and asked them to start rotating in a clockwise direction. Next she presented *rotated* versions of the original stimulus ("probes") in accordance with calculations based on the subject's known speed. For example, if the subject rotated at 50 degrees per second, she might present a 50-degree rotated figure after one second, or a 100-degree rotated figure after two seconds, or a 75-degree rotated figure after 1.5 seconds. Subjects were instructed to say

[2] J. Metzler and R. N. Shepard, "Mental Rotation of Three-dimensional Objects," *Science*, 171 (1971): pp. 701–703.

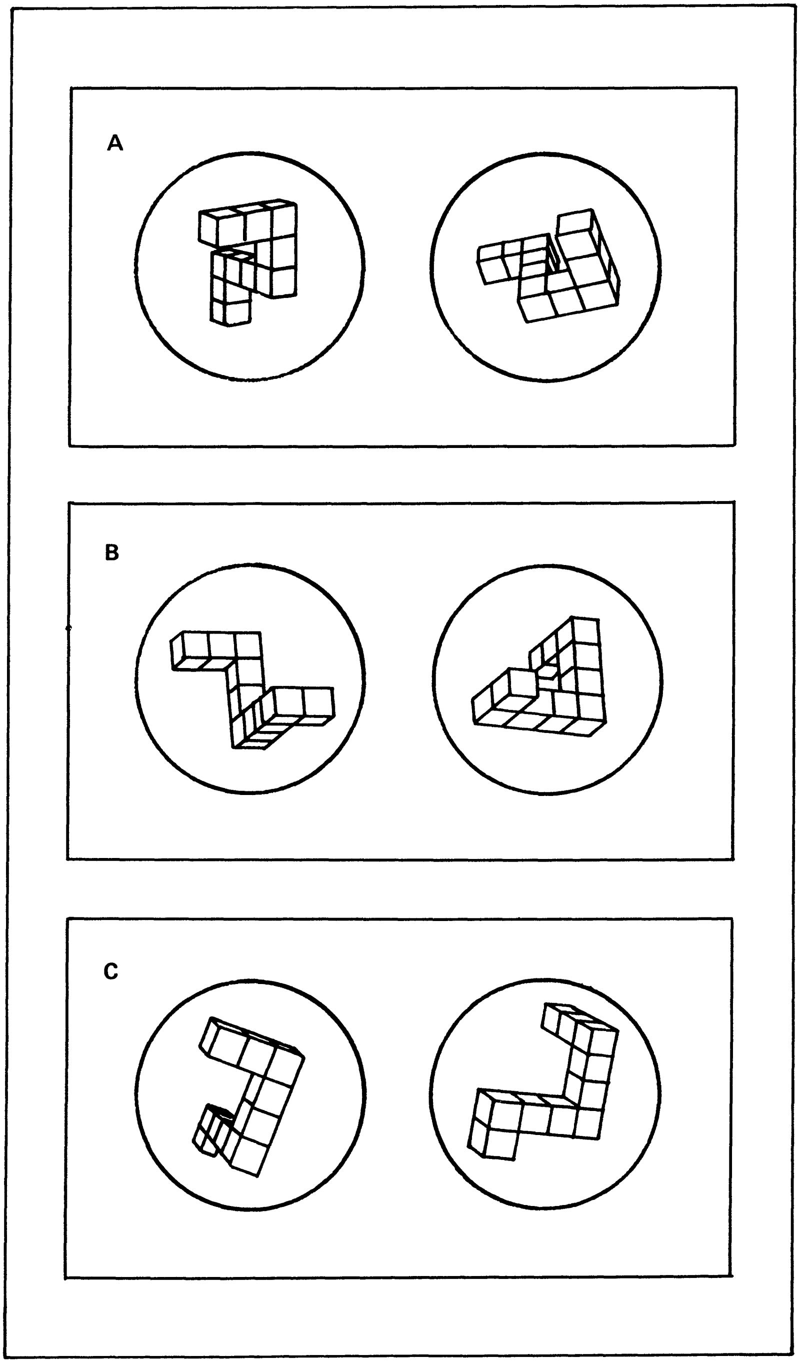

Figure 1.
(This figure was reprinted with the permission of *Science,* Vol. 171, no. 3972 (February, 1971), p. 701.)

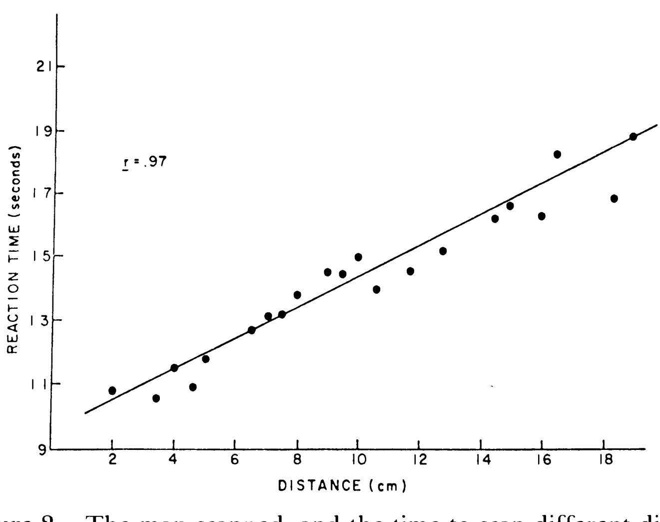

Figure 2. The map scanned, and the time to scan different distances across an image of it. (This figure was reprinted with the permission of Stephen M. Kosslyn from his paper, "On the Demystifaction of Mental Imagery," p. 137 of Ned Block, ed., *Imagery,* Cambridge: MIT Press, 1981.)

whether the original figure (whose image they had been asked to rotate) and the "flashed" rotated figure were congruent or not. What she found was that the time it took subjects to respond (to congruent probes) was always about the same, independently of the angular separation of the first and second figure. That is, whether she presented a 50-degree rotated figure after one second or a 100-degree rotated figure after two seconds made no difference to response time. And the response times were short—about the same as the shortest response times of the Shepard and Metzler experiment just described. This is interpreted as indicating that she *caught* their rotated images at *just the right point* with her probe.[3]

MAP SCANNING

Subjects studied the map in Figure 2 until they could draw it, getting the seven locations marked by red dots (Xs in Figure 2) approximately right. In one of the experiments done by Kosslyn and his colleagues, subjects were asked to "zoom in" on one of the objects on the map, say the well, until that object "filled their entire image."[4] They were told that they would soon hear the name of another object that might or might not be on the map. They were instructed to "see" the named object (if it was in fact on the map), pressing a button when they had done so. All pairs of the seven objects were used. Result: time taken to "see" the second object was directly related to the distance on the map between the two objects (correlation = .89). This result is taken to indicate that mental images can be *scanned*. Further, this experiment is used to support the idea that parts of mental images can be *unconscious*. For when subjects "zoomed in" on an object on the map, they were not aware of imaging the *other* objects on the map. Yet their scanning times were far faster than would be expected if they were generating

[3]J. Metzler, unpublished Stanford University Ph.D. thesis, 1973. Described in Shepard and Cooper, op. cit., pp. 59–62.

[4]T. Ball, S. Kosslyn, and B. Reiser, "Visual Images Preserve Metric Spatial Information: Evidence from Studies of Image Scanning," *Journal of Experimental Psychology: Human Perception and Performance* 4 (1978): pp. 47–60. Figure 2 was also reprinted with the permission of the American Psychological Association, Copyright, 1978.

images of the other objects anew—based on independent measurements of how long it takes to produce images.

In another scanning experiment, Finke and Pinker showed subjects patterns made up of dots. They turned off the dot pattern, asked the subject to form an image of the pattern, and then flashed an arrow, which, in half the trials, pointed to where a dot had been. Subjects were asked to indicate if the arrow did point at one of the imagined dots. If the arrow did point towards a dot on the original pattern, time taken to answer was directly related to the distance between the tip of the arrowhead and the dot (mistakes were rare).[5]

How do these experiments support the picture-in-the-head view? If it were an intelligent *robot* that exhibited this sort of phenomenon, we would surely be tempted by the hypothesis that the robot was constructing, manipulating and examining pictorial representations somewhere inside its body. Many of us are less tempted to postulate internal pictures in *us,* partly because we have tentative competing hypotheses about what is happening in the experiments and partly because we know no pictures have been found in our brains (more about these matters later).

Experiments such as the ones mentioned above have given rise to considerable controversy. One side, *pictorialism,* supposes that the mental representations of imagery represent in the manner of pictures, for example, of President Reagan being shot. The rival, *descriptionalism* (a popular view, especially in the artificial intelligence community), says that the representations of imagery (and

[5]R. Finke, and S. Pinker, "Spontaneous Imagery Scanning in Mental Extrapolation" *Journal of Experimental Psychology: Learning, Memory and Cognition* 8, no. 2 (1982): pp. 142–147 and R. Finke and S. Pinker, "Directional Scanning of Remembered Visual Patterns," *Journal of Experimental Psychology: Learning, Memory and Cognition,* forthcoming in June, 1983. Zenon Plyshyn has objected to Kosslyn's map scanning experiments on the ground that subjects merely mimic features of perceptual interactions of which they have tacit knowledge. They know it takes longer to move their gaze a longer distance, so when they recreate the experience, they make the longer image scan take longer. However, Finke and Pinker in the latter of the two papers just mentioned show that subjects exhibit patterns of *errors* in their dot scanning experiment of a rather surprising sort: there are more errors closer to the point of the arrow. This can be predicted, given the instruction set (and it happens in actual visual scanning), but it is doubtful that it can be explained via tacit knowledge.

all other internal representations as well) represent in the manner of familiar symbol structures such as those in a computer or those in English (e.g., the sentence 'President Reagan was shot'.)[6]

Now that the reader has caught a glimpse of the pictorialism/descriptionalism controversy, I can better explain the plan of this paper. The paper has two quite different argumentative strands. The first, which begins after ground-clearing remarks (on perception, the vagueness of 'mental image', and introspection), is a defense of pictorialism from three rather flat-footed objections:

(1) *The No Seeum Objection.* When we look in the brain, we don't see any pictures.

(2) *The Leibniz's Law Objection.* Mental images can be orange- and pink-striped, but bits of the brain never have such properties.

(3) *The Paraphernalia Objection.* Rotating and scanning pictures in the head would require an internal eye, hands, and other things that are not in fact to be found in the head.

My strategy is to argue that the sort of problem indicated by these objections is one to which descriptionalism is just as vulnera-

---

[6]I shall use 'picture' as short for 'pictorial representation', thus speaking of pictorialism as postulating internal pictures. No commitment to paper or canvas is intended. For a look at the pictorial/descriptional debate, see Zenon Pylyshyn's "Imagery and Artificial Intelligence," and Stephen Kosslyn's and James Pomerantz's "Imagery, Propositions and the Form of Internal Representations," in my *Readings in Philosophy of Psychology,* Vol. 2 (Cambridge: Harvard University Press, 1981). Pylyshyn is a descriptionalist; Kosslyn and Pomerantz are pictorialists. For a second Pylyshyn vs. Kosslyn faceoff, see Pylyshyn's "The Imagery Debate: Analog Media versus Tacit Knowledge" and Kosslyn's reply, "The Medium and the Message in Mental Imagery: A Theory" in my *Imagery* (Cambridge: Bradford/MIT Press, 1981). See also my "The Photographic Fallacy in the Debate about Mental Imagery," *Nous* Vol. 17, No. 4 (1983).

An interesting pictorialist theory of imagery is presented in R. N. Shepard, "Psychophysical Complementarity," in M. Kubovy and J. Pomerantz, eds., *Perceptual Organization* (Hillsdale, N.J.: Erlbaum Associates, 1981) pp. 279–341. A good recent review of descriptionalist and pictorialist theories is given by S. Pinker and S. M. Kosslyn, in "Theories of Mental Imagery," A. Sheikh, ed., *Imagery: Current Theory, Research and Application* (New York: Wiley, 1983).

ble as is pictorialism, and, more importantly, that there are defenses available to the descriptionalist which are also available to the pictorialist. In short, pictorialist and descriptionalist cognitive science must stand or fall together against such objections. This mode of defense of pictorialist and descriptionalist cognitive science brings in my second and more important argumentative strand. I argue that the common defense of pictorialist and descriptionalist cognitive science reveals a potentially serious limitation on the explanatory power of theories in cognitive science. The key concept in this argument is the concept of a primitive process, a process whose input-output function is appealed to by cognitive science explanations, but whose mode of operation is not itself within the domain of cognitive science. The mode of operation of a primitive process is rather in the domain of a "lower level" science: neurophysiology, in the case of humans. The potentially serious limitation comes in with the possibility that primitive processes may bear most of the burden in explaining human thought. If so, cognitive science "bottoms out" in primitive processors much sooner than expected.

In the course of presenting this point, I shall point out that cognitive science has a ready defense against Wittgensteinian criticisms about representations and rules, but that the employment of this defense leads once again to the potentially serious limitation just mentioned.

THE MEANING OF 'MENTAL IMAGE'

The term 'mental image' is too vague to allow a determinate answer to the question of what the nature of mental images is. In the controversy over what the mental imagery experiments show, 'mental image' is *usually* used to denote the internal representations involved in mental imagery, but even in this literature, 'mental image' sometimes denotes the *experiences* we have in imagery, sometimes mental states that *include* these experiences plus more, and sometimes even abstract "imaginary" objects. This sort of vagueness is pretty common in the history of science. Seventeenth-century Europeans did not distinguish among heat, temperature, and even perceived temperature. Often, a single word was used for all three in situations in which context did not disambiguate. I think construing 'mental image' as denoting the internal representations involved in imagery helps to make sense of the controversy over whether mental images are pictures in the head.

To avoid confusion, it may be useful to distinguish mental images in my sense (the internal representations of imagery) from what might be called *phenomenal* images, the "seeming objects of image experiences," the things, if such there be, which are more or less vivid, sometimes orange-and-pink-striped, etc. (even though they reflect no light).[7] Much of the discussion in philosophy over whether mental images *exist* or are objects is usefully construed as concerning phenomenal images, and perhaps not mental images in my sense. I happen to think that phenomenal images are identical to mental images in my sense, though I shall not assume this is so (except to rebut an objection that assumes it) or argue for it here. If phenomenal images are not identical to mental images, then I don't know (or care) what phenomenal images are. If phenomenal images *are* identical to mental images, then according to both pictorialist and descriptionalist cognitive scientists, they are symbols or symbol-structures in the head.

Imagery and Perception

The experiments that motivate the pictorialist perspective are often taken to support a different hypothesis: that the representations and processes of imagery are like the representations and processes of *perception*. This hypothesis is often confused with pictorialism, in part because many people seem to find it difficult to understand the possibility that the representations of perception *might not be pictorial* (more on this later). There are many experimental results that support the similarity of imagery and perception without supporting pictorialism. For example, it is known that visual acuity for vertical stripes is better than for oblique stripes.

[7]The distinction and the quoted phrase are from Georges Rey's "What are Mental Images?" which appears in my *Readings in Philosophy of Psychology,* Vol. II, op. cit. A closely related distinction between metaphysical and scientific iconophobia appears in Dan Dennett's "Two Approaches to Mental Images," *Brainstorms* (Montgomery, Vermont: Bradford Books, 1978), (reprinted in my *Imagery,* op. cit.).

This is called the oblique effect. If the experimenter holds up a grating and slowly moves it away, asking subjects to indicate when the stripes blur, subjects report the vertical stripes beginning to blur further away than the oblique stripes. Amazingly, the oblique effect occurs when the subjects are asked to *imagine* a grating moved slowly away from them. If the imagined grating is vertical, blurring occurs at a greater subjective distance than if it is oblique.[8]

Many perceptual phenomena such as the oblique effect have been found to occur with imagery.[9] While the experimental literature does not distinguish as clearly as one would like between the effects of representations and the effects of the processes that manipulate representations, the evidence is now strong that the representations of imagery and perception are of the same kind.

But—and here is the point—the claim that the representations of imagery and perception are of the same kind is irrelevant to the controversy over pictorialist vs. descriptionalist interpretation of experiments like the image scanning and rotation experiments mentioned earlier. The representations of imaging and of perceiving should be lumped together as either *both* pictorial or both descriptional. But we still want to know *which!* The descriptionalist is committed to regarding the representations of perception as descriptional (given the evidence just mentioned).[10] By lumping the representations of imagery and perception together, we reduce two problems to one problem, but we have not thereby solved the one problem.

---

[8]Pennington and Kosslyn did this experiment. See Kosslyn's "The Medium and the Message in Mental Imagery," op. cit.

[9]See R. A. Finke, "Levels of Equivalence in Imagery and Perception," *Psychological Review* 187, no. 12 (1980), and R. M. Shepard, "The Mental Image," *American Psychologist* 133 (1978). I should add that there are considerable differences among people in their susceptibility to these phenomena—and these differences correlate with independent measures of the vividness of their images and with ability to reparse images. See Finke, op. cit., and J. A. Slee, "Individual Differences in Visual Imagery Ability and the Retrieval of Visual Appearances," *Journal of Mental Imagery* 4 (1980): pp. 93–113.

[10]It is not at all unusual for descriptionalists to recognize this. See Pylyshyn's "Imagery and Artificial Intelligence" (op. cit.) for example. Descriptionalist views of perception are common in artificial intelligence approaches to vision.

Imagery and Introspection

The picture-in-the-head theory of images profits from an apparent *convergence* of experiment and introspection. As we look inward, our mental images often seem to us to be pictures in the head, and by golly, when the experimental results come in, they back up this introspective judgment. Part of the basis of the introspective judgment could be put like this, "Of *course,* my mental image of my daughter is pictorial; it looks just like this photograph of my daughter; I carry one around in my wallet and the other in my head." I want to reject firmly this line of reasoning; introspective judgment in this case is a cover for fallacious reasoning.

When someone says his mental image looks like a picture, he must be appealing to the fact that his image experience is like the experience he gets on looking at the picture. But this is a poor reason to suppose that his image *is* a picture. After all, the image experience is *also* like the experience he gets on looking at his daughter *herself.* And one could as well argue from this fact that his image is a *person.*

If I experience x and y as similar, this could be due to a similarity in x and y, but it could be due instead to something about me and my relation to x and y. I experience my psychiatrist and my father in the same way, but the explanation is not any similarity between my father and my psychiatrist. I would have experienced the psychiatrist as like my father even if either of them had been quite different along almost any dimension you choose. I experience the two as similar because of a neurosis having to do with my father, and the familiar transference of the psychiatric interaction.

Returning to the similarity between the way I experience my mental image of my daughter and a photo (or a wax dummy) of her, we *have* a readily available explanation in terms of something about me, so we don't have to postulate a similarity between the image and the photo or wax dummy. The available explanation appeals first to the fact that the representations of imagery and perception can be experientially similar. Second, a glimpse of a good wax dummy of my daughter can give rise to the same retinal stimulation as a glimpse of my daughter herself, and hence the same visual representations. So we can have a similarity between imaging and perceiving my daughter, and a similarity between perceiving my daughter and perceiving a 3-D or a 2-D picture of her. So the similarity between imaging my daughter and perceiving a picture of her should be no surprise.

Perhaps what I've just said is not quite enough to quiet the claims of introspection. Consider the following reply:

> Your account of the basis for our intuitive judgment that mental images are pictorial just sets up a straw man. Here is the real story:
>
> (1) My mental image of my daughter resembles my daughter, as is typical of pictorial representations of my daughter. Further, there is a detailed point by point correspondence, for example, the top of the image resembles the top of my daughter, the bottom of the image resembles the bottom of my daughter, etc. This is typical of *good* pictorial representations of my daughter.
> (2) My mental image of my daughter resembles good pictures of my daughter (again, in a point by point fashion) just as each picture typically resembles the others.
> (3) My image does not resemble either words, phrases or sentences about my daughter, nor do any of the other types of pictorial representations of her (movies, wax dummies) resemble language about her.
> (4) Assuming that my mental image is a representation, and that we are deciding whether to categorize it as pictorial (along with photos and wax dummies) or as descriptional (along with words and phrases) it would therefore be idiotic to choose the latter.

Rejoinder: if we were shown a new form of representation (e.g., a hologram, when they were first invented) which we see in the manner in which we see objects, including pictures, sentences, and my daughter herself, and if the new representation looked like my daughter and pictures of her, but unlike language about her, we might well be reasonably persuaded to classify the new form of representation as pictorial. The trouble with this reasoning as applied to mental images is that we don't *see* them. My daughter, a photo of her, and the phrase 'my daughter' are things we see; mental images, by contrast, are things we *have*. The argument depends on a false parallel between our relation to our mental images on the one hand, and our relation to pictures, people, and printed words on the other. Perhaps the representations we have when we perceive—the representations that mediate perception—are themselves descriptional; this is the crucial possibility that the introspectionist argument fails to take account of.

In sum, the fact to start with is that we have both introspective and experimental evidence for similarity of representations of imagery and perception: imaging my daughter and seeing her can yield functionally and experientially similar representations. Now a good picture of my daughter is designed to produce a representation experientially like that produced by seeing my daughter. *That* is why my image "looks like" the picture. The important point is that this explanation of why my image "looks like" the picture assumes nothing about whether the internal representations are pictorial or descriptional.

No Seeum Objection

Perhaps the most straightforward objection to the view that mental images are pictures in the head is simply that when one *looks* in the head, one doesn't see any pictures. This is the no seeum objection. In locating what is wrong with this objection, it is useful to note that it is *not* so tempting to make this sort of objection against the chief cognitive science rival to the picture in the head view, descriptionalism.

Now it is clear enough why descriptionalism is immune from the argument that we don't *see* descriptional representations when we look in the brain. Before you can recognize a descriptional message, two conditions must be satisfied. First, you must know the representational system, and second, there must be a visible message in that representational system. The hairs on the spy's head are arranged so as to spell out a 500-word message for his barber, his link to the KGB. But *you* don't see any message when you look at his hair.

Here is my point: the reply that protects descriptionalism from the no seeum objection equally protects pictorialism. What *any* representation represents, and *how* it represents—pictorially or descriptionally—depends on the system of representation within which it functions. Ink marks that your tribe uses as a picture could be a letter in *my* tribe's descriptional alphabet. Perhaps some of the inscriptions that archeologists treat as pictures that give insight into long-dead civilizations really functioned originally as symbols in a language. Conversely, imagine an ancient symbol system that archeologists think is a language, one that they have successfully decoded. They write each other letters in this language, and even form a club where only this language is spoken. But actually the language is their own creation; the original symbols they found were really pictures.[11]

This relativity to system of representation also holds *within* the category of descriptional representations and also to some extent within the category of pictorial representations. In the descriptional case, it is illustrated by Davidson's example: the sound we make in saying "Empedocles leaped" means one thing in English and another in German. In the pictorial case, cross-cultural study of picture perception reveals some cultural relativity, though of course nothing like the degree of conventionality found in language. For example, a picture which we see as having an elephant in the background was seen in another culture as having a very small elephant suspended in air in the foreground.[12]

My point is not just that what represents what is relative to a system of representation. Rather, my point is that you can't tell for sure whether you are looking at a representation at all *just* by looking. Whether a thing is a representation at all, and if so, whether it is a pictorial representation or a descriptional representation, cannot be determined if one limits one's attention to the thing itself. One has to determine how the thing functions. Since we know next to nothing about how representations in the brain function, we can hardly expect to tell, just by looking, *which* struc-

[11]I had planned a more detailed exposition of this point, but dropped it when Robert Schwartz sent me his "Imagery—There is More to It than Meets the Eye," which contains an excellent treatment. Schwartz's article is in *Imagery,* op. cit, and is forthcoming in Asquith and Giere, eds., *PSA 1980,* Vol. 2.

[12]Cultural relativity of picture perception is discussed in R. J. Miller, "Cross-cultural Research in the Perception of Pictorial Materials," *Psychology Bulletin* 80 (1973). Gombrich has long stressed a limited cultural relativity of systems of pictorial representation. See, for example, his "Standards of Truth: The Arrested Image and the Moving Eye," in *The Language of Images,* edited by W. J. T. Mitchell (Chicago: University of Chicago Press, 1980).

tures in the brain are representations, much less how brain representations represent.

Now that I have allowed that the pictorialist's pictures in the head can't just be *seen* to be pictures, and the descriptionalist's sentences in the head cannot just be *seen* to be sentences, surely I must say something about what the difference between pictures and sentences in the head is supposed to be. What I would like to do here is introduce an adequate account of the pictorial/ descriptional distinction. Unfortunately, though many accounts are in the literature, none that I know of are very satisfactory.

One condition on pictorial representation that has something going for it is this: if a picture represents something, S, then at least one part of the picture must represent part of S; further, typically some of the relations among the parts of the picture represent relations among the parts of S, and some relations between parts of the picture and the whole picture represent relations between the parts of S and S itself.[13] Obviously, such a condition need not apply to descriptional representations. 'The first son of Elizabeth II' refers to Charles, but no part of it refers to part of Charles. Though this condition has some appeal, I do not want to place a lot of weight on it. For one thing, it is doubtful as a sufficient condition. (Part of 'Britain's empire' refers to part of the empire.) Secondly, in the cases in dispute, there is generally as much of a problem about the application of the condition itself as there is about the application of 'picture'. Finally, this account is particularly unrevealing with respect to the very issue I've just been talking about: the dependence of a thing's representational properties on the system of representation in which it functions. The condition takes for granted 'represents' and 'part', concepts which are themselves functional.

I do not know how to characterize what it is about function within a system of representation that makes a representation pictorial or descriptional. The best I can do is to try to evoke a feeling for the difference by means of examples.

[13]Such a criterion figures prominently in S. Kosslyn, S. Pinker, S. Schwartz, and G. Smith, "The Demystification of Mental Imagery," *The Behavioral and Brain Sciences* 2 (1979); in Kosslyn's *Image and Mind*, op. cit.; and in Aaron Sloman's *The Computer Revolution in Philosophy* (Sussex: Harvester, 1978).

Consider what goes into a computer by way of storing a picture of, say, a vertical line. Consider a matrix n squares wide. Each of the squares can be light or dark. A vertical line would be represented if each lighted square is n squares distant from the first lighted square (counting by row, as on a calendar). The descriptionalist sees the line representation as a set of sentences. If n = 7, the set of sentences could be: '1 is dark', '2 is light', '3 is dark', . . . '9 is light', '10 is dark', and so on. The descriptionalist sees a representation of a line in the *brain* along the same lines as a set of descriptional representations, akin to a set of sentences. The pictorialist, on the other hand, sees the representation of a line in the brain as like the *matrix itself,* not the corresponding set of sentences.

But what does the distinction come to between the set of sentences and the matrix itself? Aren't they just *different ways* of inscribing the same representation with the same semantic properties?

I say "No." The key is the way the representations *function.* Consider how the descriptionalist would *rotate* his line. A small counterclockwise rotation could be accomplished if the first lighted square stayed lit, the next number of a lighted square was increased by 1, the next by 2, the next by 3, and so on. In terms of the example, 2 would stay lit, 10 would be lit instead of 9, 18 instead of 16, 26 instead of 23, etc. So the descriptionalist's new set of sentences would be '1 is dark', '2 is light', . . . '10 is light', . . . '18 is light', . . . '26 is light', and so on. The important point is that the computer's "rotation" calculation just involves the *numbers,* not the arrangement of the numbered squares. The matrix display is *for us* and plays no role in what the computer does. From the point of view of the computer's calculations, the squares *could as well be arranged in a line or a circle rather than in a matrix.* Real live computer graphics works this way. The machine manipulates numbers. The programmers think of the numbers as numbers of cells in matrices, and they put in numbers that correspond to matrices of visual interest. They program the computer to operate on the numbers in ways that correspond to visually interesting or useful matrix changes. These correspondences are what makes the computer's number crunching *graphical,* but the correspondences play no role at all in the number crunching itself.

Once we see what the computer *does,* we realize that the representation of the line is descriptional. But what sort of process would show that a representation is pictorial? I don't have a ready answer, since we have no machines that manipulate pictorial representations in the systematic ways that computers manipulate descriptional representations. But I can try to give the flavor with an example, albeit a physiologically absurd one.

Imagine that a space is to be represented by electrical potential and current at points in the visual cortex. The system of representation is polar coordinates, and current represents angle while potential represents distance from the origin. (For those who have forgotten all they learned in high school, in polar coordinates a point is specified by a pair of numbers, one of which gives the distance from the origin, while the other gives the angle (measured counterclockwise) of the arrow pointing from the origin to the specified point.) Suppose there is a region of the visual cortex in which all points have the same current (and thus fall on a line) but differ in potential, being distributed between 0 and 10 units (so we have a line 10 units long). Suppose also that in its present mode, all the points in this region are linked by circuits that cause current to flow from regions of high current to regions of low current. Result: the linkage keeps the region homogeneous in current value and the points represented stay on a line. Now if the current value is smoothly increased, we would have a smooth counterclockwise rotation. Imagine that the current of all the points is increased as a result of increasing the current of a few points (the linkage among currents causing the other currents to increase). This would be like turning a whole stick by applying torque to a small part of it.

I am emphatically not claiming that it is a *sure thing* that there is a legitimate distinction to be made between pictorial and descriptional representations in the brain. I do think, however, that the possibility of a genuine difference here deserves to be taken seriously. Such a difference, if it obtains, is certainly easier to recognize than to characterize.

Leibniz's Law Objection

The claim that mental images are physical particulars in the head—let alone pictorial particulars—is widely held to be absurd. In the context of an early defense of a general mind-body identity claim, J. J. C. Smart considered the objection that if an image is orange and identical to a brain state, then the brain state must be orange too.[14] For Leibniz's Law reminds us that if x *is* y, then any property of x is a property of y. Smart did not want to be committed to orange neurological entities, so he countered by saying that the identity theorist does not suppose that there are orange mental images, but rather only *experiences* of orange images. The experiences are not orange, so there is no Leibniz's Law problem for materialism here. Materialists who have discussed the matter have been nearly unanimous in supporting one or another variant of a Smart-style "disappearance theory" of mental images. If these materialists are right in supposing mental images aren't physical particulars, then mental images certainly aren't pictures in the head either.

Now the pictorialist can agree with the disappearance theory that if there is a sense of 'mental image' in which it is of the essence of mental images to be, for example, orange or striped or vivid, then no mental images in *that* sense of the term exist. But the pictorialist can go on to disagree strongly with the disappearance theorist, insisting on the existence of mental images in *his* sense of the term (the internal representations of imagery). He can avoid the Leibniz's Law problem by allowing that mental images in his sense are *not* orange, vivid, etc., explaining away our temptation to ascribe colors, shapes, etc., to our mental images by noting that when we say a mental image is orange, what is really the case is that it *represents* something as orange. Having the mental image is phenomenally similar to seeing an orange thing. And this is presumably because the neural entity (with which the materialist pictorialist identifies the mental image) has physiological properties of the sort typically produced by seeing an orange thing.

On this view, it is no surprise that we describe the mental image as orange even though, strictly speaking, it is not. For it is easy to slip into ascribing to representations the properties of what they represent. People who work routinely with graphical representations of sounds (e.g., oscilloscope readings) often speak of them as

---

[14]"Sensations and Brain Processes," *Philosophical Review* 68 (1969): pp. 141–156.

if they had the properties of the sounds they represent—for example, being loud or high pitched.

This line of response to the Leibniz's Law problem meshes with the illustration given earlier of a pictorial representation in the brain, a representation of a rotating straight line. This pictorial representation was not itself literally straight, nor did it literally rotate. Rather, the homogeneity in current value represented straightness, and the smooth change in current value represented rotation.

Though this response to the Leibniz's Law problem is adequate, I want to go on to mention another pictorialist option, one that does *not* take it to be literally false that mental images can be orange.

There is some reason to think that there is a systematic difference in meaning between certain predicates applied to physical objects and the same predicates applied to mental particulars. Consider a nonimagery example: the predicate '____ is in ____'. This predicate appears in each premise and the conclusion of this argument:

The pain is in my fingertip.
The fingertip is in my mouth.
Therefore, the pain is in my mouth.

This argument is valid for the "in" of spatial enclosure, (e.g., the sense of 'x is in y') which means that every point of x is surrounded by points of y (though not conversely), since "in" in this sense is transitive. But suppose that the two premises are true in their *ordinary* meanings: my fingertip hurts and I've thrust the finger deep into my mouth. The conclusion obviously does not follow, so we must conclude that "in" is not used in the spatial enclosure sense in all three statements. It certainly seems plausible that "in" as applied in locating pains differs in meaning systematically from the standard spatial enclosure sense.[15] My suggestion is that the same

[15] This point derives from a draft of a paper on Leibniz's Law objections to materialism that Jerry Fodor and I wrote jointly long ago but never finished.

sort of systematic difference in meaning applies with respect to 'orange' and other terms applied to mental images. 'Orange' has one meaning applied to mental images, and another applied to fruit.

Call the sense of 'orange', etc., in which such terms apply to mental images the *phenomenal* sense. Phenomenal orange represents real orange, just as the orange pigment on a painting represents real orange. Of course, orange representing pigment is orange in the nonphenomenal sense too. Imagine a system of pictorial representation for the blind in which texture represents color. Squishiness, say, represents orange. Squishiness, like phenomenal orange, represents real-world orange without *being* real-world orange.

Objection: "But the orange of orange juice *looks just like* the orange of my orange mental images. So how can you deny that my mental images are orange in the very same sense as orange juice?"

Reply: this is the very fallacy scouted earlier in the section on introspection. Recall that you cannot argue from a similarity in the way I experience x and y to a similarity between x and y. (As I noted, I experience my father and my psychiatrist similarly, but not because they are similar.) I experience orange juice and orange images similarly because both *experiences* involve the same quality (what I call phenomenal orange), not because orange images and orange juice are both orange in the same sense.

This color term ambiguity fits into patterns of less controversial systematic ambiguities. "Nude" (e.g. in 'looking at nudes') is ambiguous between *representation of an unclothed body* and *unclothed body.* 'Painting trees' is ambiguous between *making representations of trees, using paint,* and *putting paint on actual trees.* Similarly, 'orange image' can be viewed as ambiguous between *image that is colored real-orange* (i.e., the color of real orange juice), and *real-orange-representing image,* where the latter is what normally would be meant.

This "ambiguity" line of response to the Leibniz's Law problem differs in one simple respect from the line of response with which this section began. The first line takes talk of orange mental images to be a literally false way of talking of orange-representing images. The present line takes talk of orange mental images to be (sometimes) literally true in one sense of 'orange' (the phenomenal sense), and cashable in terms of orange representing images in another (real world) sense of 'orange'. Both lines fit together with the idea that if two of my images represent real-world orange, then there is some shared neurological property in virtue of which they so represent, and that there are two correspondences (at least in a given person at a given time): one between imageable properties of objects and neurological properties of their brain representations, and another between the relations among imageable properties of objects and the relations among corresponding neurological properties.[16]

* * * * *

Thus far I have been concerned more with pictorialism than with cognitive science. Now I shall shift the emphasis to the presuppositions of cognitive science about internal representations and the processes that operate on them. Then I shall return to pictorialism, bringing to bear my conclusions about representation and processing in cognitive science, arguing that pictorialism may carry the seeds of the destruction of cognitive science as it is currently conceived.

Paraphernalia Objection

Suppose that neurophysiologists discover that the reason that they hadn't seen pictures in the brain is that they hadn't thought of flattening out its highly convoluted surface or of looking at it under ultraviolet light. When they do this, they find the long sought-after pictures: red and white cane-shaped pictures of candy canes, and so forth. Such a discovery would get around the no seeum and the Leibniz's Law objection, but would it really *explain* the phenomena

[16]These correspondences are what Shepard had in mind by "second-order isomorphism." See Shepard's review of Neisser's *Cognitive Psychology* in *The American Journal of Psychology* 81 (1968): pp. 285–289, where he first states this idea. See also R. N. Shepard and S. Chipman, "Second-order Isomorphism of Internal Representation: Shapes of States," *Cognitive Psychology* 1, pp. 1–17. Shepard later shifted his focus from this pictorialist view to the claim that mental imagery representations are similar to perceptual representations, but he has recently returned to second-order isomorphism in his original sense in "Psychophysical Complementarity," op. cit.

illustrated by the experiments described earlier? Obviously not. In one of the experiments, images seemed to be *scanned.* But the presence of pictures in the brain wouldn't explain how images can be scanned without an internal *eye* to do the scanning. And if one acknowledges images of perception, one may need *another* eye in the internal eye's brain. In another experiment, images seemed to be rotated, and in different dimensions. In other experiments, images are expanded and contracted. Pictures in the brain wouldn't explain this unless there were also something to do the rotating and expanding and contracting, for example, internal hands. And since it is dark inside the skull, the eye would require an internal flashlight.

In sum, the presence of pictures in the brain wouldn't explain anything without an internal eye, hands, and other paraphernalia that in fact do not exist. Further, the internal eye claim is not only empirically false, but may lead to a vicious regress as well.

This objection is defused by points already made in connection with the Leibniz's Law objection. Recall that when we speak of mental images being rotated, scanned, and the like, 'rotate' and 'scan' must not be understood in the senses of the terms in which we speak of sticks being rotated and maps being scanned (both lines of reply discussed in the Leibniz's Law section agreed on this). This reply is not entirely satisfactory, however, for a version of the paraphernalia objection applies equally well to descriptionalist cognitive science as to pictorialist cognitive science, and the line of reply just mentioned does not extend to descriptionalism. The descriptionalist postulates internal sentences that are transformed in various ways. How can the sentences be transformed without an internal eye to read them and internal scissors and glue to change the parts? Here it would not do to say that internal sentences are not transformed in the same sense of 'transform' as external sentences.

I shall argue that a deeper reply to the paraphernalia objection protects pictorialism and descriptionalism equally well. The reply depends on two basic presuppositions of contemporary cognitive science which I will sketch below. Of course I don't expect you to accept these presuppositions on faith. Rather, my point is that if the reply works for descriptionalist cognitive science, it also works for pictorialist cognitive science, and further, it is needed for both, so the objection is irrelevant to the pictorialism/descriptionalism controversy. I also have an ulterior motive, for in explaining the cognitive science reply to the paraphernalia objection, I intend to uncover a deeper problem for cognitive science.[17]

The first presupposition of cognitive science is that our thoughts are coded in a system (or perhaps a number of systems) of structures (neurologically instantiated in us, but instantiable in different ways in different beings), at least some of which have syntax and semantics roughly akin to those of familiar natural and artificial languages. Computers use on-off states of binary elements to do this, but no one has any idea how it is done in the brain. *However* the brain does it, there is a finite alphabet of elementary structures. Complex structures—"words" and "sentences" are built up out of elementary structures. The word and sentence structures have meaning, and the meanings of sentences are determined by the meanings of words plus syntax. Cognitive science is committed to these descriptional internal representations, but it need not deny that there are nonlinguistic, pictorial representations as well. The pictorial/descriptional controversy is, of course, a controversy *within* cognitive science.

The second presupposition is that there are (neurally instantiated) processors that operate on the symbolic structures on the basis of the physical properties of the structures. Thus, though the structures have meanings, the processors can take account of these meanings only to the extent that these meanings are reflected in their physical properties.[18] It is important to note that while both

[17]When I speak of cognitive science, I have in mind an ideology—stated below—presupposed by a significant line of work in cognitive psychology and artificial intelligence, and to some extent linguistics and philosophy. Many—and perhaps most—cognitive psychologists do not subscribe explicitly to such a point of view, though most do work that implicitly involves it. Many writers on this topic prefer to use 'cognitivism' to refer to the ideology I have in mind, treating 'cognitive science' as a neutral disciplinary term. I see 'cognitive science' as an ideologically loaded term (as is 'sociobiology'), since the research programs which parade under the banner of cognitive science would not form a field at all without the ideology.

[18]Jerry Fodor speaks of the class of physical properties that processors can attend to as in the category of "shape" or "form." See Fodor's "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology," *Behavioral and Brain Sciences* 3 (1980): pp. 63–72. Of course, the 3-volt/1-volt difference that codes 1 and 0 in some computers isn't natu-

the representations and the processors discussed by cognitive scientists are assumed to be neurological entities, and the processors are assumed to operate on neurological properties, the *descriptive machinery* of cognitive science nowhere takes account of this assumption of its practitioners. The descriptive machinery of cognitive science assumes only that the processors have no direct access to the meanings of representations. Cognitive science presupposes *mechanism,* not *materialism.* The talk of representations and processing in cognitive science is compatible with electronic or hydraulic realization. Indeed, it is compatible with nonphysical realization. If the representations were drawn in ectoplasm and the processors were souls, so long as the souls fitted the right regularities, the explanations of cognitive science would remain the same.

Some processors search lists of representations, others compare representations, checking for matching, and others transform representations in various ways. The processors are sometimes described as black boxes, sometimes more imaginatively, as homunculi. The task of cognitive science is seen as one of modeling human thought by constructing networks of homunculi that examine and change representations, and send messages to one another. Each homunculus is decomposed into further networks of intercommunicating homunculi and their representations until the whole system is grounded in processors whose operation is to be explained neurophysiologically rather than in terms of operations on representations.[19]

---

rally described as a shape or form difference. This is really a terminological matter, but there is a *substantive* dispute about this class of physical properties as well. Fodor seems to hold that the class of physical properties is restricted to monadic properties (as "shape" and "form" suggest), while others suggest that relational properties that contribute to causal role (such as location) may be important too. See the criticism of Fodor by Sylvain Bromberger and me, "States Rights," in the same issue, and Fodor's response, also in that issue. Fodor's mistake, according to me, is missing the possibility that primitive processors can respond to relational properties.

[19]This picture was first sketched as far as I know in Fodor's "The Appeal to Tacit Knowledge in Psychological Explanation," *Journal of Philosophy, LXV,* 20, October 24, 1968. See also, Dan Dennett's "Why the Law of Effect Won't Go Away," *Journal for the Theory of Social Behavior* (1975); reprinted in *Brainstorms* (Montgomery, Vermont: Bradford Books, 1978). Cummins's "Functional Analysis" puts the same idea in a wider context of explanation outside psychology. The relevant part is reprinted in my an-

The ultimate aim is to "discharge" the homunculi, in Dennett's apt phrase. The first step is to give an account of an homunculus in terms of another network of homunculi and *their* representations, if this can be done. But in the end, discharging depends on finding homunculi that cannot be explained in this way, but only in a different way, nomologically, that is, in terms of laws of nature or their consequences.

The bottom level processors, the ones whose operation can only be explained nomologically, can be called *primitive* processors (this is a definition).[20] For example, a current cognitive science theory of language understanding takes the form of a flowchart which contains a box labeled, "Decide if input is a word." This processor

---

thology, *Readings in Philosophy of Psychology,* Vol. 1, op. cit. These articles, however, tend to scant the representationalist presupposition. This presupposition is given more weight in Fodor's *The Language of Thought* (New York: Crowell, 1975); John Haugeland's "The Nature and Plausibility of Cognitivism," *The Behavioral and Brain Sciences* 2, (1978): pp. 215–225; and Zenon Pylyshyn's "Computation and Cognition: Issues in the Foundations of Cognitive Science," *The Behavioral and Brain Sciences* 3, no. 1 (1980): pp. 111–132. Though Haugeland's article attends to the representationalist presupposition, he rules out pictorialism by assuming that the representations must be descriptional; so what he characterizes is a descriptionalist *school* within cognitive science (one that is more representative of AI than of cognitive psychology) and not cognitive science proper.

[20]Terminological question: should *parts* of primitive processors also be counted as primitive? The definition I'm using (primitive processors can be explained only nomologically) would say yes, though one could say instead that primitive processors can be explained only nomologically *and* are not parts of larger systems that can be explained only nomologically. The latter definition captures the idea better, but the former one is simpler, and so I will continue to use it instead.

Approximately this conception of primitive processor appears in Fodor's tacit knowledge paper (op. cit.). Unfortunately, this idea of primitive processor coexists in Fodor's paper with a quite different idea: ". . . an elementary operation is one which the normal nervous system can perform but of which it cannot perform a proper part." But this latter idea is shown to be defective in Thomas Nagel's comment, "The Boundaries of Inner Space," *Journal of Philosophy, LXVI,* 14, July 24, 1969. The main problem is this: consider a process, P, that is primitive by Fodor's first criterion. Biofeedback (and other kinds of intervention) can often allow the subject to perform any of a number of minute parts of P in isolation. But it won't do to regard these minute parts of P as the *real* primitives, since in general when P is performed by itself, it is performed as a whole, just as each minute part is performed as a whole when it is performed by itself.

can send a "yes" to another box (or rather, the processor symbolized by the box) or a "no" to still another box. But determining whether a given form is a word is not a good candidate for a primitive operation. So a deeper level theory is called for in which the "Decide if input is a word" box is replaced by a network of other boxes, with their own representations; that is, the processor that decides whether something is a word is "decomposed" into other processors. According to one current theory, one more elementary processor fetches words from lists of representations ("dictionaries" in this case), one at a time; another checks whether a fetched word matches the target form, etc. The job is done when all the postulated processors can be reduced to primitive processors.[21]

Now that I have sketched the presuppositions of cognitive science's computation-representation explanations, let us return to the paraphernalia objection. Note to begin with how cognitive science's playmate, the computer, manages to manipulate descriptional representations without internal hands or eyes. Commands given to a computer in a "higher level" language are, in effect, translated into the machine's machine language, a language it *just uses* (in a sense to be explicated). Consider a computer in which addition (but not multiplication) is accomplished by a single ma-

[21]A "machine language" for one machine can be a "higher level language" for a second machine, implemented by the second machine's own machine language. From the point of view of the programmer, the two machines may appear to have the very same primitives. Indeed, the real primitives of the second machine may be inaccessible to the programmer. (To take an extreme case, each primitive processor of the first machine may be simulated by a separate digital computer with its own primitives in the second machine.) In computer science, the set of primitives that a machine appears to the user to have are often taken to define a " virtual machine," and one speaks of the virtual primitives and their interrelations as specifying the "functional architecture" of the virtual machine. Two machines can have the same functional architecture, even though their *real* primitives are quite different. These ideas are explained lucidly in Pylyshyn's "Computation and Cognition . . . " op. cit.

I am ignoring the notion of functional architecture (as distinct from real architecture), first, because I doubt that the distinction between functional architecture and real architecture has any application to human mentality. (The point is that it seems unlikely that computer scientists' clever ways of simulating one machine using another reflect anything mother nature had any need to do.) Secondly, even if the distinction between functional and real architecture does apply to people, it is the real architecture (with the corresponding real primitives) that play a role in my argument.

chine language instruction. The answer to "How does it multiply n times m?" is: "It adds n to itself m times." The answer to "How does it add x to y?" however, is given not by an algorithm (or by a description of how an algorithm is computed) as with multiplication, but rather by an electronic or mechanical or hydraulic mechanism that is explainable only nomologically.

Suppose the string of zeroes and ones (this is a way of talking about binary elements—literally on-off switches operated by electromagnets in some old machines) that stands for addition, say '1000', appears in the control register. '1000' in this register simply causes the zeroes and ones in two other registers to interact so as to produce a series of zeroes and ones in yet another register that represents the sum.

The answer to "How does it multiply?" has a computational-representational answer in terms of an algorithm (or its implementation); the answer to "How does it add?" is given via a mechanism; that is why addition, in this machine, is where computational-representational explanation bottoms out. Addition is explainable only nomologically. But the border between computational-representational explanation and nomological explanation cannot always be drawn sharply. The line between cognitive science accounts and neurophysiological accounts, like most disciplinary borders, is often fuzzy. The important point for the idea of a primitive processor is that *eventually* the cognitive science "decomposition" strategy leads to processors that *clearly* can only be explained nomologically, even if, on the way, there are mechanisms that cannot be easily classified as computational-representational or as nomological.

For example, one way of constructing an "adder" is by connecting gates together. An "and" gate is a device with two inputs and one output; the output fires if and only if impulses reach both inputs. One can build the adder out of "and" gates and "or" gates. Now an "and" gate *can* be viewed as a representation manipulating device—one that emits a '1' just in case it sees two '1s'. So one may wish to regard the gates rather than the adder as the primitive processors. Further, even the gates can be constructed out of processors that can be viewed as representation manipulators, for example, one could make them out of people. One rather old-fashioned way of making gates is constructing them from electromagnetic switches ("relays"). Once one has gotten to the level of these switches, however, one has definitely hit bottom. The only way to explain the operation of such a switch is nomological: electrical pulses in the switch's coil produce a magnetic field (in a way explainable only by laws of physics) which attracts a piece of metal at one end of the coil, closing a circuit.[22]

The application to cognitive science is straightforward, and that is no accident, since cognitive science derives from the metaphor of the mind as a computer. We can explain how the computer accomplishes complex operations, such as multiplication, in terms of a decomposition into operations such as addition, operations which cannot be explained in terms of further commands and representations, but whose only explanation is in nomological terms, for example, in terms of electronics. Call these bottom level operations here, as before, the primitive operations.

Similarly, cognitive science postulates complex processors whose

---

[22]Another way of defining the level of primitive processors (though one approprite to functional architecture, not real architecture) would be to regard the "decomposition" process described in the text as having reached the level of primitive processors when in ascribing representations one requires a change of subject matter (cf. Haugeland, op. cit.). In the computer example just mentioned, the add mechanism would be primitive because the gates' representations should be regarded as having a different subject matter, or so I would argue. The key fact is that the algorithm involved in decomposing multiplication into addition (multiply n by m via adding n to itself m times) is independent of notation (e.g., it works as well in decimal as in binary notation). By contrast, if one views the network of "and" gates, "or" gates, etc., which make up the adder as computing an algorithm, the algorithm is one that works only in binary notation. The multiplier's symbols refer to numbers, but if the gates' symbols refer at all, they are better taken as referring to numerals.

The detail of how the network of gates works brings more conviction on this point. One example: the basic idea of one common binary adder is that when two binary digits to be added are the same, the sum digit in their column is a '0', while if they are different, the sum digit is a '1'. The reason is: $1 + 0 = 1$, $0 + 1 = 1$, $0 + 0 = 0$, since 1 and 0 are designated in binary as in decimal, but in binary, $1 + 1 = 10$, since decimal $2 =$ binary 10. The rightmost digit of the result of $1 + 1$ and $0 + 0$ is the same, viz., '0'. So the rightmost column can be added by an "exclusive or" gate that writes '0' when the input digits are the same and '1' when they are different. When the gate's inputs are two '1's, what it writes is not the sum of 1 and 1, but rather the rightmost numeral of a particular notation for expressing this sum, and that is why it is more natural to regard the gate as a numeral cruncher than as a number cruncher.

operation can be explained in terms of a decomposition into interacting primitive processors whose operation in turn can be explained only nomologically, in terms of physiology. The explanatory "buck" in cognitive science stops at the primitive processors, just as the explanatory buck in the multiplication example stopped at the hard-wired circuitry that allowed the command register to affect the other registers in the appropriate way.

Cognitive science can propose to handle pictorial representations mechanistically in the same way that it handles descriptional representations. Both are to be processed ultimately by mechanisms whose only account is nomological. This answers the Paraphernalia Objection, as it applies both to pictorial and descriptional representations. We don't need hands, scissors, flashlights or eyes if we have mechanical processes that do the job.

This being said, it is worth noting that descriptionalism is in a better position viš-á-vis the paraphernalia objection than is pictorialism. Existing computer models do give us a glimmer of how perception and imagery might be modeled descriptionally, but we have something less than a glimmer of what a pictorial story would be like.

The Background Objection

Hubert Dreyfus has argued that cognitive science is doomed to failure because of its assumptions about internal representations.[23] The basic idea (applied to language understanding) has recently been amplified by John Searle. Dreyfus and Searle claim that to understand language, we require a *common sense background* that cannot itself be fully captured by a set of representations. Searle considers the suggestion that the common sense background required to understand language could be spelled out explicitly as a set of representations. He says:

> But this suggestion is unfillable for three reasons. First we would never know when to stop in spelling out the background. Even if, for example, we described the practices of our culture, those practices

[23]H. L. Dreyfus, *What Computers Can't Do* (New York: Harper, 1972). The argument I refer to is summarized on pp. 203–204 of the revised edition.

themselves depend on certain very general facts about what nature is like and what human beings are like, e.g. that grass grows, that it doesn't eat human beings, that grass is softer than steel, that grass growing and cutting goes on at the surface of the earth—and so on indefinitely (though not infinitely). And secondly, for each of our attempts to spell out the "assumptions" we will have to use words in sentences, and those words in sentences determine their own truth conditions only relative to yet other sets of assumptions, which in turn we would have to spell out. And third, it is not at all clear that "assumptions" is even the right word to describe what it is that makes meaning and understanding possible at all, since the expression implies that these assumptions all have propositional content, that they are all representations. But from the fact that any element of the background can be formulated as a representation, it does not follow that prior to that formulation it existed and functioned as a representation. Have we, for example, always believed unconsciously that grass does not eat humans? Well, if I ever saw a strand of grass eating somebody I would certainly be astonished, and that is at least evidence that one of my intentional states is unsatisfied; but it is at best misleading to assimilate that case to the case where, for example, I expect that it is going to keep on raining and find myself surprised to see the sun come out. The conditions which make representations possible need not themselves all be representations, even though each of them is representable or formulatable as a representation.[24]

The Dreyfus-Searle objection could be summed up as follows:

1. Cognitive science assumes that all of our knowledge is coded in representations in our brains.
2. Our common sense background knowledge cannot be so coded.[25]
3. So cognitive science is committed to an assumption that cannot be true.

Premise (2) obviously needs more backing up than the passage

[24]"The Background of Meaning," in J. Searle, F. Kiefer, and M. Bierwisch, eds., *Speech Act Theory and Pragmatics* (Dordrecht: D. Reidel Publishing Co., 1980), p. 228.

[25]In the last of Searle's three arguments (just quoted) for this premise, he appears to be denying only that our common sense background *need* be coded as representations. Of course, proponents of the position that Searle and Dreyfus are arguing against (which is not the cognitive science position, as I shall point out) could say, "We only insist that in fact the common sense background *is* so coded, not that it must be."

quoted gives it. For example, Searle argues that each represented background assumption will determine truth conditions only relative to others which will also have to be spelled out. But there is no *obvious* difficulty here. Nothing Searle says counts against the possibility of spelling out the *whole* background via a set of representations *each* of which has its truth conditions determined by others. Perhaps attempting to fill this gap in the argument, Dreyfus has recently argued in support of (2) that the satisfaction conditions of mental representations depend on *skills* that are not codable as representations.[26]

Tempting as it is to enter the fray on premise (2), I won't since my real target is premise (1). Cognitive science simply does not assume that all of our knowledge is explicitly coded as representations in our brains. Indeed, cognitive science is more nearly committed to denying this claim.[27]

To make this clear (if it isn't already), I want to introduce the notion of syntactic twins. Syntactic twins can be any devices that fit with the cognitive science ideology. They have internal representations and processors, ultimately grounded in primitive processors. Syntactic twins are defined as pairs of such devices that are *molecule for molecule the same, except inside the primitive processors.* The syntactic form of the representations and the network of primitive processors are the same for syntactic twins, but the input-output function of the primitive processors need not be.[28] For example, two

[26]"Dasein's Revenge: Methodological Sopipsism as an Unsuccessful Escape Strategy in Psychology," *The Behavioral and Brain Sciences* 3 (1980): pp. 78–79.

[27]One type of nonrepresentational containment of information has been discussed explicitly in connection with David Marr's theory of vision. As Shimon Ullman points out in "The Rigidity Assumption in the Interpretation of Structure from Motion" (*The Proceeding of the Royal Society of London* B, 203, (1979): pp. 405–426.), computation in the visual system takes transformations to be motions of a rigid body if there is a unique rigid body interpretation. We think of the rigidity assumption as being "in" the system because it is presupposed by what the system does, not because it is explicitly represented in the system. I don't discuss this matter in the text, since my focus here is on another type of implicit representation of information, viz., information that is implicit in primitive processors.

[28]Since the input-output function of primitive processors (though not their internal operation) is within the domain of cognitive science, these syntactic twins should not be thought of as *cognitive science* twins.

identical digital computers with identical programs could be syntactic twins—and they could *remain* syntactic twins with identical arrays of 0s and 1s, and in that syntactic sense, identical representations, even if we changed the input-output functions of the primitive processors of one (or both) of them, and even though they no longer responded in a similar manner to inputs.

Suppose that the enterprise of cognitive science succeeds, and allows us to build intelligent machines—not necessarily digital computers, since they are only *one* type of device that fits the cognitive science model. Suppose we have two molecule for molecule identical intelligent machines, and once again we change the input-output functions of the primitive processors of one of them. Again, after we have fiddled we still have syntactic twins, since the two machines are the same down to the level of primitive processors—including having the same representations, at least syntactically speaking.[29] Furthermore, the device that we did *not* fiddle with is ex-hypothesi intelligent. But its syntactic twin may turn out not to be at all intelligent, or it may have a different "style" of intelligence and possess different information. Since the two machines have the same representations, cognitive science cannot hold that possession of information or intelligence or "styles" of intelligence must be entirely a matter of representations; rather, these matters depend on the primitive processors as well. So the Dreyfus objection sticks cognitive science with an assumption that it ought not to accept.[30]

[29]The primitive processors of two syntactic twins could be so different that given the relativity of *syntactic* category to process (much discussed in the beginning of this paper), we would not want to regard the two as having even syntactically identical representations. For example, the processors of both twins might skip every other symbol (i.e., treat every other symbol as a "space"), but in one twin the symbols in odd-numbered positions might be so ignored, while in the other twin the even-numbered ones might be ignored. The result: 'CDAOTG' would be regarded as 'DOG' by one twin and 'CAT' by the other. Think of the twins in the examples in the text as having sufficiently similar processors to ensure *syntactically* identical representations.

[30]John Searle tells me that he agrees with this objection (and indeed thought of it himself), so he now thinks that the cognitivists can escape the problem posed by the common sense background. However, I think Searle gives in too quickly. As I shall point out below, the background objection *does* pose an obstacle to the program of cognitive science—at least to the extent that this program is tied to the computer metaphor.

Wittgensteinians sometimes suppose that cognitive science must miss the insight that thinking cannot be just a matter of representations and explicitly represented rules for manipulating the representations. They argue that instructions for using rules would require still further instructions for *their* interpretation, and so on. So we would have an infinite regress.[31] The Wittgensteinian way out is to say that there comes a point at which we have enough instructions for using the rule and we just go on in a certain way. Martians might go on differently, but we go on *this* way. However, the Wittgensteinian appeal to the way we go on has a parallel in the picture I just sketched—the cognitive scientist's appeal to how the primitive processors manipulate representations.

Consider an "add 1" primitive processor, such as the mechanism of a car's odometer. Compare it to a "Wittgensteinian add 1" device, a similar mechanical device that adds 1 to numbers from 1 to 100, adds 2 to numbers from 101 to 200, etc. These devices could be placed respectively in syntactic twins, and cause them to go on in response to "Add one!" in different ways.

This difference between the two "add 1" devices (or perhaps the twins in which they are placed) *can* be seen as an informational difference. And where there is an informational difference there is information. Clearer cases of information implicit in primitive processors require an appreciation of the processor's function in a knowledge possessing system. For example, some birds navigate via internal mechanisms that contain pieces of magnetic material—biological compasses. To the extent that one is willing to think of birds as having navigational knowledge, one can locate directional information in the compass + reading mechanism.

One example illustrating how information can be implicit in the primitive machinery (the primitive processors and their interconnections) derives from the famous Lewis Carroll/W. V. Quine point about *modus ponens*. Imagine a logic machine that has an explicit representation of *modus ponens*, viz.:

Given premises of the form:
(1) Such and such is an A,
(2) All As are Bs,
conclude: (3) Such and such is a B.

[31]See, for example, Dreyfus, *What Computers Can't Do*, p. 203.

Still, not all cases of *modus ponens* reasoning in this machine can be controlled by this explicitly represented rule, since the application of the explicitly represented rule *itself* requires *modus ponens* type reasoning. To see this, imagine a sample argument to which the explicit rule is to be applied, say: All men are mortal; Socrates is a man; conclusion: Socrates is mortal. In order for the machine to apply the explicitly represented version of *modus ponens* to the premises of the Socrates argument to draw the conclusion, it would first have to check that the argument *does have* the specified form (a (1)-style step). Then it would have to take account of the fact that whenever the premises are of the right form, the conclusion of the corresponding form can be drawn (a (2)-style step). Then it could draw the conclusion (which is of the form of (3)). Clearly, this reasoning involves *modus ponens* itself. Conclusion: even in a logic system that has *modus ponens* explicitly represented, the primitive machinery must have the effect of implicitly representing *modus ponens*.[32]

The Computer Metaphor

I have been arguing that cognitive science can regard some of the information available to an information processing system as implicit in its primitive machinery. The concession avoids the Wittgensteinian and Background Objections, but it is not without cost

[32]This point is spelled out in more detail in my "Philosophy of Psychology," in P. Asquith and H. Kyburg, eds., *Current Research in Philosophy of Science* (East Lansing, Michigan: Philosophy of Science Association, 1979). (The same passage is reprinted in the introduction to my Readings in the *Philosophy of Psychology,* both volumes, p. 4.) Barry Stroud makes essentially the same point from a Wittgensteinian point of view in "Inference, Belief and Understanding," *Mind* 88, no. 35 (1979): pp. 179–196. Dan Dennett makes a similar point in *Brainstorms,* op. cit. Note that the implicit information can be "embedded" in the interactions among primitive processors, rather than in any single primitive processor. A Turing machine, for example, can execute *modus ponens* reasoning even when its primitive processors are all just "and" neurons plus peripheral transducers. Any computer with information embedded in a primitive processor (e.g., a computer with a primitive "adder") can be simulated by a Turing machine whose primitives are simple "and" neurons. Of course, this does not show that the original machine does not really have adding information in its primitive adder. A machine which lacks property X can often simulate a machine which has property X.

for cognitive science. To the extent that our knowledge is to be explained by appeal to the nature of primitive operations, to that extent the kind of computational explanations at the heart of cognitive science explain less.

Let us return to an example used earlier: psycholinguists have made considerable progress in characterizing how we understand language. One current theory has it that an initial stage of this process determines whether an input is a word. This is held to be done for functor words such as 'and', 'in', 'the', and 'all' via a different mechanism than for content words. The two types of words are held to reside in different lexicons which are accessed differently. But suppose psycholinguists find new evidence showing that the processor that determines whether one's current input is a word is actually primitive. Then "How do we determine whether our current input is a word?" would not have a cognitive science answer (such as the one in terms of accessing the lexicons), but rather it would only be explainable in terms of physiology. That is, cognitive science explanations would "bottom out" sooner than expected. The primitive operations of cognitive science are not themselves explainable by cognitive science, so to the extent that the explanation of psychological phenomena requires appeal to primitive processes—to that extent, the psychological phenomena do not have cognitive science explanations.

According to my definition of "primitive processor," primitive processors have only nomological explanations (and not explanations in terms of manipulation of representations). In one of the many senses of the unfortunate term, 'analog', this comes to the claim that primitive processors are analog devices.[33] The distinc-

[33]I fear that the introduction of 'analog' into this discussion may be counterproductive. Please note that 'analog' in the sense in which I have just introduced it applies to processes, not representations. So it is to be distinguished from (for example) Goodman's sense of the term. Note that in this sense of 'analog', whether a device "creeps" like a UHF television tuner or "clicks" like a VHF one is irrelevant to whether it is analog. Jerry Fodor and I advocated a sense of 'analog' defined in terms of nomologicity in "Cognitivism and the Analog/Digital Distinction," a paper circulated in 1971, but never published. Similar notions of analogicity have been used by S. Palmer ("Fundamental Aspects of Cognitive Representation" in E. Rosch and B. Lloyd, eds., *Cognition and Categorization* (Hillsdale: Erlbaum, 1973) and Zenon Pylyshyn, "Computation and Cognition," op. cit. Fodor

tion between analog in roughly this sense, and digital (in the sense of: explainable in terms of manipulation of representations) has loomed large in the writings of some opponents of cognitive science. Their arguments turn on the implicit claim that cognitive science only *postulates digital processes.*[34] But, as we have seen, *cognitive science cannot get along without appeal to analog processes,* for they are the ones that *ground* the digital processes.

Once we see that cognitive science must postulate *both* analog and digital processes, it becomes natural to ask just *how important the analog processes are.* Proponents of the computer metaphor (Pylyshyn, for example—see footnote 6) often talk as if the burden

---

and Pylyshyn have recently applied this idea ("How Direct is Visual Perception?: Some Reflections on Gibson's 'Ecological Approach'," *Cognition* 9, no. 2 (1981)) in defining the concept of a detector (a detector is a primitive peripheral perceptual mechanism). Fodor and Pylyshyn think of an analog device as one "whose output is lawfully dependent on the character of the input." Such proposals make the mistake of trying to characterize analogicity *directly* in terms of nomologicity.

The trouble is that if one uses a narrow notion of law (say something of the sort one might expect to find in a science textbook), then the primitive devices of some computers (or perhaps the human brain) will fail to be analog. For example, one could design primitive mechanisms for a digital computer whose input-output relations are counterfactual supporting but not covered by textbook-type laws or chains of them. I have an old-fashioned Swingline stapler whose staples will point in or out depending on the position of the anvil (the metal plate upon which the stapes are bent when the top is pressed). The following counterfactual-supporting input-output generalization applies: anvil in—staple points out; anvil out—staple points in. Thus the staplier is a binary device which, together with the appropriate interfaces, could be a "flip-flop" in a rather pitiful digital computer. Of course, the input-output relations of such a device will follow from a collection of textbook-type laws *plus descriptions of the parts of the device* at the appropriate level of detail. But in *this* sense of nomological (i.e., the input-output relation follows from textbook-type laws plus descriptions of the particular mechanisms), the input-output relation of *every* input-output device is nomological.

My strategy in this paper was to define analogicity more obliquely as having *only* a nomological explanation. The defect of this type of definition is vagueness: other types of explanation lurk in the background, but are not explicitly characterized.

[34] I would ascribe this claim to Dreyfus and Haugeland, though neither it nor the specific analog/digital distinction I am using is explicitly endorsed by them. See Dreyfus, *What Computers Can't Do,* op. cit., and Dreyfus and Haugeland, "The Computer as a Mistaken Model of the Mind," in S. C. Brown, ed., *Philosophy of Psychology* (London: Macmillan, 1974).

of proof is on the pictorialist to demonstrate the existence of analog processes in us. But the *existence* of analog processes is presupposed by the computer metaphor itself; the real issue about the adequacy of the computer metaphor is how much of an explanatory role the analog processes play. And here there is neither a burden of proof nor a clear thrust of the evidence.

The computer metaphor has directed attention to models in which everything is accomplished by elaborate combinations of extremely simple primitive operations such as "check for a match," or "move the element in the nth place to the mth place." The explanatory guts of these models are in the lists of representations and the ways the primitive operations are combined. In this respect, cognitive science is very like its predecessor, associationism, which attempted to build intelligence out of a few simple mechanisms of association of ideas. In my view, this is where the computer metaphor may have done real harm. We should consider the possibility that the primitive operations are more complex and have to shoulder more of an explanatory burden.

For example, suppose that cognitive scientists investigate dolphins and find that they are able to take advantage of what by our standards are sophisticated aspects of the hydrodynamics of their environment. Psychologists dominated by the computer metaphor might investigate models which contain explicit representations of the differential equations of hydrodynamics and digital computing machinery for solving these equations. But they might be wrong. The right approach might be to look for neural mechanisms that themselves obey differential equations of the same form as those of hydrodynamics. Models of this sort would cast the dolphin's navigational equipment as analogous to a model airplane in a wind tunnel.

The relevance of the pictorial/descriptional controversy to the viability of the computer metaphor in cognitive science should be becoming visible. The computer metaphor goes naturally with descriptional representations, but it is not at all clear how it can work when the representations are nondescriptional. This is easy to see, but hard to prove. One can imagine computer-type operations such as list searching and movement from one list to another with pictorial representations, but such operations depend on "matching" processes (the way an item on a list is identified is via matching it with a target item). And there is good reason to think that such "match" operations will be useless with nondescriptional representations. The utility of match operations depends on an "alphabet" of symbols. Of course, one can attempt to get "overlap" measures in order to match pictures, but the history of such attempts strongly suggests this is a blind alley. The trouble is that a picture of a horse and a couch can "overlap" more than the couch does with another couch (or even with the same couch pictured from another angle). "Template matching" approaches in artificial intelligence and psychology have failed miserably, even for letters of the alphabet. Success comes only with restriction to specified type fonts.[35]

The hypothetical dolphin example gives a better idea of the sort of processes that go most naturally with nondescriptional representation. The dolphin has a representation of itself and the currents in its environment, and they interact in a way explainable only nomologically.

Note that I am not *denying* that cognitive scientists may well come up with primitive (analog) processors appropriate for pictorial representations. Rather, my point is that such processors would probably carry much more of the load of explaining how the mind works than is envisioned by current proponents of the computer metaphor. For example, suppose an analog device is developed that can do about as well as a person in recognizing handwritten letters of the alphabet or in saying whether two pictures from different angles are pictures of the same object. Such an analog device would be *far* smarter than any digital device anyone knows how to make today. Indeed, it is commonly believed by cognitive scientists that "cracking" this sort of pattern recognition problem is the key to understanding the mind. My point is that this "key" may be an analog one.

One of the reasons that the artificial intelligence community has been hostile towards pictorialism is that computers don't have primitive operations appropriate to pictorial representations, for example, "rotate." Of course, one could couple an analog computer containing a "rotator" to a digital computer. But the real danger

[35]The *locus classicus* for this point is Ulric Neisser's *Cognitive Psychology* (Meredith, New York: Newpsychology, 1967), pp. 61ff. More recent evidence along the same lines is presented in E. Smith and D. Medin, *Concepts and Categories* (Cambridge, Mass.: Harvard University Press, 1981).

for artificial intelligence is that the model might soon become an unimportant digital computer coupled to an important analog computer. Rosch's work supports the idea that in some kinds of thought about, say, birds, people employ a representation of a prototypical bird such as a sparrow. There is some evidence for the claim that these representations of prototypical objects are the same sort of entities as mental images. There is also evidence for the role of imagery in reasoning, for example, in inference based on the transitivity of a term such as 'taller than'.[36] Note that I am not saying that these items of evidence are conclusive, or even compelling, but only that they raise a difficulty for standard cognitive science that must be taken seriously.

Where are we? I have been arguing that cognitive science explanations may "bottom out" in analog processes too soon for the comfort of cognitive scientists. My argument depends on taking two possibilities seriously: first, that mental images are pictorial, and second that cognitive processes involving images or image-like representations (e.g., prototypes) are important in cognition. Even if these possibilities are actualities, I should concede that cognitive science may nonetheless have substantial explanatory power. Just *how much* explanatory power depends on the "depth" of the layer of "digital" processes that employ imagistic primitives. At one extreme, we can imagine that the brain is a single analog processor—in which case cognitive science is a flat-out dead end. At the other extreme, we can imagine that imagistic primitives are involved in a few peripheral realms of cognitive activity, and that connections among different types of imagistic processes are handled by a central descriptional system. In this case, cognitive science is unscathed.

---

[36]For a recent review of Rosch's work, see C. B. Mervis and E. Rosch, "Categorization of Natural Objects," in M. R. Rosenzweig and L. W. Porter, eds., *Annual Review of Psychology*, Vol. 32, 1981. For some evidence on the imagistic nature of prototypes, see Rosch's "Cognitive Representations of Semantic Categories," *Journal of Experimental Psychology: General* 104, no. 3 (1975): pp. 192–233.

For evidence on the use of imagery in reasoning, see J. Huttenlocher, "Constructing Spatial Images: A Strategy in Reasoning," *Psychological Review* 75 (1968): pp. 550–560. A more recent review is: R. J. Sternberg, "Representation and Process in Linear Syllogistic Reasoning," *Journal of Experimental Psychology: General* 109 (1980): pp. 119–159.

Dreyfus and Haugeland have long advocated the view that much of what cognitive science seeks to explain digitally can only be explained by appeal to analog processes. However, they don't offer the type of argument presented here.[37] Haugeland does discuss imagery in his critique of cognitive science "The Nature and Plausibility of Cognitivism," (op. cit.) but he gives it short shrift, supposing that it doesn't matter much to cognitivism if it cannot handle imagery. He says:

> Since any Cognitivist theory must include some mechanism for getting from retinal images to cognitive descriptions of what is seen, I don't see why that same mechanism couldn't also take inputs from some precognitive visual "tape recorder" (perhaps one with "adjustments" for orientation, size and location). Then playbacks from the recorder would have whatever nondiscursive, "image-y" quality perception has, and Cognitivism would be unruffled.[38]

But Haugeland does not take account of the possibility that our thought processes are *importantly* imagistic. To the extent that imagistic representations play a role in thought, if cognitivists follow Haugeland's advice of sticking imagery phenomena into the "precognitive" machine, they will be throwing out the baby with the bathwater.

Epilogue on the 'Meaning' of Picture

Objection: "When you started this paper, I thought you meant by 'picture' what people normally mean by the word. But it soon turned out that in your sense of 'picture', something could be a picture without *looking* like one; indeed, your pictures in the head aren't even perceived at all (except, perhaps, accidentally by the odd brain surgeon, who, as it happens, doesn't even know he is seeing pictures). Then it turned out that a picture that represents

[37]Dreyfus briefly mentions Rosch's claim that categories may be coded in the mind as images in the introduction to the revised edition of his book (op. cit., p. 24). But his purpose there is only to "highlight the fact that there is no empirical evidence at all" for P. H. Winston's supposition that categories are represented in the mind as sets of necessary and sufficient conditions.

[38]P. 271 in Haugeland's anthology, *Mind Design,* (Cambridge, Mass.: Bradford, 1981).

something as striped need not itself be striped (in the sense in which tigers are striped). Isn't your sense of 'picture' rather a peculiar one?"

I have been saying that the way a thing functions in a system of representation is what makes it a pictorial representation. Something which is a picture relative to a Martian system of representation may not be a picture relative to a human system. Certainly the system within which our *ordinary* pictures function involves their being perceived. Perhaps conceptual analysis would reveal that nothing could be a picture unless its function involved being perceived. I take no stand on this, and thus no stand on the issue of whether the sense of 'picture' in which mental images could be pictures in the head is an *extended* sense of the word. My point required only that mental images be *like* pictures in respects that make it plausible that primitive operations on them would not be of the sort suggested by the computer metaphor. In short, the objection may be right, but it doesn't matter to my point.

Authors on both sides of the pictorial/descriptional controversy that I've been talking about often implicitly accept the objector's point. Some think of pictorialism as the view that mental images are "quasi-pictorial" (Kosslyn's term). Others say that mental images are *analog* representations. I've not introduced the term 'analog' as applied to representations until now because 'analog' is no clearer than 'picture', and just raises a different set of problems and confusions.[39]

[39]There are a number of somewhat different quite precise characterizations of 'analog' in the literature. See Nelson Goodman's *Languages of Art* (Indianapolis: Bobbs Merrill, 1968). An improved version of Goodman can be found in John Haugeland's "Analog and Analog," *Philosophical Topics* 12, no. 1 (1981). See also David Lewis, "Analog and Digital," *Nous* 5 (1971). 'Analog' is even vaguer than 'pictorial', being a vague *technical* term with no home in everyday thought. Since there is less for a definition of 'analog' to be responsible to, adequate precise characterizations are easier to come by. Unfortunately, *incompatible* precise characterizations are equally justified. Note that 'analog' in the sense in which it applies to representations is to be distinguished from 'analog' in the sense in which it applies to *processes*. This latter sense is the one mentioned in connection with primitive processes earlier. I shall now indicate how it is that the two kinds of analogicity need not go together.

A digital processor can use analog representations. Consider a wheel whose rim is divided into 1000 segments; one has negligible electrical resistance, the next segment has a resistance of 1 Ohm, the next one 2 Ohms, and so

'Analog' would also have done as well as 'pictorial' for my point against cognitive science, though the concepts of analog and pictorial are clearly different. To see the difference between analog and pictorial, recall the condition on pictorial representation given earlier: if a picture represents something, S, then at least one part of the picture must represent part of S. But values of charge density can represent distance in an analog computer, even though it is wrong to think of part of the charge density as representing part of the distance.[40]

As far as this paper is concerned, there is no problem if a defense of pictorialism must use 'picture' in an extended sense; a problem arises only if the notion of pictorial (or analog) deployed on behalf of pictorialism is totally *obscure.*

I hope I have clarified the doctrine at least a little here, though, as I have conceded, I have no real characterization of 'pictorial' to offer. With respect to obscurity, the descriptionalist is in a better

---

on up to 999 Ohms. These are analog representations of numbers (being represented by values of a fundamental physical magnitude). The wheel is the main part of an adding machine which works as follows. The starting state is one in which an ohmeter's stationary prongs span the zero segment of the wheel. If n is to be added to m, the wheel is turned m segments, then n segments more; to get the sum, the resistance of the segment touching the ohmeter's prongs (the m-plus-nth segment) is measured. The converse point also holds: *an analog processor can use digital representations.* See the article by David Lewis for an example of a hydraulic analog computing device in which numbers are represented by amounts of water. The device can function equally well with B-Bs as the fluid, number of B-Bs doing the representing. This is an analog processor with digital representation. Lewis's discussion and Haugeland's critique (in "Analog and Analog," op. cit.) fail to distinguish analog processing from analog representation. (The italicized points were made in Fodor's and my "Cognitivism and the Analog/Digital Distinction" mentioned in an earlier footnote.)

[40]Holographic film contains unintelligible "swirls" recording the interference patterns produced by laser beams bouncing off a to-be-pictured object. A holographic image of the object is produced by sending a laser beam through the film and splitting the beam with mirrors, reversing the process by which the pattern was produced. The holographic film has the property that small parts of the film generate a degraded version of the same image as the whole film. Parts of the film do not generate parts of the image (in any usual sense of 'part'). One can perhaps conceive of a processor that changes the image in meaningful ways by processing the "swirls" on the holographic film directly. If this is possible, then holographic film representations would be analog, though not pictorial by the "part" criterion.

position than the pictorialist, because he can anchor his idea of 'descriptional' in a mechanical device that genuinely uses descriptional representations in a way we can understand: the digital computer. We have a conception of mechanistic models of imagery using descriptional representations, and no corresponding conception for pictorial representation.

The core of pictorialism is a striking analogy: experiments suggest that the internal representations of imagery are analogous to pictures in their mode of representation. What this claim amounts to is perhaps no more than that the internal representations of imagery are like pictures in *unspecified* respects. Indeed, it would be surprising if the internal representations of imagery are pictorial in *all* respects. Wouldn't it be a remarkable coincidence if we developed external representations that represented in exactly the same way as the ones nature designed for use by our cognitive subsystem?

Though the thesis of pictorialism is vague, it would be a mistake to write it off on that account. For it may nonetheless be importantly true, and the truth in it may challenge the computer metaphor in cognitive science.[41]

Massachusetts Institute of Technology

[41] I am grateful for comments on an earlier draft to Richard Boyd, Sylvain Bromberger, Susan Carey, Hubert Dreyfus, Kathy Emmett, Paul Horwich, Patricia Kitcher, Bob Mathews, Steven Pinker, Georges Rey, Lance Rips, and the editors of *The Philosophical Review*. Earlier drafts were read at UCLA, at a conference sponsored by Bryn Mawr and Temple, at Yale, Northern Illinois, University of Chicago, and Johns Hopkins. The writing of this paper was supported by the Cognitive Science Program of the University of California at Berkeley, and the Cognitive Science Center at MIT; in both cases from funds supplied by the Alfred P. Sloan Foundation.