ELLIOTT SOBER

MENTAL REPRESENTATIONS*

I

The use of mental images as explanatory constructs has swung in and out of favor in the history of psychology. The study of mental images was so central to the enterprise of 19th century German psychology that Kulpe's claim against Titchener that imageless thought occurs counted as a revolutionary theoretical alternative. More recently, psychologists have increasingly rejected behaviorism's proscription of mental images, and so theories in cognitive psychology have been increasingly peppered with talk of mental representations, both pictorial and nonpictorial. An unfortunate feature of this pendulum effect has been a lack of clarity as to the difference between kinds of representations and a lack of sensitivity to the problems involved in showing that the mind uses one kind of representation rather than another for a certain range of cognitive tasks. This family of questions – why representations are needed in psychological explanations, what distinguishes pictorial from linguistic from other kinds of representations, and how one might go about deciding what the representational systems of thought are – forms the focus of this paper.

A principal feature of behaviorism's attack on introspectionist psychology was the claim that talk of mental pictures is irrelevant to the enterprise of psychology and is without explanatory content. Psychologists like Titchener and Kulpe thought of introspection as being *the* distinctive psychological technique for discovering the form and content of mental representations. For behaviorism, the use of introspective techniques was blatantly unscientific because scientific method was supposed to be intersubjective and public. In rejecting the *method* of inquiry, behaviorists took the further step of abjuring the alleged *subject* of investigation; mental images were seen as unscientific because they were investigated by unscientific procedures.

Another objection to mental images has a much older history, dating back to empiricist theories of meaning. Berkeley criticized Locke for thinking that an idea can be simultaneously general and pictorial. What

Synthese **33** (1976) 101–148. *All Rights Reserved*

picture could correspond to the general idea of triangularity, since any picture of a triangle must be a picture of an equilateral, isoceles, or irregular triangle? The essential *specificity* of pictures seems to mitigate against their capturing the contents of general ideas.[1] Berkeley tried to solve this problem by claiming that specific pictures represent general ideas by a process of abstraction – a picture of some specific kind of triangle represents the general notion of triangularity by our ignoring certain of its special features. On this view, pictures survive as mental contents but now work in partnership with something decidedly non-pictorial. We might say that Berkeley's idea was that pictorial representations work in tandem with *rules* governing how they are to be applied. Berkeley's argument has the force of showing that positing a purely pictorial representation system is not *sufficient* to explain our mental life. This raises a kindred question which we will consider later: Is positing pictorial representations *necessary* in explaining psychological phenomena?

It is now fairly commonplace for philosophers to pursue Berkeley's line of questioning and conclude that some terms fail to have corresponding mental images. This point of view was anything but a truism for 19th century German psychology. Titchener, for example, claimed that all thought was imagistic and sought by introspection to describe the pictures associated with words like 'but', 'patriotism', and 'triangularity'. For him, the mental image corresponding to the general idea of triangularity was

> . . . a flashy thing, come and gone from moment to moment; it hints two or three red angles with the red lines deepening into black, seen on a dark green ground. It is not there long enough for me to say whether the angles join to form the complete figure or even whether all three of the necessary angles are given.[2]

With psychology done in this way, it is no wonder that Frege (in *The Foundations of Arithmetic*) formulated the *variability* argument against the view that the meanings of terms are mental images. Since different people communicate in a natural language, they must associate the same, or nearly the same, meanings with the words and phrases of the language. But the mental images people associate with words vary considerably from person to person. Hence, the meanings of terms are not mental images. According to this argument, mental images cannot play a central role in communication because they are not suitably *invariant.* By generalizing this point, one might claim that mental images cannot figure importantly in any psychological process whose workings possess a great deal of intersubjective uniformity.

Behaviorists like Skinner have also argued against the use of mental images as explanatory constructs on the grounds that their use involves the empty positing of a homunculus. If mental pictures occur in thought, there must be a little man in the brain who looks at the pictures and interprets them. But nothing is explained by this fairy tale, since the same problems concerning how a person perceives reoccur in full force with respect to explaining how the homunculus works. Image theorists need to absorb the point of Molière's often-quoted joke: One does not explain why a potion causes sleeping by saying that it has a sleep-inducing property.

Each of these standard arguments against the introduction of mental images into psychological explanations can be met. The argument from introspection fails to notice that mental images may be posited as purely explanatory constructs. Without any appeal to introspection, one can argue that their existence is confirmed by the way in which imagistic theories of cognition succeed in predicting what responses occur in what stimulus conditions. The variability argument depends on the use of introspection since the claim that mental images are idiosyncratic derives its force from such reports. Once the positing of mental representations is viewed as a theoretical procedure, there is no reason to expect this variability to be an essential feature of mental representations. Indeed, the methodological maxim that councils one to infer that like effects have like causes supports the view that cross-subject invariance of behavior is to be explained by certain invariances of mental structure; imagistic theories for invariant psychological processes should posit nonidiosyncratic mental images. Hence, the idiosyncratic images that we *are* aware of are either not central to invariant cognitive processes or subjects idiosyncratically misdescribe them. In either event, the introspective reports that the variability argument makes so much of are shown to have little significance.

The argument that pictures are limited in their expressive power by their specificity is well-taken. But this at best argues for the incompleteness of any account of the mind which uses only pictorial representations; the fact that pictures can't represent everything does not indicate that they play no role at all in psychological processes. And last of all, the homunculus argument attacks a straw man. One can easily describe operations on pictures which do not presuppose the full intelligence and mental capacities of people. For example, dividing a picture into a grid of squares and counting the number of red squares is an operation on pictures that a fairly primitive machine can perform. A psychological theory which posits the existence of pictorial representations and specifies operations on pictures of only this rudimentary kind would be immune from the homunculus objection. Homunculus explanations are empty if the homunculi are endowed with all the capacities attaching to the people they inhabit. But explanations positing *stupid* homunculi need not be empty.[3]

Discounting or sweeping aside the standard objections to mental images in this way makes room for the possibility of mental pictures as explanatory constructs. However, if one takes seriously the notion of information processing models of cognition, one can give an argument for something stronger: the necessity of the use of representational systems as explanatory concepts. This in itself does not guarantee the need for mental pictures, since they are but one kind of mental representation. But the need for representational systems in psychological models does set the stage for the question 'What distinguishes different kinds of representational systems and which ones are involved in cognition?' which we will take up shortly.

Information processing models of cognition must show how inputs are transformed into outputs and must characterize the items so transformed as containing information. For example, stimuli and responses might be the ultimate inputs and outputs, and the intervening stages of processing might be characterized as operating on hypotheses or sentences, since these are bearers of information. It is crucial that we get clear on the notion of information here at stake. Let us begin by contrasting the syntactic Shannon-Weaver idea of information with a semantic idea of information, for example, that proposed by Carnap and Bar-Hillel.[4] Any system which has different states that occur with determinate probabilities can be characterized by the Shannon-Weaver idea of information. The information content of a particular state of the system is just the improbability of that state's occurring. The rough idea here is that the more unexpected a state is, the more information it conveys when it occurs. What is completely expected provides no surprises. Notice that there is nothing specifically cognitive about this idea of information. One could as easily construct information processing models of digestion as information processing models of cognition, if this were the notion of information used.

By contrast, the semantic idea of information concerns the likelihood that certain hypotheses (or more generally, representations) are *true*. Thus, for example, 'Fred wore a hat' contains less information than 'Fred wore a red hat and carried a walking stick'. The latter is less probable than the former; there are more 'possible worlds' in which the latter is false. On this notion of information, the information content of a hypothesis is spelled out via the Popperian idea of falsifiability: the more likely it is that a sentence is false, the higher is its information content. Notice that this notion of information can be applied only to representations, to those items for which it makes sense to talk of truth and likelihood to be true. If information is in some way more central to cognition than to digestion, it becomes natural to expect the idea of information in information processing models of cognition to be a semantic one.

It is customary to illustrate the difference between these two ideas in the following way. Suppose you are receiving signals over a radio. The signals are sentences of some specified language; suppose that there are 1000 of them that can be transmitted. Now with each of these 1000 signals we can associate a probability – the probability that the next signal you receive will be that one. Presumably we could formulate hypotheses about these probabilities by observing the frequency of occurrence of different signals in some sufficiently large and diverse sample class of signals. The more likely it is that the next signal you receive will be such-and-such, the lower the syntactic Shannon–Weaver information contained in that signal. Notice that on this notion of information there is nothing 'intrinsic' to a sentence which determines its information content – all that matters is its probability of transmission. It may turn out that you receive the sentence 'It will rain or snow tomorrow' far less frequently than 'It will snow tomorrow and the stock market will go up five points'. On the Shannon–Weaver notion, the former signal contains more information than the latter, although according to the semantic notion of information, the former hypothesis would presumably contain less information than the latter, since it tells us less about the world, and is presumably more likely to be true. As Shannon and Weaver have emphasized, the syntactic notion of information has nothing to do with the *meaningfulness* of signals, but merely concerns the probabilities of different states of a system.

It might be objected that the argument succeeds in distinguishing the concepts of syntactic and semantic information only because the signals considered are not considered with respect to their truth and falsity. If all the signals are *occasion* sentences (i.e., sentences which can be true on one occasion of use and false on another) and if the sentences are only transmitted when they are true, then the relative frequency of occurrence of a signal will be the same as the relative frequency with which the sentences are true. In this situation, syntactic and semantic information appear to be interdefinable. Notice that this argument has no force when *standing* sentences (i.e., sentences which if true are true forever) are considered. The frequency of transmission of standing sentences does not help us define the probability of their truth. Even if true standing sentences are sent in every signal transmission and false ones are sent in none, syntactic information still fails to define semantic, for, in this case, the syntactic information of all true standing sentences is 0 and that of all false ones is 1; the semantic information provided by contingent standing sentences presumably falls somewhere in between.

This alleged link between the syntactic and semantic informations of occasion sentences can be defeated in a very general way, in that we can show that for any language containing infinitely many occasion sentences, no signalling device can be such as to render syntactic and semantic information interdefinable. In the situation considered above, the signalling device transmits *only* true occasion sentences. But if frequency of transmission is to correspond to probability of truth, then every transmission must contain *all* the occasion sentences that are true at the time. That is, imagine a series of signals, each consisting of one or more occasion sentences. For relative frequency of transmission to equal relative frequency of truth, each signal must be *complete*. But no signalling device can transmit an infinite number of distinct sentences on a single occasion of transmission. Thus, the signals must fail to be complete in the required sense, and so the syntactic information of signals will fail to define their semantic information.

The two kinds of information correspond to two kinds of probability, which might be called physical and epistemic.[5] Where the objects considered are not representations, there is no temptation to confuse the physical probability of a system's assuming a certain state with the epistemic probability that that state is true, since this latter concept makes no sense at all. But where the objects considered are simultaneously physical objects (like inscriptions or signals) *and* representations, both concepts apply and one is attracted to the possibility of collapsing them. I take it that the above argument provides an important sense in which this cannot be done.

In so far as psychology is concerned with the question *What* information is transmitted and processed in cognition, the concept of information must be semantic. This question is distinct from that of *How much* information is transmitted and processed; this latter question can be posed using either concept. The interest of recent psychologists[6] in describing *what* the frog's (or cat's) eye tells the frog's (or cat's) brain is an interest in describing the *significance* of various neural signals. In calling a given retinal area an edge detector, for example, one thereby specifies what information is transmitted when the retinal area is stimulated. To put this point more generally, within a judgemental theory of visual perception, one tries to characterize different perceptions as hypotheses about the world perceived, and one tries to show how these judgements are inferred from more primitive hypotheses which in turn ultimately derive from causal contact between physical impingements and sensory mechanisms. One is concerned not merely to show how likely it is that a person has a certain perceptual experience, but to show how the judgemental *content* of that experience is predicted and explained by the environment's effect on the information processing mechanisms of perception.

Hence, the items postulated by an information processing model of cognition must be bearers of *semantic* information. From this, can we conclude that the positing of representational systems is required? Representations have syntactic structure; in a representational *system*, representations are generated from a finite set of basic elements by formation rules. For us to answer the above question in the affirmative, we must show that the bearers of information that are justifiably postulated by cognitive theories have syntax and are generative.

Propositions are supposed to be entities that lack syntax: Any two different sentences which are logically equivalent ('contain the same information') are supposed to express the same proposition. It is traditional to say: propositions are not representations; rather, they are what representations express. Yet, propositions are supposed to contain information and have meaning. Why can't an information processing model posit them as the basic items of cognition and bypass talk of representations altogether? Several arguments can be given that show the implausibility of taking propositions as the objects to belief. If propositions were the objects of belief, then it would be impossible to believe one but not the other of two logically equivalent sentences. Mathematical ignorance and mathematical learning would be impossible. In a similar vein, there would be no difference between having two beliefs that are patently inconsistent and having two beliefs that are not at all obviously inconsistent. Rationality should not require logical omniscience; yet this is precisely what follows from taking propositions as the objects of belief. In order to avoid such untoward consequences, information processing models of cognition should posit bearers of information whose principle of individuation is more fine-grained than that of propositions.

These reflections can be supplemented by a host of reaction time experiments which provide evidence for characterizing the *form* in which beliefs are stored. One such experiment will be discussed in Section IV. In general, such experiments work by postulating a form for the beliefs and a set of operations on them as having psychological reality. Certain problems are posed for the subject and the time needed to yield answers is computed. If the conjecture concerning the form and operations is true, then reaction time should correspond to the number of computations: the more computations needed according to the hypothesis to produce an answer, the longer the reaction time should be. By suitably varying the stimulus problem questions, one can differentially confirm pairs of competing hypotheses, each of which makes claims about the form and transformations of beliefs occurring in thought.

One can imagine a continuum from propositions at one extreme to sentences of a natural language at the other. The items at the propositional end have no syntactic properties; the items at the sentential ends have no end of such properties. Intermediate between the two are the equivalence classes defined by Carnap's notion of intensional isomorphism.[7] If intensional isomorphism is used to define the notion of same belief, some pairs of logically equivalent sentences will express the same belief while some others won't. Other positions along the continuum will be defined by how much syntactic commonality there must be between two sentences for them to express the same belief. I have argued above that the weakest requirement (that they merely be logically equivalent) is too weak; propositions are not the objects of belief. How far along the continuum in the other direction should we go? Presumably, the other extreme is no good either: If we let distinct sentences always represent distinct beliefs, we would have to count it as intelligible to say that a person believes '7 + 5 = 12' but does not believe '5 + 7= 12'. A principle of parsimony demands that we endow the objects of belief with no more syntactic texture than is required for purposes of explanation. This constraint serves as a magnet drawing us towards the propositional end of the continuum. Notice that it is an empirical question just how fine-grained the objects of thought are; we can think of different possible species as having objects of belief with different principles of individuation. Whether belief in a given sentence is automatically accompanied by belief in this or that logical consequence of the sentence will turn on the amount of logical acumen that is, as it were, part of the involuntary inferential equipment of the representational system of thought. Experimental results can provide evidence for giving rather precise characterizations of the form that beliefs take in our species. A mere inspection of two logically equivalent sentences accompanied by the feeling that there could not be any difference between believing one and believing the other is no reason at all for thinking that the representational system of thought is not sufficiently fine-grained for there to be a difference between believing one and believing the other.

Granting that the bearers of information in human cognition have syntax, why must they be items in representational *systems*? What would it be for them to fail to form a system? For this to be the case, there must be an infinite number of representations that cannot be described as constructable from a finite list of basic elements via a finite number of transformations. It has been argued that languages which fail to be generative in this way cannot be learned.[8] Learnability is said to require generativity because without generativity there would be no finite characterization of the language that would delimit all and only the sentences of the language, and what learning occurs in mortal flesh is limited to the acquisition of what can be finitely encoded. But why must the set of all representations used in thought be learnable? Maybe the brain, so to speak, never learns them, but has them from the very start. Why insist that the representational system of thought be generative?

The set of representations must be generative because the brain is essentially finite. Since there are finitely many different kinds of signals that sensory receptors can send to the brain and finitely many different physical transformations that the brain can apply to these impulses, generativity is guaranteed. From this finite set of inputs and transformations, all possible representations are constructable. This argument rests on the claim that the physical characteristics of a system place an *upper bound* on the complexity of the system's psychology. A similar conjecture seems equally plausible: The psychological characteristics of a system place a *lower bound* on the complexity of its physical structure. Thus, for example, no system with the psychology of a human being could have the physical homogeneity of a bowl of jello. One of the most pressing problems facing a functional account of psychology is to explain how physical characteristics of a system relate to psychological ones. Perhaps the two principles just cited regarding complexity preservation can help to fill in the details.[9]

This completes our argument that information processing models of cognition presuppose the use of representational systems: such models must view cognition as consisting of the formation and transformation of representations, where these are understood to have syntactic properties and to be part of a generative system. Assuming that there is such a representational system of thought, what are its properties? Although our previous arguments have in the main attempted to articulate presuppositions that psychologists have left tacit, there has been a great deal of discussion of this latter question by psychologists. Recent discussions have frequently claimed that there are at least two representational systems of thought: one of them linguistic, the other imagistic or pictorial. At times, psychologists have distinguished the two systems in terms of their developmental relationships: Bruner and Piaget, for example, see the linguistic system as emerging later than the imagistic one and as being qualitatively different from it. Paivio, on the other hand, has argued that the two systems differ in degree and not in kind, and that they are not distinguished developmentally. Paivio also argues that the two systems differ in terms of the properties dynamic/static, parallel/serial, and concrete/abstract.[10]

As mentioned before, the proliferation of claims about the representational system of thought has been plagued by a lack of clarity as to two questions: What distinguishes pictorial from linguistic systems? How is one to decide whether cognition involves the use of the one, the other, both, or neither? A merely intuitive answer to the former question is as theoretically empty as is a decision about the latter based on introspection. When you are asked how many windows there are in your house, you might think of yourself as answering this question as follows: 'I conjure up images of successive rooms, as if I were walking from one room to the other glancing at the walls, and by this process I glance around the picture of each room and count the windows'. What are we to make of such a claim? Does it provide evidence for the use of pictorial representations in thought? I suggest that it is possible that the *form* of representations used in cognition is opaque to introspection, even though the contents of particular cognitions may be somewhat transparent. Could this mental process, so handily described in terms of images, be described as occurring within a wholly linguistic representational system? If so, how are we to decide which system, if either, is used?

In the next two sections, I will try to characterize the difference between pictorial and linguistic representational systems by attempting to *reduce* pictures to sentences. That is, I will argue that every representational picture has a sentential counterpart of a certain form which captures the content of the picture with which it is identified. By giving linguistic characterizations of crucial pictorial properties, operations, and relations, I will argue that pictorial representational systems can be viewed as a special kind of impoverished linguistic representational system. However, in Section IV, I will argue that this reductive strategy ultimately fails; some pictorial systems are *analog* and these cannot be completely reduced to a *digital* linguistic system. Yet the specification of approximate reductions remains an open possibility and a useful way of understanding nonlinguistic representational systems. The consequences for psychological theory of this view of the relationship between pictorial and linguistic systems are then explored.

II

A representational picture gives it to be understood that a certain state of affairs obtains. If a picture is *true*, then the state of affairs *does* obtain. For this reason, we will identify each representational picture with an existential hypothesis which posits the existence of certain specified objects. A picture will thus be true if and only if the world contains the kinds of objects demanded by the sentential counterpart. True pictures are thus ones that have *verifying instances.*

To facilitate our discussion of the relation between pictures and their sentential counterparts, we will talk about a function $I(\,)$, which we will call the *method of interpretation.* This function maps pictures into sentences. In a rough and preliminary way, we might say that $I(p)$ is the representational content or meaning of the picture p; $I(p)$ specifies the information that p provides. The constraint described in the previous paragraph might be characterized as saying that for every representational picture p, p and $I(p)$ must be *equivalent.*

A unicorn-picture (roughly, a picture that gives it to be understood that there is a unicorn . . .) and a centaur-picture are both false on our above definition, since each has a sentential counterpart which claims that there is a certain kind of entity, and, as is well known, there are neither unicorns nor centaurs. Yet, we do not want the linguistic surrogate of a unicorn-picture to be the sentence 'There is a centaur . . .'. Nevertheless, the two pictures seem to have the same verifying instances – namely, nothing at all. A way around this problem is to demand that the domain of individuals that can serve as verifying instances not be limited to the actual. That is, if $I(p)$ is the sentential analogue of a picture p, then in all *possible* situations, p pictures a scene just in case that scene is a verifying instance of $I(p)$. 'There is a centaur . . .' fails to be the linguistic counterpart of a unicorn-picture because there is a possible situation (e.g., one in which a unicorn romps across a meadow), such that the picture represents the scene, but the sentence fails to have the scene as a verifying instance.

Theories of representation differ in the ways they propose to understand the function $I(\,)$. I will briefly sketch two alternatives; my own solution will fall in between them. On a *causal*[11] theory of representation, a picture will have a sentence 'There is an x such that . . .' as its linguistic counterpart only if the item uniquely picked out by the sentence is causally related in the appropriate way to the representation. As with all causal theories, it is notoriously difficult to specify what the 'right' kind of causal connection is; whatever the nature of the causal relation between

Napoleon and a picture that makes the picture a picture *of* Napoleon, it is not that Napoleon placed a pistol to the artist's temple and exclaimed 'Paint a Corsican seascape!' According to the causal theory, a picture may represent Napoleon and yet not represent another person who looks precisely the same as Napoleon. Because of this, the sentential counterpart of the picture would not be 'There is a man wearing a three-cornered hat with his right hand thrust inside his shirt . . .', since this sentence has both Napoleon and his look-alike as verifying instances. Rather, the sentence would presumably look like this: 'There is an x where x is Napoleon and . . .'. Notice that a proper name occurs in the latter sentence, although not in the former. This connection between a causal theory of representation and the presence of proper names in the linguistic counterparts of pictures will be discussed later.

A different theory of representation is the *projection theory*. On this view, a picture represents any object that would project onto the picture (by some standard technique of projection – say, rectilinear point projection) when suitably oriented. Any object or objects that project (a) in Figure 1 would have to count as a verifying instance of its sentential counterpart. Even the disconnected array of sticks pictured from a different angle in (b) of Figure 1 would have to count as a verifying instance of (a).[12]

How would the projection theory specify the linguistic analogue of the Napoleon-picture discussed above? The sentence would not be 'There is an x such that x is Napoleon and . . .' nor would it be 'There is a human being with a three-cornered hat and . . .', since arrays of objects can project the picture in question and yet be neither Napoleon nor human. Disconnected bolts of cloth and pieces of plastic can project onto the picture here considered in much the same way as the scene pictured in (b) can project onto (a) of Figure 1. For this reason, presumably the projection theory would have to require that the linguistic surrogate of a picture p have the form 'There is an array of objects that projects p.'

The causal theory is as narrow as the projection theory is wide. On the causal view, one but not the other of two visual identicals can count as a verifying instance of a picture, whereas on the projection theory, the most bizarre and unexpected arrays of objects can be verifying instances of what we naively think of as a picture of a human being. The view that I propose falls in between these two, in that it says that Napoleon and his

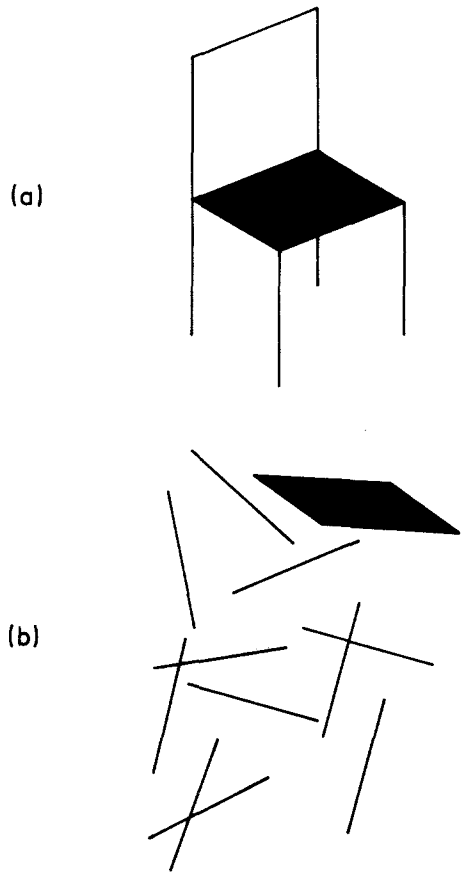

Fig. 1.

look-alike are verifying instances of the picture described above, although not just any hodge-podge of things is a verifying instance of the picture. I will call this view an *inferential* theory of representation. Roughly, it says that the linguistic counterpart of a picture is the hypothesis that a person would infer from the picture (as being its representational content) by using his *pictorial competence* alone.

A person may infer all sorts of things by looking at a picture: 'The artist must have eaten cheese for breakfast', 'It must be time to take the bus', or 'Napoleon looks uneasy' all may come to mind as one looks at a Napoleon picture. Yet, I suggest that it is both harmless and fruitful to introduce a notion of pictorial competence that is strongly analogous to the more familiar notion of linguistic competence. Although a person's ability to understand English sentences may not always be completely separable from his knowledge of the physical world, it is often useful to distinguish some kinds of knowledge from others. As a first approximation, those skills that are constitutive of the ability to speak English are part of what all speakers of English have in common. Similarly, given some antecendently fixed group of people who share the use of a representational system, pictorial competence will be part of that community's shared knowledge. Notice that this idea provides a necessary condition for a given belief's being part of symbolic competence: it must be shared by all those who use the same symbol system. This necessary condition is no more problematic than the idea of people using the same communication system (that idea is already of sufficient theoretical complexity!). But describing a sufficient condition for symbolic competence presents further difficulties. Here one has to show how some but not all of the beliefs shared within a community of symbol users are constitutive of competence in the symbol system. Although I don't think that providing such a sufficient condition is impossible, we need not attempt to do so here. Fortunately, we can refute the causal theory of the function $I(\,)$ without appeal to it.

Pictorial competence alone does not enable someone to look at a picture of Napoleon and decide that the picture gives it to be understood that there is someone who is Napoleon. On the other hand, pictorial competence does allow someone to know more than the mere fact that there is an array of objects that projects onto the picture. The causal theory overestimates the power of pictorial competence; the projection theory errs in the opposite direction. Saying exactly what pictorial competence does allow one to 'read off' from a picture is a major theoretical problem. Perhaps the value of the function $I(\,)$ in this case is roughly this: 'There is a man wearing a three-cornered hat . . .' where the dots are to be filled in with a further specification of the man's stature, appearance, clothing, and surroundings. Clearly, pictorial competence will depend on the ability to use visual cues to formulate perceptual judgements; this connection between representation and preception will be elaborated later on. The main point here is that the linguistic counterparts of pictures will posit the existence of *kinds* of objects; the values of the function *I*( ) will never contain proper names.

One way to test this hypothesis is to ask a variety of people who one has antecendently decided to have pictorial competence (within some modern Western system of representation, say) to provide a proper name in answer to the question 'Who or what does this picture represent?' for a sufficiently diverse range of pictures. If providing proper names does not figure in pictorial competence, then the subjects' successes and failures should be nonuniform, a function of their particular backgrounds and not a part of their commonly held competence. If this is the case for most proper names (as it surely is: think of all the photographs that you can't put a proper name to), then the relatively few cases in which almost everyone can provide a proper name (if such there be) are more smoothly handled by thinking of *all* proper names as failing to fall within the linguistic counterparts of pictures. This strategy makes for a simpler theory than would excluding all proper names except 'Jesus Christ' and a few others from the sentences that the functions *I*( ) can have as values. The ability to provide proper names in the above situation is no more constitutive of pictorial competence than the ability to give proper names in answer to the question 'What is described by the following general terms . . .?' is constitutive of linguistic competence.

Although I have located the inferential theory of representation between two other alternatives, I have not explained what is inferential about it. To do this I will discuss a series of experiments which sought to discover which properties of a two-dimensional picture determine the kind of three-dimensional object the picture is taken to represent.[13] Although each of the hexagons in Figure 2 is a projection of a cube, subjects tend to see only 2 and 4 as three dimensional: 1 and 3 are usually seen as flat hexagons and not as pictures of cubes at all. Hochberg *et al.*

(1) (2) (3) (4)

Fig. 2.

found that a three-dimensional interpretation is more likely than a two-dimensional one if more information is needed to specify the three dimensional interpretation. In the above figure, the three-dimensional interpretation 'There is a cube . . .' is more minimal in information than 'There is a hexagon of such and such a kind' for items 2 and 4, but the reverse is true for item 1 and 3. Of course, spelling out the notion of information involved and deciding how the various possible pictorial interpretations are to be formulated (i.e., which properties of the figures and objects are to be explicitly described in the interpretations) are far from trivial problems. An important factor in their solution seems to be that the greater the number of angles and segments and the greater the number of different sized angles and segments in a configuration, the more information is required to specify it.

This minimum principle has obvious affinities with the idea of simplicity that seems to be a constraint in hypothesis choice: it seems to be a special case of Ockham's principle of parsimony which bids us posit no more than is needed to explain the data. The hypothesis 'There is a cube . . .' is more parsimonious than those competitors which posit more angles and line segments and more different sized angles and line segments. The other constraint that Hochberg *et al.* saw as governing the choice of interpretation is that the posited object should *project* onto the drawing considered. This requirement is analogous to the criterion of hypothesis choice which says that any hypothesis must fit the evidence. Thus, pictorial interpretation obeys constraints of simplicity and evidential fit, just as hypothesis choice does in general. It is for this reason that I have called the theory of representation advocated here an *inferential* one.

The inferential view of representation just sketched also serves nicely as an account of perception, and its generalizability in this direction is another point in its favor. We can conceive of perceptions as judgements. Every perception has a propositional counterpart which is again an existential hypothesis. The true perceptions are the veridical ones – those whose existential hypotheses have verifying instances. Hallucinations have false existential claims as their sentential surrogates. If a person sees a tomato, then there must be *some* existential hypothesis generated by the information processing system. The content need not be 'There is a tomato . . .'; indeed, the sentence need not even have any tomato as a verifying instance. But without some such conceptualization, there would be no perception at all. A difference between the linguistic treatment of pictures and that of perception is that the linguistic counterparts of perceptions contain indexical elements. They will read 'There is a basketball on the ground *in front of me*' and not merely 'There is a basketball on the ground'. This indexical element in perception is replaced by a relativity to point of view that is to be found in all pictures. This will be discussed in the next section.

Previously, we have discussed the problem of reducing pictures to sentences. Just now we touched on the problem of reducing perceptions to sentences. What is the connection between these two questions? The former concerns the relation between two representational systems; the latter concerns the possibility of a certain kind of information processing model in psychology. If one can conceive of perceptions as having a two-dimensional input, the problems seem to be intimately connected. A two-dimensional mosaic of colors is processed in perception and gives rise to a hypothesis about the three-dimensional world of physical objects. As we know now, the mosaic of colors is a fiction if it is conceived of as a stable pattern on the retina; a stabilized pattern tends to degenerate or disappear altogether.[14] However, it may be useful to think of the grid of colors as having reality at some slightly later stage of processing, in which case the two problems of explaining picture interpretation and visual perception may come close together.

The values of the function $I(\,)$ thus specify the information content of pictorial representations and of perceptions; I have argued that the problem of explaining the nature of this function is best dealt with in an inferential theory. But the task of explaining the relational notion '*x* represents *y*' and '*s* perceives *y*' are distinct from that of characterizing $I(\,)$, and it is just here that causal accounts find their proper place. A given picture represents Napoleon in virtue of a causal relation of the appropriate sort holding between it and Napoleon (it is a symptom of the complexities that must lurk within this causal relation that fictitious objects can be represented). A given person *sees* Napoleon in virtue of a similar causal connection. But representations may be inaccurate just as perceptions may be unveridical; even though Napoleon may be the object represented or perceived, the representation or perception may contain misinformation about him. Where a representation is inaccurate or a perception is imperfectly veridical, the representation or perception p may have $I(p)$ as its content and yet the object represented or perceived may not be a verifying instance of $I(p)$. But where representation is accurate or perception veridical, the predicate obtained from $I(p)$ by dropping the existential quantifiers will be true of the object(s) represented or perceived. Thus causal and inferential theories complement each other, the former specifying the *objects* represented or perceived, the latter characterizing the *information content* of representations or perceptions.

In registering our objections to the causal and projection accounts of the function $I(\ )$, we have restricted out attention to pictures drawn from standard Western systems of representations. However, it is essential for the purposes of our argument that our conclusion in favor of some sort of inferential theory not be parochial to a single representational system. Our eventual goal is to formulate a view of what pictorial representation *is* and then to use this account to answer questions about the nature of representational systems that are psychologically real. The affinities noted earlier between representation and perception may perhaps support this more general conclusion: That perceptual judgements characteristically consist in the existential positing of physical objects is no quirk of culture. That pictorial representation should assign interpretations of the same canonical form to pictures also seems to me to be a highly invariant feature of the use of pictures.

In the next section, we will discuss how certain physical properties of, operations on, and relations between pictures are semantically significant. Here again, familiar systems of representation are treated as examples. Yet, it should not be thought that all human pictorial systems attribute just the same significance to the physical properties, operations, and relations that are accorded importance in our system. Our goal is generality, but not of this kind. Rather, the physical properties, operations, and relations that are important in our representational system are *accidental* to the nature of representation; other physical properties, operations and relations *could serve the same functions*. What is crucial is that a representational system contain *some* way of representing certain semantical features. This is a familiar point in language: In English, a conjunction is constructed by writing 'and' between two English sentences and appropriately altering punctuation. But using these three letters and placing the conjunction operator between the two conjuncts are inessential. A language may achieve the desired result in endlessly diverse ways. The crucial point is that a natural language contain *some* device for representing conjunction.

Similarly, it should be noted that in applying our arguments about representational systems to psychology, we need not suppose that the brain contains little canvases and newspaper columns, nor that there are physical operations that the brain performs that duplicate what we can do to the objects in a museum. Rather, the representational systems of thought are instantiated by various physical structures; the representational systems we use in our public discourse are instantiated by others. If there is a language of thought, we can expect it to have a syntax and a semantics; we cannot automatically expect it to be instantiated by accoustical wave packets or by physical inscriptions drawn on a contrasting background. Endowing representational systems with psychological reality requires neither cerebal sepia, neural notebooks, nor a little man in the brain to write and read.

There are too many pictures for us to go through them one by one and detail their linguistic counterparts. Arguably, there are infinitely many pictures: Take two different representations and place them side by side. The picture thus formed will be distinct from the two that went into its construction. This procedure gives rise to infinitely many pictures. So it is not just *difficult* to specify linguistic surrogates piecemeal; it is *impossible.* Consequently, we will give a *normal form* for the sentential counterparts of all representational pictures. A crucial feature of pictorial representation is that every linguistic surrogate has a certain logical form. In this regard, pictorial systems are quite distinct from linguistic ones. Sentences can have various quantifier structures, but with pictures the different possibilities are much more narrowly constrained. It will turn out that similar restrictions attach to maps, musical notation, and graphs. In the following section we also will describe some of the ways in which the interpretations of pictures are constructed out of the interpretations of their subparts. This task corresponds to the linguistic problem of devising a *generative semantics.*

III

In natural languages, certain syntactic operations on sentences achieve certain semantic results; for example, as noted before, placing the word 'and' between two English sentences has the result of effecting a conjunction: the interpretation of the conjunction is a determinate function of the interpretations of the conjuncts. Is there some operation one can perform on pairs of pictures which has the same result? That is, is there an operation, denoted by '+', on pictures p and q such that

$$I(p+q) = I(p) \;\&\; I(q)?$$

If there were, + would be a pictorial conjunction operator. Similar definitions can be provided for the pictorial analogues of the other sentential operators. In a similar vein, we can define pictorial logical truth and contradiction in either of the following ways:

> p is a logically true picture iff $I(p)$ is logically equivalent to '$s \vee -s$'.
> p is a logically contradictory picture iff $I(p)$ is logically equivalent to '$s \;\&\; -s$'.
> p is is a logically true picture iff $I(p)$ is implied by every sentence.
> p is a logically contradictory picture iff $I(p)$ implies every sentence.

Pictorial implication and ambiguity are defined similarly:

> a picture p logically implies a picture q iff $I(p)$ implies $I(q)$.
> a picture p is ambiguous iff $I(p) = s$ and $I(p) = t$ and s and t are not logically equivalent.

Given these definitions, we now can ask what operations on pictures might correspond to or approximate the various sentential operations. This will throw some light on the expressive power of pictures in general.

Concatenating pictures comes very close to forming a picture whose interpretation is the conjunction of the interpretations of the pictures concatenated. But strictly speaking, concatenation fails to exactly exemplify pictorial conjunction. The closest that concatenating pictures p and q can come to conjunction is to yield a picture whose interpretation is: $I(p)$ and then immediately to the right $I(q)$. Notice that placing p to the right of q yields a different picture, and of course a different value for the function $I(\,)$, than placing p to the left of q. Concatenating pictures usually implies a particular *spatial* relation between the scenes pictured. However, in pictorial forms as diverse as Medieval paintings of the stages of the cross and newspaper cartoons, concatenation has temporal, but not spatial force. As one moves from left to right, one imposes 'and then' between one's interpretations of the member pictures to form the interpretation of the whole picture.

For concatenation to yield conjunctions, the interpretations of the concatenated pictures must be fairly independent of context. That is, it must be the case that the interpretation given *p* in isolation does not radically change when *p* is placed along side of *q*. This is true for many cases, but not for all. At times a picture will have a determinate interpretation when viewed in isolation, but a wholly different one when it forms a subpart of a larger picture. It is this phenomenon that makes for the possibility of *clever cropping,* which is achieved when a picture is cut into subparts in such a way that the separated subparts have radically different interpretations from the ones they had before. Thus, not every representational picture that is a subpart of a given picture is a *subpicture* of the one divided. This has a sentential analogue. The form of words 'you fall' is a sentence of English, and occurs within the larger sentence 'To you fall the rewards'. However, the semantic interpretation of the larger sentence does not turn on the semantic interpretation of 'you fall'; the sequence of words is a subpart, though not a subsentence, of the larger sentence in which it occurs.

That concatenation is a conjunction operation receives further confirmation when one reflects on the relationship between a representational picture and one of its subpictures. The interpretation of the larger picture will logically imply the interpretation of the subpicture; a conjunction logically implies each of its conjuncts. Just as a pictorial analogue of conjunction is concatenation, so a pictorial analogue of implications is containment. Moreover, I conjecture that these two operations on pictures are the *only* general counterparts of the two sentential operations within our representational system.

The impossibility of pictorial logical truth and contradiction is indicated by several considerations. On the one hand, it is a mark of the logical truths and falsehoods that no nonlogical terms essentially occur in them. 'Napoleon' does not essentially occur in 'Napoleon won at Waterloo or it is not the case that Napoleon won at Waterloo' because the sentence is logically equivalent to sentences in which 'Napoleon' fails to occur. The absence of nonlogical terms in such sentences is one sign that they cannot be the linguistic counterparts of any representational picture, for such pictures always have the force of giving it to be understood that objects of such-and-such a kind *exist*, where the properties of the posited object(s) go beyond those that every object must logically have (like self-identity), but fall short of those that no object *can* logically have (like nonself-identity). A further argument for there being no pictorial representations corresponding to logical truth and falsity derives from the remarks above about logical implication. If one starts with a representational picture and cuts it down so as to extract a subpicture of the original, one has constructed a new picture whose interpretation is a logical consequence of the interpretation of the picture with which one started. If there is such a thing as a pictorial tautology, and if containment is indeed *the* pictorial analogue of implication, then it should be possible by successive snippings of pictures to yield a subpicture of a subpicture of a subpicture . . . which has a tautology as its sentential analogue, this being the case because tautologies are implies by every sentence. But clearly, such a procedure never yields a pictorial tautology. One might start with a picture of a rabbit, an eggplant, a baseball, and an alchemist, and by three successive croppings yield three subpictures, each implied by the one before. But given a picture of a baseball, further snippings will do one of two things: either they will destroy the representationality of the picture, or the sentential analogue of the picture produced in this way still be an existential claim which asserts of some nonlogical property that it is instantiated (e.g., 'There is something that is white'). If the first alternative were ever true, some pictures would be nonrepresentational; if the second alternative were always the case, then every picture would be representational. Which of these two alternatives is correct will be discussed later. But notice that in either event, successive snippings fail to yield a pictorial tautology.

An argument in the other direction points to the impossibility of pictorial contradiction. Since logical falsehoods imply every sentence, and if, as we have claimed, concatenation is *the* way of forming conjunctions of pictures, then it should be possible by successive concatenations to build up pictures with successively stronger sentential counterparts. One can move from a picture that gives it to be understood that there is a baseball to one that gives it to be understood that there is a baseball and an eggplant, etc. But this procedure, carried out as long as you like, never results in a pictorial contradiction. An even stronger picture is always possible, no matter how bloated the existential commitments of the one just constructed.

Berkeley's critique of Locke, mentioned in Section I, is relevant to this question of whether pictures can have logical truths or falsehoods as their sentential counterparts. Pictures have a persistent specificity that general terms and their logical combinations often lack: according to Berkeley, there is no picture of a triangle that fails to be a picture of an equilateral triangle or an isosceles triangle or an irregular triangle. If there were such a thing as pictorial tautology, such pictures would be the most universal images possible. But the Berkelian point is stronger than that: it isn't merely that the expressive power of pictures is limited to sentences which are neither logical truths nor falsehoods, but that within the class of such sentences, certain *weak* sentences are not the linguistic counterparts of any representational picture. Could there be a picture whose *full* sentential analogue is 'There is a fire engine'? How could a picture be a fire engine picture without there being more content than this to its linguistic counterpart? It isn't that the picture must have parts which are interpreted as representing wheels. Nor is a hook and ladder required. But how would we justify the fire engine interpretation of a picture without pointing to some subparts of the picture and justifying our viewing them as representing subparts of a fire engine? Typically, a representational picture will have subpictures; and the interpretation of the picture is a determinate function of the interpretations of the subpictures.

Like the fire engine picture, every representational picture is *complex*. Every representational picture has representational subpictures; successive snippings never destroy representationality. It follows that every picture is a representation. Since even the most abstract canvas can be embedded in a picture that is straightforwardly representational, even the most abstract canvas can be obtained by snipping away at a picture that is clearly representational. Even a homogeneous white canvas has a sentential counterpart, if only the rather impoverished assertion 'Something is white'. Such pictures are properly treated as representations because they enable us to know about the object pictured, if the picture is accurate. Moreover, the method of producing an abstract picture (e.g., by camera) may be precisely the same method that gives rise to more familiar representations.

Even though every picture is a representation, not every picture is used as a representation. The Mona Lisa has a representational content that is taken to be part of its aesthetic significance. A Jackson Pollock canvas likewise has a representational content, but this is not viewed as aesthetically significant. Our conventions of interpretation might have been precisely the reverse, letting the representational content of the Mona Lisa count for nothing, but holding that the representational content of the Jackson Pollock has aesthetic importance. The difference between so-called 'representational' and 'nonrepresentational' pictures lies not in whether some are assigned interpretations by the function $I(\,)$, but in whether this assignment, which every picture receives, is aesthetically important.

It is an interesting feature of the pictures which we ordinarily view as nonrepresentational that their interpretations come to no more than, roughly, 'Something projects this picture'. A homogeneous white canvas claims no more than that something is white; its interpretation fails to specify which amongst the scenes satisfying the projection requirement is claimed by the picture to exist. In this way, such pictures differ from the examples of Figures 1 and 2. There we saw that the method of interpretation we use selects amongst the possible interpretations satisfying the projection requirement and singles out a kind of object which is the one that the picture posits.

The subpictures of a homogeneous white canvas are just more of the same; the subpictures of a fire engine picture must in some way picture parts of a fire engine. Something like a principle of complexity matching seems to be at work here. That is, the more complex the associated sentence (as viewed by the community of picture users), the more complex must a picture be to have that sentence as its linguistic counterpart. Hence, for us 'Something is white' may have a homogeneous canvas as its associated picture, while 'Something is a fire engine' may not.

To this it may be objected: 'Why can't I simply *stipulate* that a certain homogeneous red canvas represent the Empire State Building? This stipulation is intelligible and can be absorbed into the way we understand and use pictures, but no part by part interpretation is involved.' Several points need to be made in answer: First, it is not clear whether this red canvas is to be treated as a *pictorial* representation or as an item in a language, or in some other representational system. That the item in question happens to be a canvas is no guarantee that it is functioning pictorially. Also, the possibility of such a convention is consistent with our claim: all that is being asserted is that such a convention would count as a fundamental alteration of the laws governing the representational system we use. And last, it is fruitful to consider the analogous arguments that have been made against the kinds of generalizations that transformational grammarians seek to put forward concerning human natural languages. Granted, people are able to absorb such a convention; the point is that the diversity of *actual* human languages happens not to include such a convention, and the nonoccurrence of such a convention is *no accident.* The best explanation for its absence is that a constraint occurs in the laws that govern the kinds of representational systems people use that in some way precludes such a convention.[15]

Berkeley's point about the specificity of pictures is false as formulated above: there *could* be a picture of something triangular which leaves indeterminate the kind of triangularity the thing has. Perhaps the object is pictured as being in the distance, at an angle difficult to determine, and partially obscured by fog. The specificity of pictures does not consist in their being completely determinate; it need not be that every predicate is such that either it or its negation figures in the sentential counterpart. Nor is specificity to be understood as requiring that if the sentential counterpart of a picture includes the predicate *P*, and if falling under *P* implies falling under *Q* or under *R* or under *S*, then *Q*, *R*, or *S* must also figure in the interpretation. Rather, the specificity of pictures consists in this: There are concepts that *cannot stand alone* in the interpretations of pictures. No picture gives it to be understood that something is triangular full stop. In this, 'triangle' resembles 'fire engine' and contrasts with 'white'. But perhaps 'white' cannot stand alone either: a specific shade of white is what is required. The specificity of pictures is thus a kind of interpretational *complexity.*

Above, I argued against the possibility of pictorial contradiction by appealing to the fact that successive concatenations of pictures fail to yield a picture with a contradiction as its sentential counterpart. But this claim would be undermined if there were a pair of pictures whose interpretations were of the form p and $-p$ respectively. For if there were such a pair, they could be concatenated to yield a picture whose interpretation would approximate something of the form $p \& -p$. Of course, finding such a pair of pictures would not guarantee the existence of pictorial contradiction precisely because of the ways in which concatenation fails to *perfectly* exemplify conjunction. But the non-existence of a picture whose interpretation has the form $p \& -p$ is guaranteed by more than just the absence of a perfect conjunction operator. It is also guaranteed by the absence of a pair of pictures of the requisite form: If every picture has the form 'There is an object such that . . .', no picture has the form 'It is not the case that there is an object such that . . .'.

The absence of pictorial negation goes deeper, deriving from the property of persistent specificity noted earlier. If there is a picture with an interpretation of the form 'there is an x such that Fx . . .', then there is no picture whose interpretation has the form 'there is an x such that non-Fx . . .'. Although there may be pictures giving it to be understood that there is a river such that . . . , there are no pictures which give it to be understood (merely) that there is a nonriver such that . . . . The only way the existence of a nonriver can be posited by a picture is by its positing something *specific* (e.g., a basketball) in its stead. To put this idea more generally: If a predicate can figure in the sentential counterpart of a representational picture, then its negation cannot. This, in itself, does not say which of 'river' and 'nonriver' can so figure, but only asserts that both cannot. The question that this distinction gives rise to is this: What properties of a predicate and its negation determine which can enter into the interpretations of representational pictures? This will be taken up later.

It will be helpful to conceive of the construction of sentential counterparts in the following way. A set of predicates is available for the construction. Our interest is in the logical form that the construction will take. So far I have argued that the normal form begins with a series of existential quantifiers, and that conjunction but not negation is possible. For these admissions and restrictions to make sense, we must relativize them to some specification of predicates. This can be seen by considering whether the predicate 'nonriver' is negated or not. The mere presence of the letters en-oh-en means nothing, since we could arbitrarily specify a definitional substitute for 'nonriver' which lacked the tell-tale prefix. So let us conceive of a *P-system*, which is a set of predicates, perhaps including 'river' or some definitional equivalent. Our stricture against negation now has substance: if a predicate is in the *P*-system, its negation is not. We might imagine there being many possible *P*-systems, and then go on to inquire which one or ones human being use. For example 'green' is within the arsenal, but presumably 'not green' is not.

The kind of question considered in Section I as to the identity conditions of the objects of belief – whether they are propositions or sentences or a kind of thing whose criterion of individuation is somewhere in between – can be pursued for the items in the *P*-system. Again, the two extreme cases are probably unsatisfactory – what we have called the 'predicates' of the *P*-system should not be conceived as being syntactically fixed to such a degree that 'rouge' might be plausibly included but not 'red', nor is the criterion of individuation so coarse-grained that any definitional substitute of a predicate occurs in the *P*-system if the predicate does. The fact that 'green' occurs should not guarantee that 'green before the year 2000 or green after the year 2000' also occurs.

Thus, as we have seen, the absence of negation has two parts: On the one hand there is no operation on a picture that produces a picture whose imterpretation is the negation of that of the one operated upon. On the other, predicates occuring in the interpretation of any representational picture are such that their negations never so occur. A similar limitation attaches to the operation of disjunction: There is no operation on pairs of pictures which effects their disjunction, and if two predicates occur in an interpretation, their disjunction never so occurs. These two strictures presumably come to much the same thing because the disjunction of the two existential claims 'There is an x such that Fx' and 'There is an x such that Gx' is equivalent to the single existential sentence 'There is an x such that Fx or Gx'.

I do not have a knockdown argument for the absence of a disjunction operator. But notice first of all that *ambiguous* pictures do not provide counter-examples to my claim. Ambiguous pictures have two interpretations, not a single disjunctive one. Inspection of candidates for a disjunction operator will perhaps increase the plausibility of my negative thesis. Now consider the second part of the claim about disjunction.

Suppose that lemons and limes look precisely the same. Couldn't there be a picture whose sentential counterpart includes a clause reading 'There is an x such that x is a lemon or lime'? To argue that no picture is properly construed in this way, I must show that some nondisjunctive interpretation is always more highly valued by the set of constraints figuring in the function $I(\ )$. Perhaps in this case the designated interpretation would be 'There is a yellow ellipsoid fruit . . .'.

The previous comments on conjunction, negation, and disjunction are formally identical with the line I took in *Simplicity* on the idea of propertyhood. I will make this analogy explicit by first outlining the connection between propertyhood and simplicity judgements on pairs of hypotheses. It is generally agreed that a hypothesis which says that the world will change with respect to a given property is less simple than one which says that the world will not change. Thus 'all emeralds are green until the year 2000 and thereafter blue' is supposed to be less simple than 'all emeralds are green'. As in physics, the ideas of change and no-change make sense only relative to a rest frame. If we take *colors* as our frame of reference, the first hypothesis is a change hypothesis, while the second is a no-change hypothesis.

However, suppose we take *grulers* as our frame of reference. Just as green is a color, grue is a gruler. An enduring physical object is grue if it is green before the year 2000 and blue thereafter; it is bleen if it is blue before and green thereafter. Now relative to this frame of reference, the hypothesis 'all emeralds are green' is a change hypothesis, since it says that emeralds change from grue to bleen at the year 2000. Similarly, 'all emeralds are green before the year 2000 and blue thereafter' counts as a no-change hypothesis relative to grulers as a frame of reference, since this hypothesis says that all emeralds remain grue. The relativity of change and no-change to a rest frame guarantees that a theory of simplicity can mirror the intuition that change hypotheses are less simple than no-change hypotheses only if it contains a parameter for representing a frame of reference. Within the theory I propose, one hypothesis is simpler than another relative to a P-system, where the P-system is a set of predicates that are regarded as specifying *natural properties.* For us, colors are natural properties whereas grulers are not, so we see 'all emeralds are green' as simpler than 'all emeralds are grue'. However, it is possible that a creature should take grulers to be the appropriate frame of reference and have the opposite intuition about the relative simplicity of these two hypotheses.

Thus, if a person thinks one hypothesis is simpler than another, we can use this fact as evidence for making hypotheses about what properties this person regards as natural. By eliciting this and other evidence we can construct a smallest set of predicates which suffices to specify a frame of reference for the totality of a person's simplicity judgements. This *P*-system will be the set of predicates that the person regards as natural. The question then arises: What truth-functions of those natural predicates are themselves natural? I answered this question as follows: If a truth-function of predicates is nomologically equivalent to some conjunction of predicates in the *P*-system (i.e., the equivalence is a law), then the truth-function picks out a natural property. If the truth-function is materially equivalent to some conjunction, then it picks out a natural set. Thus, suppose by the merest accident 'is a prime number or a unicorn' picks out the set of prime numbers. Then the predicate determines a natural set, but not a natural property. On the other hand, suppose that it is no accident that all and only the integers are picked out by the predicate 'is a prime number or a composite number'. Then the predicate picks out a natural property.

By this argument it can be seen why the negations of natural predicates are not natural. What truth function of natural predicates picks out the same class as that picked out by 'nonriver'? No conjunction does this–at best an elaborate disjunction is required. This underlies our suspicion that irreducibly negative and disjunctive predicates are mere artifices of language and fail to pick out sets of things that really have something in common.

If propertyhood abhors negation and disjunction, why should pictorial representation? Clearly, *language* allows us to delimit all sorts of artificial sets and nonnatural properties; why should pictorial representation be any less profligate? Perhaps considerations like the following suggest why: Pictures posit objects by the following process: Pictures are decoded by specifying *properties* that they have, and then these are transformed into an interpretation which gives it to be understood that there are objects having certain *properties*. This claim is not trivial, in that it excludes the picture being specified, for example, as 'There is a *grue* patch of color . . .', and similarly pictorial interpretations are not constructed with such nonnatural predicates either. On this view, the systems that process pictures transform specifications of properties into specifications of properties. But predicates that are irreducibly disjunctive or negative (relative to some suitable *P*-system) do not pick out properties, and so they are not encoded.

An information processing model of visual perception would represent perceptions as judgements constructed within some representational system. Human beings do not have perceptions that are properly characterized by the form of words 'There is something grue in front of me such that . . .'. Visual perception is supposed to allow one to adapt to and adapt one's environment. For this reason, it is supposed to pick out properties of the environment that are significant in terms of prediction and explanation. It would be disfunctional if the perceptual mechanism were clogged with existential posits that focused on nonnatural categories. Thus, it is to be expected that evolutionary advantages accrue to perceptual systems that represent in terms of the natural properties, rather than to ones that lump together unrelated elements by means of artificial categories.

The procedures whereby a linguistic interpretation of a picture is generated are formally quite similar to the procedures whereby perceptual judgements are constructed from sensory stimuli. In both cases two-dimensional inputs are transformed into posits of three-dimensional objects. This formal similarity between the mechanisms of picture interpretation and visual perception, taken together with our earlier claim about the way in which perceptual judgements are focused on properties, argues for a construal of pictorial interpretation as similarly focused on propertyhood. Hence one should expect pictorial interpretations to fail to include irreducible disjunctions and negations, since these fail to pick out properties.[16]

Evolutionary advantages accrue to organisms whose visual representations are constructed out of predicates that in fact pick out properties. But perfect adaption is rare, and so there will be systems whose representational systems fail to exhibit such perfect harmony with the environment. All systems, regardless of how well adapted they are, are to be regarded as constructing perceptual judgements out of their *P*-system, where this *P*-system is the set of predicates that *they* regard (perhaps mistakenly) as natural. Moreover, since all perceptual mechanisms can be viewed as *aiming* at picking up on the relationships between properties in the world, we can view the admissable constructions as always precluding irreducible negation and disjunction, since propertyhood is not preserved under such constructions.

This rather abstract argument for the logical form of the sentential counterparts of representational pictures can be reinforced by a more mundane consideration. In perceiving ambiguous scenes or pictures, the perceptual system 'flips' back and forth between distinct but equally 'good' (in terms of the criteria for choice of interpretations sketched in Section I) interpretations. In the young girl/old hag and rabbit/duck phenomena, for example, one's perception does not have the form of an unchanging disjunctive judgement. Rather, one switches back and forth between two nondisjunctive alternatives. The judgement is not the univocal 'There is a rabbit or duck . . .', but a vacillation between 'There is a rabbit . . .' and 'There is a duck . . .'. This further supports our claim that the perceptual system abhors a disjunction.[17]

Our view of the normal form of the sentential counterpart of representational pictures so far comes to this: initial existential quantifiers followed by conjunctions of predicates drawn from a set of predicates (the *P*-system); no negations or disjunctions allowed. Thus, the interpretation of a landscape might read: 'There is a tree in a meadow such that . . .', where the rest or the sentence describes the tree and meadow. Now imagine a scene which conforms to this description, and also has a dog in it that is plainly visible. Would this scene count as a verifying instance of the picture? It seems to me that it would. The picture posits the existence of various objects, but it is not part of the meaning of the picture that there is a scene which contains *only* the objects posited. Yet, as with other kinds of representations like maps and sentences, there is a convention governing the use of pictures which stipulates that the pictures are *complete* with respect to an implicit standard of fineness of detail. A map whose detail is only fine enough to indicate towns (symbolized by labeled points, say) is viewed as showing *all* the towns there are in the mapped terrain. But the absence of any mark in a certain area of the map does not allow one to infer that there are no houses there, but only that there are no towns. On the other hand, if the map were fine-grained enough to posit individual houses, then the omission of a mark from a given area of the map would allow one to conclude that there are no houses in the corresponding place. Pictures have the same property: imagine an aerial photograph or a painting of an area of land; although there is no pretense that each blade of grass is shown, those objects that are explicitly pictured do define a standard of fineness of grain which allows us to apply the convention of completeness. However, we do not build this into the sentential counterparts of pictures (e.g., by adding a clause which says 'and that is all that is visible from a given point of view') for the same reason that it is not part of the meaning of a sentence that it completely describes the scene it is about. A person who describes, maps, or pictures an area of land may perhaps provide an incomplete description; but it may yet be true of its intended subject and hence the completeness claim does not figure as part of its meaning.

It is a striking fact about representational pictures that each has a *point of view*. Two pictures may posit the same object and yet may differ in their points of view with respect to the posited object. The relationship between two such pictures seems to be relevantly different from that between two identical pictures, which similarly manage to make identical existential claims. In order to account for these facts, we will augment our characterization of the normal form of pictures to include a clause which reads '. . . and there exists a point of view relative to which the objects posited are related thusly . . .'. Two photographs of the same object taken from different angles need not be *logically* equivalent, and so their interpretations, even in positing the same object, should not be logically equivalent. That pictures have points of view explains how this is possible.

Although I argued above that there is no such thing as pictorial contradiction, this should not be taken to mean that there is no such thing as pictorial impossibility. There are more sentences that couldn't be true than just the logical falsehoods, and some of these are picturable. I will call these nonlogical impossibilities 'impossible pictures' for short. Some impossible pictures are constructed by concatenating pictures which are themselves quite possible. Here again, pictures parallel sentences. Thus consider the devil's pitchfork, represented in Figure 3. This figure can be decomposed into the two pictures in Figure 4. Notice that each of the two pictures in Figure 4 can be taken to represent a possible state of affairs. But consider the interpretations that we assign to the areas marked with an '*x*'. In the lower left-hand picture, the area is interpreted as representing empty space; in the upper right-hand picture, it is interpreted as

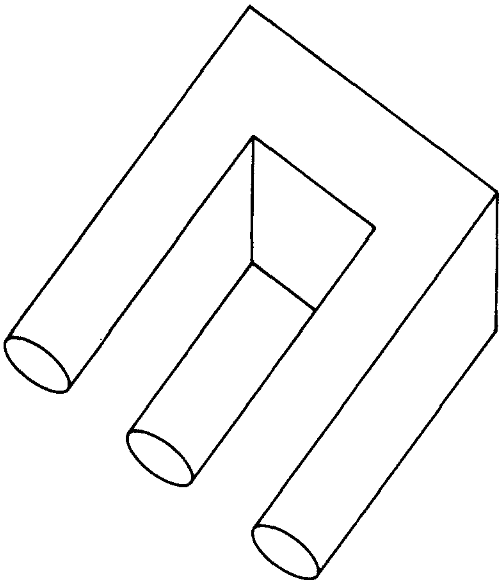

Fig. 3.

picturing part of the surface of an object. When the pictures are concatenated, the interpretations are conjoined and the area with the '*x*' is interpreted as simultaneously representing empty space and the surface of an object. Hence the impossibility of the devil's pitchfork. This way of understanding our interpretation of Figure 3 accords well with the fact that our perceptual judgements derive from quick successive fixations in which the eye darts from detail to detail of the perceived scene. Successive parts are interpreted, and the interpretation of the whole is generated by a quasiconjunction of the interpretations of the parts.[18]

The same explanation can be given of the impossibility of a painting of Magritte's called 'Ready-Made Bouquet'. This is a picture of a street with a house on it. The sky is bright with midday sun, but the house and street are set in midnight darkness. The impossibility (or if not impossibility, unexpectedness) of this picture is accounted for by its interpretation being a conjunction of the interpretations of the top and bottom halves: Under the fused interpretation, the picture gives it to be understood, roughly, that there is a scene in which it is noon and midnight at the same

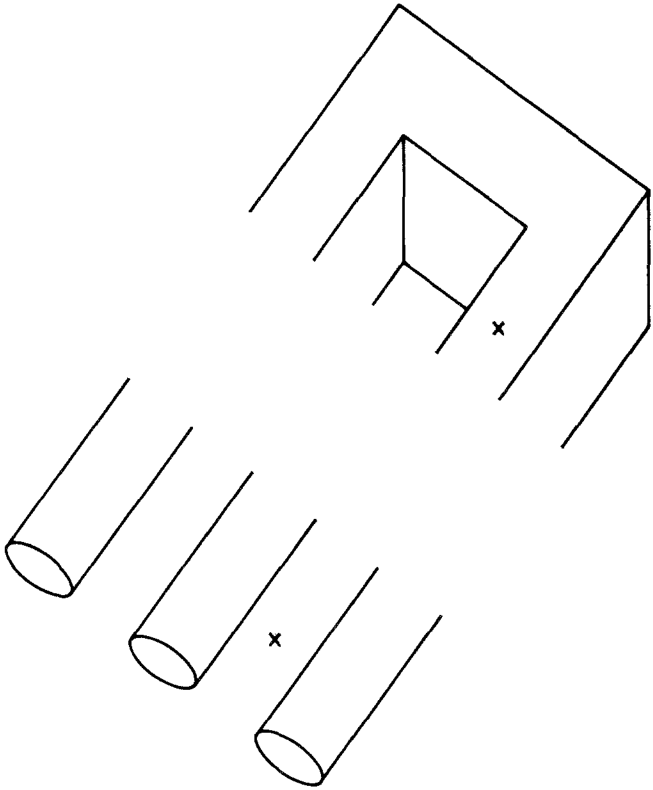

Fig. 4.

time. The impossibility in virtue of perspective of certain pictures by Hogarth and Escher can also be explained by this part-by-part analysis.

Our characterization of the normal form of the interpretations of pictures allows for there being two ways in which two pictures can have all and only the same scenes as verifying instances. As noted earlier, a unicorn- and a centaur-picture both have the same (null) sets as their extensions in the actual world. Another way in which two pictures can so agree is that they be pictures of the same scene from different points of view, and the possibility of this sort of equivalence is insured by a clause in the normal form. In addition to these nonlogical coextensivenesses, it should be mentioned that blowing up a picture is an equivalence transformation. A picture and its blow-up have precisely the same sentential counterparts, since the size of a representation is of no representational significance; the only relevance that size can have is by way of the *relative* size of the different subpictures of a given picture. Hence a picture and its blow-up are necessarily equivalent.

The fact that blowing-up is an equivalence operation allows us to answer a question raised earlier concerning the degree to which the interpretations of concatenated pictures are independent of each other before and after concatenation. I claimed that if concatenation is to be a conjunction operation, the concatenated pictures must have determinate interpretations before concatenation, and the interpretation of the picture formed by concatenation must be a determinate function of the interpretations of the pictures before concatenation. This is not always the case. All blow-ups of a picture are logically equivalent. Now consider two different representational pictures *A* and *B* which we are going to concatenate. Notice that it matters a great deal to the interpretation of the resulting picture whether one concatenates a large version of *A* and a small version of *B*, or vice versa. Yet, choosing a large version of *A* or a small version of *A* does not affect the interpretation that *A* receives in isolation. This shows that concatenation can give representational significance to a property of the concatenated pictures (namely, their relative size) which makes no difference at all to the representational significance of the pictures when taken in isolation.

In our discussion of the differences between the causal and inferential views of representation, I claimed that proper names do not figure in the sentential counterparts of representational pictures. This point can be put more generally by employing Kripke's distinction between rigid and nonrigid designators.[19] A rigid designator is an expression that picks out the same object in all possible worlds in which it exists, whereas for nonrigid designators, the object selected may vary from world to world. An example of the former is 'Benjamin Franklin'; an example of the latter is "the inventor of bifocals". It seems to me that every representational picture will have some nonrigid designators in its sentential counterpart. For example, in a still life with lemon, one part of the existential posit that the picture specifies is that there is an object that is yellow, and 'yellow' is nonrigid. Pictures as a class are nonrigid – they may have one thing as verifying instance in this world but something quite different will serve that purpose in a different possible world (e.g., Edgar Allen Poe may in a different possible world be the verifying instance of a picture that in this world pictures Napoleon). The verifying instances of pictures can vary from world to world, but the referents of rigid designators cannot so vary; hence the sentential counterparts of representational pictures cannot be formed by linking existential quantifiers to rigid designators.

The approach we have pursued with respect to pictorial representation can be applied to any representational system at all. To illustrate, I will consider very briefly what the normal forms are of maps, musical notation, and curves plotted on Cartesian coordinates. In all these cases, it will turn out that concatenation is a conjunction device and inclusion is an implication relation. Perhaps these are universal properties of human representational systems.

Maps constitute an interesting combination of pictorial and linguistic properties. The basic quantificational form is pictorial: a map gives it to be understood that there is an area with certain properties. The linguistic properties of maps derive in an obvious way from the use of place names as labels. We argued above that pictures do not have proper names in their sentential counterparts. But it is clear that maps do have such names in their sentential counterparts, and this comes from the occurrence of place names in the map itself. All this is not to deny that pictures sometimes have names in them – think of a mug shot of Al Capone which includes the label 'Al Capone'. The point is that if the occurence of the name in the picture is treated pictorially, then the picture can have persons other than Capone as verifying instances.

Musical notation may be construed as a description of sounds; this description may be used (among other things) as instructions for producing a performance, or as a standard for checking a given performance. The logical form of musical notation comes close to that of pictures: the score gives it to be understood that there is a set of sounds related to one another in a specified way.

If pictures and musical notation have the same logical form, how do the two systems of representation differ? The answer is twofold: in the details of the predicates used to specify their linguistic counterparts, and in the properties of the representation that have representational significance. Pictures represent visual properties of the world; scores specify *aural* features. Pictures usually grant representational significance to the color of the picture surface (pen and ink drawings constituting an exception), although musical notation rarely does. Representational systems whose values for the function $I(\ )$ have the same form can be expected to differ in these two ways.

In contrast to the above two cases, where the logical form of representations closely parallels the logical form of pictures, consider a curve that describes the relationship between two properties on Cartesian coordinates. The curve itself has the logical form of a universal quantification 'for any object s and number n, if the x-value of s is n, then the y-value of s is $f(n)$.' The further inclusion of *data points* alters this form, however. Suppose the graph considered represents the relationship between the period of a pendulum and its length. The inclusion of data points on this curve indicates that *there exists* a pendulum with such-and-such a period and length. This goes beyond what the curve itself asserts, in that the curve leaves logically open what pendulums (if any) there are: it says only that for any object, *if* it is a pendulum then its length and period are related in such-and-such a way.

IV

So far, I have described how each representational picture can be assigned a linguistic counterpart. Distinct pictures have distinct sentential analogues, and significant operations on pictures, like concatenation and blowing-up are mirrored in counterpart operations on surrogate sentences, like conjunction and equivalence. If every picture and every significant property, operation, and relation of pictorial representations had a linguistic counterpart, we could argue that pictures *reduce* to language.

However, if pictures are to reduce to sentences, with distinct pictures identified with distinct sentences, then the number of distinct pictures must not outstrip the number of distinct sentences. But this is precisely what happens in virtue of the *digital* character of language, if the pictorial system in question happens to be *analog.* There is a countable infinity of sentences in English, but a continuum of distinct representations in the pictorial representational system we are familiar with using. We mentioned earlier that the relative size of two picture parts is representationally significant; within our representational system, the ratio of the lengths of two picture parts can represent the ratio of the lengths of the two things pictured. But how many different such ratios are there? There is a continuum of different ratios. Hence the expressive power of analog pictorial representations extends beyond that of language, which contains at best a countable infinity of distinct representations. Similar comments apply to the representational significance of the *orientation* of two picture parts. Hence, not all pictorial systems are reducible to impoverished linguistic systems of a certain kind. The relation of analog pictorial systems to language is more complicated: With respect to logical operations on representations, linguistic systems are more powerful, but with respect to expressing specifically visual relations between posited objects, linguistic systems can be more impoverished.[20]

It may be objected that there are not continuously many different representations because of the quantized nature of physical objects. There are minimum differences between the lengths or orientations of any physical objects: Objects whose lengths are closer together than that minimum are indeterminate with respect to whether their lengths are equal or unequal. This fact seems to me to be irrelevant to the question of whether a representational system is analog or digital. The rules of pictorial representation imply, in the cases imagined above, that if the length of one picture part is n and the length of another picture part is m, then the lengths of the object pictured by the first is related to the length of the object pictured by the second by a ratio of n/m. It is not part of the rules of any representational system to say that the possible lengths of objects in the real world are limited to rational numbers; this is specified by a physical theory. But given information about the lengths of the two picture parts, the representational system uses this information to construct an interpretation of the picture. In saying that a representational system is analog, we are claiming that there is a property of representations which the system treats as significant, and which is such that if there were a continuum of values of the property, there would be a continuum of significantly different representations. An analog system no more demands that matter is continuous than a digital system demands that matter is quantized.

As is well known, analog systems can be transposed into digital ones by segmenting the continuum into a countable number of intervals. For example, a smooth curve in a plane might be represented digitally by dividing the plane into a number of small squares and then defining the curve by specifying those squares through which the curve passes. Although there are continuously many distinct curves that can be drawn on a page, there are countably (and even finitely) many sets of squares. Here, of course, the reduction to a digital representation involves a loss in precision.

There seems some reason to think that this reductive process occurs in perception. A picture or a physical object is perceived by means of a digital encoding. The packet of light rays reaching the eye causes retinal sensors to be stimulated which are on/off in character. These sensors then relay a signal deeper into the brain and the possible transformations this signal is subject to are also digital. From this process ultimately arises the perceptual judgement of the three-dimensional object.[21]

If this digital characterization of the brain is correct, two questions arise with respect to our claim that some pictorial systems are analog in character. How can there be continuously many distinct pictorial representations if the digital character of the brain forces these representations into a countable infinity of possibilities? How can the representational system of thought be analog if the brain is digital? The former is a question about the way in which properties of an organism determine the properties of the representational systems it can use; the latter is a question about the way in which the physical characteristics of a system affect the properties of its information processing systems. To answer these questions in anything like the detail they deserve must await another occasion. However, for now notice that in claiming that digital creatures like us can use analog representational systems like pictures and thermometers, I am committed to saying that the expressive capacities of a community's representational systems may well outstrip the ability of the community to use the system.[22]

As for the possibility of a pictorial system having psychological reality, I would say that the digital character of the brain as a system of neurons entails that the representational system of thought must ultimately be digital. This constraint on the relation between a physical system and its psychology follows from the requirement that for a physical system to be a realization of a given functional system, the two must be *isomorphic*. If relevant portions of the brain are realizations of various psychological processors, then there must be a mapping under which distinct psychological processes and objects are mapped into distinct items specified by a perspicuous representation of the structure of the brain. If two physical states are identical, then whatever mental state is associated with the one must also be associated with the other. Thus, the isomorphism requirement implies our earlier comments about complexity preservation: The physical complexity of a system puts an upper bound on the complexity of its psychology precisely because the system's psychology and the physical parts of the system which instantiate the psychology must be isomorphic.

It should be noted that not all pictorial systems are analog; black and white newspaper photographs are both digital and pictorial, for example. The photographic method used gives representational significance to the relative greyness of different areas of the picture. There are *finitely* many degrees of grey that the system employs. Since the printing process is not perfectly precise, two areas of a photograph may slightly differ in shading, and yet this difference has no representational significance. Given the variability of paper and ink, there presumably *is* a continuum of colorations that may appear in a photograph (we ignore here questions concerning the quantum character of light). Even so, this continuum of possibilities does not give rise to a continuum of semantically distinct representations. This is why the system is digital.

Pictorial representational systems may be analog or digital. Where they are digital, they simply *are* linguistic systems of a certain kind; where they are analog, the reduction to language can only be approximate. What unites analog and digital pictorial systems is their logical form, their specificity, and their containing a point of view. I take it that this cluster of concepts helps fix the nature of pictorial representation.

This view of the relation between pictorial and linguistic representational systems has consequences for the kinds of arguments that have been made for the psychological reality of nonlinguistic representational systems. Such arguments usually proceed by claiming that there are certain perceptual or cognitive abilities which can be handily described as involving pictures and their transformations. This is then taken as evidence for saying that the representational system of thought that is involved is pictorial *rather than linguistic*. Such proposals usually admit that a linguistic system is justifiably posited to explain other capacities.

If the only property distinguishing a pictorial system from a kind of impoverished linguistic system is its analog character, and if the representational system of thought cannot be analog, then it would seem that the claim for the psychological reality of both pictorial and linguistic representational systems must be equivalent to the claim that there are two representational systems of thought, one of them containing the expressive capacities of a full-blown quantificational language, the other being a language whose logical properties are precisely those characteristic of pictorial systems. If pictorial systems are reducible to linguistic systems, the claim that the mind contains both thus commits one to the view that the mind is slightly *redundant*. It might seem that such a proposal is methodologically unsound – why attribute two systems to the mind, one reducible to the other, instead of positing one all-encompassing (linguistic) system? This object will be discussed shortly.

An example of the kind of proposal just described is to be found in an interesting experiment of Metzler and Shepard's.[23] Subjects were shown pairs of line drawings like those shown in Figure 5, and were asked whether the two drawings were different views of the same object. Metzler and Shepard discovered that the time it takes between display and answer was proportional to the number of degrees of rotation through which one of the figures must be turned to bring it into superimposition with the other. That is, after subtracting for latency times, it takes twice as long to answer if the rotation is 50 degrees as it takes if the rotation is 25 degrees. Shepard claimed that he rotates mental

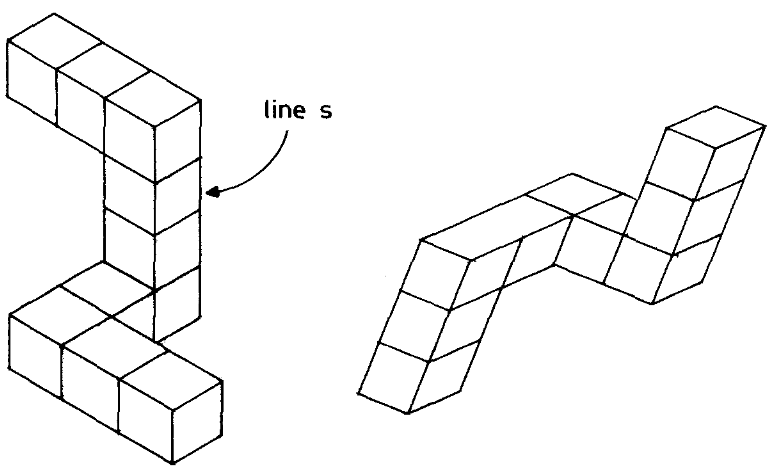

Fig. 5.

images at an angular velocity of 62.6 degrees per second. The form of the argument thus seems to be that the *best explanation* of certain abilities is that the representational system of thought is pictorial or imagistic in character.

However, this conclusion in favor of pictorial representations faces the problem that an alternative explanation in terms of sentential representations seems equally plausible. Consider the sentential counterparts of the two pictures in Figure 5. Each asserts the existence of an object of a certain geometrical kind; the sentences differ only as to their specification of the spatial orientation of the object. The spatial orientation of posited objects might be characterized by using a three-dimensional coordinate system. The left-hand picture might have in its linguistic counterpart a clause that reads: '... and the object is oriented so that line *s* is parallel to the *y*-axis ...'. Shepard claims that the mind works by rotating the left-hand picture into the right-hand one. We can characterize this rotation as involving a series of pictures, each a slight rotation through some very small distance (say one degree) of the one before. The sentential analogues of the pictures in this series will be identical save for the clause that specifies the spatial orientation of the posited object. That is, the linguistic counterparts will differ only in how they fill in the blank in the clause 'line *s* forms a —— degree angle with the *y*-axis'.[24]

Given these sentential analogues, the pictorial operation of rotation might have the linguistic counterpart of increasing or decreasing the numeral in the above clause that specifies the number of degrees of angle. The reaction time differences that are a function of the number of degrees of rotation could thus be equally well explained by claiming that the computation time on sentences of running through the numerals between '0' and '50' is twice that of running through the numerals between '0' and '25'.

Metzler and Shepard (pp. 150–51) call the pictorial system they posit 'analog', but by this they do not mean to claim that there are continuously many distinct representations operated on in the rotation phenomenon. Rather, they hold that the analog character of the representational system consists in the fact that the mental operation of comparing the two pictures in Figure 5 involves moving through a series of representations, each member of the series representing a small rotation of one of the stimulus pictures. Although this seems to me not to be what is meant by 'analog', I would concede (as stated above) that the pictorial character of the representational system of thought would not be ruled out if the representations were found to be discontinuously related to each other. However, notice that given what they meant by 'analog', purely linguistic systems can fill the bill; our own alternative explanation is one such.

The dilemma that our understanding of pictures poses for psychological postulation of pictorial representations is precisely this: I have claimed that pictorial systems can differ from linguistic systems only with respect to a property that no psychologically real representational system can have. If this is true, then arguments for pictorial systems of thought cannot succeed in showing that a psychological process is pictorial *rather than linguistic.* Pictures are reducible to languages when the domain considered is limited to systems that can have psychological reality.

As mentioned earlier, the reducibility of pictures to language does not guarantee that there is but one representational system of thought. It may be that the mind is partially redundant. One of the striking features of the processes that occur within a single processing system is the redundancy of operations; processes are duplicated so as to minimize breakdowns and malfunctions.[25] Perhaps the same sort of redundancy occurs between representational systems as occurs within a single processor. To go beyond mere speculation as to the possibility of such cross-system redundancy, one would first have to isolate a psychologically significant process which can be characterized as occurring purely within a pictorial system. Then one would have to show that this pictorial system is psychologically distinct from a more powerful linguistic system of thought. This latter task would perhaps be facilitated by cerebral localization results. This extra step of cerebral localization is required only if the reducibility claim is correct. For if pictures are not reducible to language within the restricted domain of systems that can have psychological reality (and if reduction in the other direction is likewise impossible), then the fact that a process can be explained by positing a pictorial system counts in favor of thinking of a distinctly pictorial system as having psychological reality.

Considerations of parsimony dictate that it is better to posit one representational system of thought rather than two, whether the two are irreducibly distinct or redundant. In a similar vein, it is best to minimize the content and details of the representational systems postulated; one attributes only enough structure to them to explain the totality of behavior. This second injunction is a frequent casuality in imagistic theories of cognition. Let us go back to the problem discussed earlier: By what procedures do you recall how many windows there are in your house? An imagistic account would claim that the processing system went through a series of pictorial representations, extracting information. Now let us consider these individual pictures. For them to figure in this procedure in the way claimed, presumably the pictures of the rooms contain pictures of the windows as subparts. How many panes of glass are there in each window? What kind of light is coming through the windows? What, if anything, is visible through the windows? As argued before, pictures are essentially complex: they always contain subpictures, although there is no particular subpicture that they must contain. The fact that question after question concerning the details of the mental representations fails to elict an answer somewhat undermines the claim that they are pictorial. The absence of detail suggests that the representations are linguistic, since sentences are just the sorts of thing that can be suitably unspecific.

This kind of objection to the positing of pictorial representations might be countered by claiming that the questions about details have answers although they are opaque to introspection. That is, one might claim that the window-pictures are such as to specify the number of panes and quality of light, only the subject does not have access to these features. This response is not ludicrous or impossible. Yet it does seem *ad hoc*, if such additional features never show up in behavior but only serve to keep the imagistic hypothesis afloat.

Information processing models of thought must posit representational systems in which cognitions are formed and transformed. The in principle objections to mental representations can be met, and the utility of such constructs is to be estimated by seeing what kinds of theories are fruitful. It is an empirical question whether the positing of pictorial systems will find a place in psychological theory, but considering the functional characteristics of pictorial systems and the digital nature of the brain, this possibility is best understood in terms of an impoverished linguistic system's being psychologically real. Undoubtedly, the nature of pictorial representation is far from exhausted by our comments about logical form, specificity, and point of view. A deeper understanding of varieties of representational systems will have important consequences for the philosophy of psychology and, of course, may invalidate our conclusions about the representational systems of thought. The theory of representation is even more immature than the science of psychology; still, I hope that the above discussion throws light on some important features of representation and cognition, and on the connection between them that an information processing model must forge.

University of Wisconsin, Madison

NOTES

* I wish to thank Daniel Dennett, Fred Dretske, Dennis Stampe and Robert Howell for their helpful suggestions.

[1] Berkeley was undoubtedly on to something, but the content of this notion of specificity is difficult to specify. I try to make sense of it in Section III.

[2] Quoted in R. Brown, *Words and Things*, New York, Free Press, 1958, p. 91.

[3] This point is made forcefully in F. Attneave, 'In Defence of Homunculi', in W. Rosenblith (ed.), *Sensory Communication*, Cambridge, MIT Press, 1960, pp. 777–782, and in Daniel Dennett, 'International Systems', *Journal of Philosophy* **68** (1971), pp. 87–106.

[4] See, for example, W. Weaver's 'Recent Contributions to the Mathematical Theory of Communication' and C. Shannon's 'The Mathematical Theory of Communication' in Shannon and Weaver, *The Mathematical Theory of Communication*, Urbana, University of Illinois Press, 1972, and Y. Bar-Hillel and R. Carnap's 'Outline of a Theory of Semantic Information' in Y. Bar-Hillel (ed.), *Language and Information*, Addison-Wesley, Reading, MA., 1964, pp. 221–275. I develop a nonprobabilistic concept of information in my *Simplicity*, Clarendon Press, 1975.

[5] See R. Carnap's 'Statistical and Inductive Probability' in B. Brody (ed.), *Readings in the Philosophy of Science*, Prentice-Hall, Englewood Cliffs, 1970, pp. 440–450.

[6] See, for example, D. H. Hubel and T. N. Wiesel, 'Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex', *J. Physiol.* **160** (1962), and J. Y. Lettvin, H. R. Maturana, W. S. McCulloch, and W. H. Pitts, 'What the Frog's Eye tells the Frog's Brain', *Proc. Inst. Radio. Engrs. N.Y.* **47** (1959), pp. 1940–1953.

[7] R. Carnap, *Meaning and Necessity*, University of Chicago Press, Chicago, 1956.

[8] D. Davidson, 'Theories of Meaning and Learnable Languages' in Y. Bar-Hillel, *Logic, Methodology, and Philosophy of Science*, Amsterdam, North Holland, 1965, pp. 383–394.

[9] See J. Fodor's 'Explanations in Psychology' in M. Black (ed.), *Philosophy in America*, Cornell University Press, Ithaca, 1965, pp. 161–179 and H. Putnam's 'Minds and Machines' in A. Anderson (ed.), *Minds and Machines*, Prentice-Hall, Englewood Cliffs, 1964, pp. 72–97 for this view of psychological theories as functional. In Section IV, I argue that the principles of complexity preservation just cited are consequences of an *isomorphism* requirement.

[10] J. Bruner, 'The Course of Cognitive Growth', *American Psychologist* **19** (1967), 1–15; J. Bruner, R. R. Oliver and P. M. Greenfield, *Studies in Cognitive Growth,* New York Wiley, 1966; J. Piaget and B. Inhelder, 'L'image mentale chex l'enfant', Paris Presses Universitaires de France, 1966; A. Paivio, *Imagery and Verbal Processes,* New York, Holt, Rinehart, and Winston, 1971.

[11] D. Stampe sets out a causal theory of representation in his 'Show and Tell', in B. Freed (ed.), *Forms of Representation,* North-Holland, in press. Although I argue that a causal theory does not help specify the values of the function $I(\ )$, I agree with Stampe that it does have considerable relevance to the problem of explaining the relation 'x represents y'. This will be discussed later.

[12] The two pictures in Figure 1 are copies of photographs of the famous Ames chair that were made from different points of view.

[13] J. Hochberg and E. McAlister, 'A Quantitative Approach to Figural "Goodness"', *Journal of Experimental Psychology* **46** (1953), pp. 361–364 and J. Hochberg and V. Brooks, 'The Psychophysics of Form: Reversible Perspective Drawings of Spatial Objects', *American Journal of Psychology* **73** (1960), 337–354. These and related investigations in psychology are discussed in Chapter 4 of my *Simplicity,* Clarendon Press, 1975.

[14] L. Riggs, F. Ratliff, J. Cornsweet, and T. Cornsweet, 'The Disappearance of Steadily Fixated Visual Objects', *Journal of the Optical Society of America* **43** (1953), 495–501; R. Pritchard, W. Heron, and D. Hebb, 'Visual Perception Approached by the Method of Stabilized Images', *Canadian Journal of Psychology* **14** (1960), 67–77.

[15] The laws governing language formation cannot absolutely preclude such conventions, since it is patent that such conventions can be learned and used. Instead, one must say that the laws governing language formation impose a 'preference' ordering on different languages, and such conventions are so far down the list that they are never selected in actuality. Although it is no accident that such conventions never appear, the laws do not absolutely rule them out. See my 'Computability and Cognition', forthcoming, for details.

[16] S. Barker and P. Achinstein try to solve N. Goodman's new riddle of induction (see his *Fact, Fiction and Forecast,* Bobbs–Merrill, Indianapolis, 1965 for a statement of the problem; our earlier example of 'grue' is adapted from this work.) in their 'On the New Riddle of Induction', *Phil. Rev.* **69** (1960), pp. 511–522 by arguing that 'All emeralds are grue' is not confirmable because 'grue' is not picturable. My treatment above forges a similar connection between propertyhood, representationality, and inductive confirmability. Consistent with this connection is the possibility that one may regard 'grue' as a natural predicate which finds its place in the sentential counterparts of pictorial representations and which occurs in confirmable generalizations. However, *for us* this possibility is nonactual.

[17] In *A Study of Thinking,* John Wiley, New York, 1956, J. Bruner, J. Goodnow, and G. Austin argue that disjunctions are equally problematic for that part of the cognitive system involved in concept attainment.

[18] J. Hochberg advocates something like this part-by-part view in his 'The Representation of Things and People' in E. H. Gombrich, J. Hochberg, and M. Black, *Art, Perception and Reality,* Baltimore, Johns Hopkins University Press, 1972, pp. 47–94.

[19] S. Kripke, 'Naming and Necessity', in D. Davidson and G. Harman (eds.) *Semantics for Natural Languages,* D. Reidel, Dordrecht, Holland, 1969.

[20] N. Goodman has argued that pictorial representational systems differ from linguistic ones in that the latter but not the former have notational systems associated with them. See his *Languages of Art,* Indianapolis, Bobbs–Merrill, 1968 for details on this issue as well as for a discussion of the analog/digital distinction.

[21] W. McCulloh and W. Pitts argue for a digital model of neural activity in their 'A Logical Calculus of the Ideas Immanent in Nervous Activity', in W. McCulloch, *Embodiments of Mind,* Cambridge, Mass., MIT Press, 1965.
[22] See my 'Computability and Cognition', forthcoming.
[23] R. Shepard and J. Metzler, 'Mental Rotation of Three-Dimensional Objects', *Science* **171** (1971), pp. 701–703; J. Metzler and R. Shepard, 'Transformational Studies of the Internal Representation of Three-Dimensional Objects', in R. Solso (ed.), *Theories in Cognitive Psychology,* Wiley, 1974; pp. 147–202.
[24] We assume only for the sake of convenience that the rotations occur in the x–y plane and not in three dimensions.
[25] M. Arbib stresses this feature of mental processes in his *The Metaphorical Brain,* John Wiley, New York, 1972.