Arguments Concerning Representations for Mental Imagery

John R. Anderson
Yale University

The argument for a propositional over a pictorial representation for visual imagery has largely taken the form of an attack on the logical coherence of pictorial representations. These attacks have not been valid; one can develop a coherent dual-code model involving pictorial and verbal (nonpropositional) representations. On the other hand, empirical demonstrations that are claimed to support pictorial representations fail to discriminate such representations from propositional ones. It is argued that the failure of the antipictorial and the propictorial arguments stems from a fundamental indeterminacy in deciding issues of representations. It is shown that wide classes of different representations, and in particular propositional versus dual-code models, can be made to yield identical behavioral predictions. Criteria such as parsimony and efficiency in addition to prediction of behavior may yield further constraints on representation; and, in particular, it may be possible to establish whether there are two codes, one for visual information and one for verbal, or whether there is a single abstract code. However, the conclusion of this article is that barring decisive physiological data, it will not be possible to establish whether an internal representation is pictorial or propositional.

There is no need to document the historical importance of imagery to philosophical and psychological discussion nor to explain the important place it has in current cognitive psychology. There is also no need to review the long debate over the status of the concept of imagery. Ample reviews of these matters exist in other places (e.g., Anderson & Bower, 1973; Bower, 1972; Bugelski, 1970; Fodor, 1975; Kosslyn & Pomerantz, 1977; Paivio, 1971, 1976; Pylyshyn, 1973; Richardson, 1969). Instead, I am concerned with analyzing some recent developments in the debate over the status of mental imagery.

There are those (e.g., Cooper & Shepard, in press; Kosslyn & Pomerantz, 1977; Paivio, 1976) who argue that visual imagery is encoded in terms of properties that are quite spatial and modality specific. This position is referred to as the imagery position. (This terminology will be continued here, although it is a potential source of confusion. The issue at stake is not whether imagery exists, but what mental representations underlie it.) On
the other side of the issue are the propositional theorists (Anderson & Bower, 1973; Chase & Clark, 1972; Pylyshyn, 1973, 1976; Reed, 1974), who argue that imagery is encoded in an abstract propositional format and that this same format is used to encode verbal information. In contrast, the imagery theorists want to make a sharp distinction between the codes used for verbal and visual information. My principal purpose is to provide a critical analysis of the arguments given for both sides of this debate.

This article will have the following structure: First, I will attempt to specify what a propositional theory is, and what an imagery theory is; second, I will consider those arguments that have been made against an imagery position; third, I will consider those arguments that have been made for the imagery position; finally, I will take a more abstract look at these issues and consider what evidence might be useful in deciding between the two positions.

My central thesis is that arguments concerning the representation of information have been misdirected. Any claim for a particular representation is impossible to evaluate unless one specifies the processes that will operate on this representation. Arguments for or against a particular representation are only valid assuming a particular set of processes. These arguments are not valid assuming other processes. Pylyshyn (Note 1) has made the point that a theory must be considered as a representation plus process. As he writes,

> The appropriate subject of our analysis of representation should be not the representation per se but a representational system consisting of the pair (representation, process). (Pylyshyn, Note 1)

Despite this common realization, Pylyshyn's conclusions about the representation for imagery are different than mine.

Defining Terms

Concept of a Proposition

It is classically accepted that there are three features that define a proposition (e.g., Frege, 1960): It is abstract, it has a truth value, and it has rules of formation. Abstractness is a concept originally defined with respect to sentences. A proposition can be denoted by a sentence but is more abstract than the sentence. It is not tied to the particular features of the sentence. Therefore, the attempts to equate "propositional" with "verbal" are fundamentally wrong. The notion of abstractness is related to a concept of invariance under paraphrase (see Norman & Rumelhart, 1975, or Schank, 1972). This is the idea that all linguistic paraphrases or cross-language translations would be assigned the same propositional representation. The concept of invariance under paraphrase is not part of the classical definition of a proposition and is in dispute (see Anderson, 1976; Woods, 1975). Nonetheless, it is the case that all propositional formalisms practice some degree of invariance under paraphrase.

The notion of truth values means that it makes sense to inquire whether propositions are true or false. Propositions are analyzed as bearers of truth value in logic from which propositional representations were imported into psychology. However, use of propositional theories does not imply a psychological theory of meaning based on the concept of truth.

The structural aspect of propositions means that there is an explicit set of rules for determining what is and what is not a well-formed proposition. In logic, rules of inference are framed with respect to these structural properties. In psychology, cognitive laws are formulated with respect to these structural properties (for some examples of cognitive laws so formulated, see Anderson, 1976; Anderson & Bower, 1973; Fredericksen, 1975; Kintsch, 1974; Norman & Rumelhart, 1975).

While there do exist agreed-upon criteria for a propositional representation, these criteria lack the character of operational definitions. Therefore, it is not a cut-and-dried matter to decide if a particular representation is propositional. I will try to follow what has become standard usage in the propositional-imagery debate. This usage would include as propositional representations such things as Quillian’s (1969) associative network, Schank’s (1972) conceptual dependency, or Winograd’s (1972) SHREDL data base. In each case, the representation is abstract, has at least minimal rules of units of representation as truth values. I mean to structures, stimulus-response associations (such as propagationalists), procedural structures. With respect possibility, many computerepresent propositions as initially could be a procedure when a list structure of a proposition, although it does not seem to arise in pr criterion for propositional it has clearly identifiable mental constituents.

Concept of an Image

No one seems to deny nomenon called mental in hand, there is considwhether there is a useful struct called an image, is thought to be, howe Generally, image theorists properties very much li are given in phenomenon subjects while exper Thus, these theorists might be construed as to make phenomenologic tives of the psychologic counts for these phenomen for other relevant data. In contrast, proposition attempt to derive such mechanisms that do not fit with the phenomenon.

The article by Pylyshin is highlighted concerning the role of planar concept. He to pictures underlies theception of an image. The whole vocabulary of appropriate for describing pictures. We ness of images, of scene past the study of properties depicted in pictures. Pylyshyn (1973) and (1973) have criticized
Representations for Mental Imagery

No one seems to deny that there is a phenomenon called mental imagery. On the other hand, there is considerable debate over whether there is a useful representational construct called an image. Just what an image is thought to be, however, is a sore point. Generally, image theorists ascribe to an image properties very much like the properties that are given in phenomenological reports by subjects while experiencing an image. Thus, these theorists’ use of the term image might be construed as an attempt by them to make phenomenological features the primitives of the psychological mechanism that accounts for these phenomenological reports and for other relevant data on the use of imagery. In contrast, propositional theories must attempt to derive such imagery reports from mechanisms that do not have much in common with the phenomenological reports.

The article by Pylyshyn (1973) was important in highlighting a number of issues concerning the role of an image as an explanatory concept. He argued that a metaphor to pictures underlies the image theorists’ conception of an image.

The whole vocabulary of imagery uses a language appropriate for describing pictures and the process of perceiving pictures. We speak of clarity and vividness of images, of scanning images, of seeing new patterns in images, and of naming objects or properties depicted in images. (Pylyshyn, 1973, p. 8)

Pylyshyn (1973) and Anderson and Bower (1973) have criticized this picture metaphor model of an image on many counts, for example, that it makes predictions that have been empirically disconfirmed, that there are internal inconsistencies in the use of this model, that there are better models for many phenomena, and that the model involves many inelegant, unpalatable assumptions.

Kosslyn and Pomerantz (1977) and Paivio (1976) have argued that the picture metaphor is a straw man, that few imagery theorists support such a metaphor in any literal sense. They conclude that the picture metaphor is not defensible. As Paivio (1976) writes,

The wax tablet of picture metaphor is open to the kinds of criticisms that were directed at it over the ages. Today, however, they are largely directed at a straw man because no imagery researcher accepts the metaphorical view as a working theory. (p. 2)

Kosslyn and Pomerantz (1977) write:

Pylyshyn’s attacks are based on a particular definition of imagery, namely, the picture-in-the-head hypothesis. We agree with Pylyshyn that this approach is untenable but fail to see what is gained by attacking such a strawman. No serious student of imagery holds this view. (p. 57)

The problem is that imagery theorists have failed to say precisely what an image is. It is somewhat hard to debate the status of a concept when no one will specify what the concept is. Both Paivio (1976) and Kosslyn and Pomerantz (1977) propose that an image of an object may be equated to the product of perception of the object, and they distinguish this representation from a picture. However, the current state of knowledge about perception is inadequate to decide the appropriate representation for an image. Also, it is the case that the propositional theorists (e.g., Anderson & Bower, 1973; Palmer, 1975; Pylyshyn, 1973) want to equate the output of perception with a propositional representation, so they too endorse the claim that imagery and perception have the same representation.

Consider some of the other explications of an image offered by Kosslyn and Pomerantz (1977):

Images are like surface displays generated on a cathode ray tube by a computer. (p. 70)

Images, once formed are wholes that may be compared to percepts in a template-like manner. (p. 66)
What is a display on a cathode-ray tube or a template but another name for a spatial array of light information? What is a spatial array of light information but a picture? It seems that the picture metaphor is the only available model of imagery. By "picture," I mean some form that represents information as a spatially structured array of light information. Consider the recent model by Kosslyn and Shwartz (1977). This computer simulation model is clearly the most concrete theory formulated by anyone in the imagery camp. In this model, an image is represented by a set of points encoding the location in two-dimensional space of the contour points of an imaged object. Clearly, such a representation is "pictorial" in my sense.

Thus, it seems the picture metaphor is the imagery theory. Moreover, I think all sides have prematurely abandoned the picture metaphor. What I would like to do is to consider a semirealistic picture model and show that it is not susceptible to the criticisms typically made of such a model.

What I want to defend is not simply a picture model, but that model as a part of a larger theory. This larger theory is basically the dual-code model of Paivio (1971). He proposed that there are two types of information representations, visual and verbal. I will equate the visual representations with pictures. The verbal information will be equated with word strings. These word strings can be thought of as acoustic objects temporally structured, although this interpretation is not essential to what follows. What is important is that they not be interpreted as abstract propositions. I think it is a correct reading both of Paivio's exposition of his theory and of the attacks on it by the propositional camp that nothing in his theory can be considered abstract propositional. It is important that the theory I propose be faithful to the Paivio position. If it is, it definitely should be clear that I have not bolstered the power of a picture theory by sneaking in propositions.

**Verbal–Pictorial Dual-Code Theory**

The picture model I will discuss is also somewhat similar to the idea of Kosslyn and Shwartz (1977). The principal difference is that my model is motivated purely to provide a counterexample to anti-imagery claims. Therefore, it contains some "outrageous" suggestions that I would probably want to modify were I to advance this model as a serious proposal. The point of the model is not its plausibility, but rather how a "brazen" picture model cannot be disconfirmed. If this model is not subject to any difficulties, presumably models like that of Kosslyn and Shwartz, constructed with more concern for the prevailing sensitivities of the cognitive community, will also not be subject to decisive contradiction.

Images will be interpreted as \( m \times n \) arrays of dots, where each dot can be specified according to color and intensity. We will also have as a different storage medium (primarily for verbal information) strings of words. There will also be directed associations that can connect strings and pictures.

Consider how we might represent the chessboard in an end-game position, as in Figure 1. Figure 2 shows a possible representation. Note that the situation has been broken into a number of overlapping images connected by associations (arrows). This illustrates that a single external situation may be represented by a number of images. A possible motivation for this fragmentation is the fact that there may be capacity limitations on the amount of information that can be processed. Note also that the chessboard not only stores that images need also that the black king is not captured. This illustrates that situations need not be verbal strings like "king tagged to the images. Providing meaningful interpretation..."
will be equated strings can be read as abstract propositions. This is not new, and as abstract propositions correct reading of the theory and positional camp should be considered important that it is to the Paivio theory and for the power of a propositional model.

**Circle Theory**

This discussion is also of Kosslyn and a difference is not to provide imagery claims. An outrageous suggestion to modify Kosslyn as a serious model is not its "brazen" picture. If this is the case, for concern for the cognitive subject to describe the chessboard in Figure 1. A representation is broken into a series of connected components that are represented.

The motivation for this is that there are portions of the chessboard that are not represented. This illustrates that images need not be complete. Note also that the black king is inaccurately located. This illustrates that visual images need not be veridical. Note also that verbal strings like "knight attacks king" are tagged to the images. This is one way of providing meaningful interpretations to the image.

A final feature to note of this representation is that the picture unit tends to segment out meaningful, coherent subunits of the picture. This segmentation reflects the operation of perceptual procedures that encode and interpret the original object. One can think of these as perceptual chunks such as discussed by Chase and Simon (1973). Therefore, as the imagery theorists would postulate, this representation is identified with the output of the perceptual routine.

The representation in Figure 2 seems to be a faithful embodiment of the dual-code theory proposed by Paivio (1971), that is, there is a pictorial code, a verbal code, and connections between the two. Armed with this dual-code model, let us consider some of the criticisms that are made of images and of their role in dual-code theories.

**Criticisms of Picture Theory**

*Images Are Preinterpreted*

One frequent criticism made of imagery is that there is a preinterpretation of the imagery. This can be seen in the representational system used in Figure 2. The segmentation is meaningful for myself. I admit to being a chess duffer.
theories is that images behave as if they were interpreted, whereas pictures (just arrays of light information) do not:

One's representation of a scene must contain already differentiated and interpreted perceptual aspects. In other words, the representation is far from being raw, and so to speak, in need of "perceptual" interpretation. (Pylyshyn, 1973, p. 10)

We would argue that people store perceptual interpretations of scenes rather than "raw, unanalyzed, textured details" of such scenes. (Anderson & Bower, 1973, p. 453)

Pylyshyn (1973) points out some of the many ways in which our memory for an image behaves as if it were interpreted. We can immediately retrieve an image directly from a meaningful description like "my living room." This would not be expected if we had to scan through a series of pictures. However, it is a feature of the representation in Figure 2 that it can be accessed via verbal labels. The associative interconnections between strings and pictures basically provide this conceptual indexing.

Pylyshyn also notes that we can access meaningful parts of an image such as "the light fixture in my dining room" and that when we forget part of an image, we tend to lose a meaningful part and not a physical portion of the picture that might cross many meaningful units. However, the representation in Figure 2 is organized into meaningful units—producing access to meaningful parts and a potential for loss of meaningful parts.

It is argued that it is wasteful to perceptually analyze a scene and store only the raw scene rather than store the analysis. The scene analysis would have to be performed again when the scene was retrieved. There are two remarks that can be made about this objection. First, perceptual interpretation does not seem very costly for the human system, so it is relatively insignificant whether it has to be performed again. Second, there is no reason why the output of the perceptual process might not be a representation like that in Figure 2, that is, segmented into subpictures with relations indicated among these subpictures but the structure of the subpictures not analyzed.

It is clear from considering these objections that the representation in Figure 2 is "interpreted." This interpretation is provided by the network of associations. Therefore, there is nothing incompatible between the concepts of a dual-code memory and an interpreted memory.2

**Capacity Limitations**

It is argued that the storage demands to encode these pictures would be enormous and exceed the capacity of the brain. However, given current ignorance about the capacity of the brain, this remark has virtually no force. It is also possible in the dual-code model of Figure 2 to reduce the storage demands by choosing a crude grid or by incompletely representing the objects. It is also worth considering the storage demands of alternatives to a picture model. For instance, consider the Baylor (Note 2) model advocated by Pylyshyn (1973). This model requires separately representing every vertex, line, surface, and object (and some of these redundantly) in Baylor’s I-space (contains a representation of the object organized by its vertices) and S-space (contains a symbolic representation of the object). The demands of such a representation also seem very large. The simple fact may be that there is a great deal of information in an image, and any representation of an image will have to reflect this fact.

**Representation of Vagueness**

Information can be represented vaguely in a picture, but it is claimed that the character of this vagueness does not match the character of vagueness in introspective reports of images. A picture can be of poor resolution, but the vagueness of human images seems inexplicable in terms of simple resolution. For instance, Pylyshyn (1973) points out that we may be able to recall what objects were in the room without spatial location. A piece of paper is more like spatial information than identity.

However, this is the problem with the picture again the representational possible to have two board sections without spatial location or situation as "beside," thus to the right or left.

Anderson and Bower show how it is that one lamp was in the room recall any details about lamp. This phenomenon explanations within a possible that we have "lamp" rather than a description. Another have stored a picture that picture with "it" label to indicate we saw details seriously. The transmission of information to the lamp encoding another possibility is the actual lamp with preceding that there is no.

---

2 It might be argued that the representation in Figure 2 is really a propositional representation. This could be argued on one of three bases: (a) the use of word strings like "knight attacks king", (b) the use of the associative network structure, or (c) the use of procedures that we assume will utilize this information. There are two responses to this argument. The first is to note that if one agrees that this proposition is propositional and insists it is propositional, one would be consid...
in the room without recalling their exact spatial location. A poor resolution picture of the room is more likely to preserve relative spatial information than to preserve object identity.

However, this is not an insurmountable problem with the picture metaphor. Consider again the representation in Figure 2: It is possible to have two subpictures of the chessboard sections without encoding their relative spatial location or with encoding only the relation as "beside," thus not specifying whether to the right or left.

Anderson and Bower (1973) also wondered how it is that one could remember that a lamp was in the room without being able to recall any details about the appearance of the lamp. This phenomenon has multiple possible explanations within a dual-code model. It is possible that we have stored the verbal label "lamp" rather than the picture in our scene description. Another possibility is that we have stored a picture but may have also tagged that picture with "uncertain" or some other label to indicate we should not take the visual details seriously. This might be done if the transmission of information from the scene to the lamp encoding was not accurate. Yet another possibility is that we have replaced the actual lamp with a picture of a prototype lamp. Recognizing that this is a prototypical lamp, the system is not willing to use this as a basis for reporting details about the particular lamp we saw. It might be more efficient to use a prototypical lamp than the actual lamp if it were easier to transmit information from the internal representation of a prototype than from an external scene. The possibility of an uncertain lamp or prototypical lamp illustrates an important possibility that has not been frequently recognized in the debate over imagery. This possibility is that there is internal information (i.e., the uncertain lamp or prototype) that subjects cannot report.

**Metaphors Are Misleading**

Another criticism that Pylshyn levels against the picture metaphor is that it is misleading in that it causes people to postulate mental processes by analogy to operations we can carry out on pictures.

For example, one misleading implication involved in using the imagery metaphor is that what we retrieve from memory when we image, like what we receive from our sensory systems, is some sort of undifferentiated (or at least not fully interpreted) signal or pattern, a major part of which (although perhaps not all) is simultaneously available. This pattern is subsequently scanned perceptually in order to obtain meaningful information regarding the presence of objects, attributes, relations, etc. (Pylshyn, 1973, p. 8)

However, it is an open question whether the metaphor is misleading. In particular, results such as those of Kosslyn (1973) on scanning or Shepard (1975) on mental rotation of images would seem to imply that the metaphor leads one in quite fruitful paths.

**Semantic Effects in Picture Memory**

Anderson and Bower (1973) made much of the evidence that pictures are remembered better when they can be interpreted. DeGroot (1965) finds better memory for meaningful chess positions. Wiseman and Neisser (Note 3) find better memory for Mooney pictures that subjects manage to meaningfully interpret. Bower and Karlin (1974) report similar results. Goldstein and Chance (1971) have shown poorer memory for snowflakes, which permit little meaningful structuring, than for...
faces that do. These results are certainly contrary to claims (e.g., Bugelski, 1970) about the mnemonic superiority of the image system, but they really say little about whether pictures are used as a medium of storage. It is quite possible that pure pictures are a poor storage medium. It may be that some meaningful structure as in the chessboard of Figure 2 is needed for good memory. Perhaps retrieval schemes require verbal labels to provide access routes.

**Summing Up the Picture Metaphor**

In conclusion, it has not been shown that there is anything incoherent, contradictory, or impractical in using pictures as a representational format. Whether pictures are suitable depends on the processes that one assumes operate on them. To be sure, one can assume inadequate processes for a picture representation (e.g., a forgetting process that consisted of fading rather than loss of meaningful subparts) as did Anderson and Bower, Pylyshyn, and even some of the image theorists. But, then again, one need not.

**Propositional Representations**

Besides the supposed empirical and logical inadequacies of picture theories, an independent line of argument against them is that picture representations are unnecessary. This argument has two subclaims. First, it is argued that a propositional representation is essential on independent grounds. Second, it is argued that a propositional representation can handle all the phenomena ascribed to a pictorial representation. Therefore, on grounds of parsimony, it is unnecessary to assume pictorial representations in addition to propositional representations. This argument would have some validity if both subclaims were true. However, the first subclaim has not been established.

**Are Propositional Representations Necessary?**

Pylyshyn (1973) argues that it is necessary to propose a propositional code to explain how it is that people can describe pictures in words or create pictures to illustrate verbal material. Similar proposals for an interlingua have been made by Chase and Clark (1972), by Anderson and Bower (1973), and by Fodor (1975). The abstract propositional code would serve as a neutral format into which and out of which pictorial and verbal information could be translated. It serves as a "half-way house" for the process of translating between the two peripheral codes. As Pylyshyn (1973) writes,

The need to postulate a more abstract representation—one which resembles neither pictures nor words and is not accessible to subjective experience—is unavoidable. As long as we recognize that people can go from mental pictures to mental words or vice versa, we are forced to conclude that there must be a representation (which is more abstract and not available to conscious experience) which encompasses both. There must, in other words, be some common format or interlingua. (p. 5)

However, this argument is flawed with a serious internal inconsistency. It is argued that to translate from Code 1 to Code 2, it is necessary to translate Code 1 into a new code, Code 3, and from Code 3 to Code 2. However, this argument leads to an infinite regress. To translate from Code 1 to Code 3, a new Code 4 would be needed and so on. It is simply not the case that it is necessary to have a propositional or any other intermediate code for translation. By careful analysis, it might be possible to show that an interlingua makes the translation process more efficient, but such an analysis has not been forthcoming.

Another argument used by both Anderson and Bower (1973) and by Pylyshyn (1973) is that a propositional code is needed to represent meaning. For instance, Pylyshyn argues that one needs this code to represent a relation like "attacked by" in a chessboard. His example inspired Figure 2, which shows that this information can be represented by verbal strings. The "meaning" can be contained in the processes that interpret the strings.

Anderson and Bower (1973) argue that a propositional code is needed to deal with the considerable evidence that subjects tend to remember only the meaning of sentences and not their exact wording. For instance, subjects show confusion about whether they heard an active sentence or the passive equivalent (Sachs, 1967). There are any number of explanations of the require postulations. For instance, each utterance represents the sentence in a picture. Such pictures would not be the original form that subjects refer to, but common stringent form be represented as acting this would be of the string; judging the form of the string. Another possible heard is stored in it: it does not pay attention between active and normally this is not.

**Power of a Proposition**

It seems to be a that any well-specified proposition can be represented by a proposition that can be presented in any other sense. The fact that propositions are appropriate for facts about knowledge, no one argument, although (1977) ascribes it to that argument is obvious, can be represented because the proposition will lead to the conclusion.

The power of propositional has been used by (1977) to argue a

The propositional is a convenient one is that there is no inherent problem with representing information, the propositional is so powerful, that by use of one with equal ease at finding or its content.

I will return to this point, I see this criticism A
of explanations of the phenomena that do not require postulations of a propositional code. For instance, one could propose that subjects represent these sentences in memory by pictures (e.g., a picture of a boy hitting a girl). Such pictures would represent the meaning but not the original form. Another possibility is that subjects convert various sentences to a common string format, so all sentences might be represented as active. The processor, knowing this, would be smart enough not to use the form of the string in memory as a basis for judging the form of the string that was studied. Another possibility is that the exact string heard is stored in memory, but the processor does not pay attention to the differences between active and passive sentences because normally this is not important.

Power of a Propositional Code

It seems to be a generally accepted claim that any well-specified set of information can be represented by a set of propositions. Thus, propositions can represent any information in an image or in a sentence or information from any other source. One might argue from this fact that propositional representations are appropriate for representing images. To my knowledge, no one has explicitly made this argument, although Kosslyn and Pomerantz (1977) ascribe it to Pylyshyn. The counterargument is obvious: Even if all information can be represented propositionally, it does not follow that the propositional representation will lead to the correct empirical predictions. The power of propositional representations has been used by Kosslyn and Pomerantz (1977) to argue against them:

The problem with propositional theories, on the other hand, is that they may be too powerful. They possess no inherent constraints, and the theorist must add restrictions onto his theory to make it conform with experimental observations. The propositional language is so powerful that one can use it to formulate almost any kind of theory that one desires, predicting with equal ease, it would seem, any experimental finding or its converse. (p. 62)

I will return to this question of the plasticity of representations later in this article to give it a more thorough examination. However, at this point, a specific rejoinder is required to this criticism: A representation without any process assumptions is not a theory. By making different process assumptions, it is possible to have quite different theories with the same propositional representation. We cannot test representations but only representation-process pairs. It is not an argument against a representation-process theory that there happens to be a different theory with the same representations but different processes that makes the opposite prediction.

Propositional Representations and Inference Making

It is sometimes assumed (e.g., Anderson, 1976; Fodor, 1975; Pylyshyn, 1973; Anderson & Kline, Note 4) that there is a particular affinity between propositional representations and the task of making inferences. The claim is that propositional representations are particularly well suited for the operation of inference making. Propositions, because of their abstract truth-bearing character, only represent what is necessary to judge the validity (or plausibility) of an inference. There can be, therefore, a reduction in the complexity and/or number of inference rules if they are formulated in terms of propositions. For instance, most propositional systems would represent active and passive sentences identically. Anything that follows from an active also follows from a passive. A propositional system avoids the need to have separate inference rules for actives and passives. This is probably the reason why almost every artificial in-

---

3 This conjecture about the universality of propositional representation can be seen as related to a conjecture known as Church's thesis or Turing's thesis (see Minsky, 1967). This is the thesis that any well-specified behavior can be computed by a recursive function or Turing machine. Given that any well-specified information should be capable of entering into well-specified computations and given that the recursive functions can be represented in predicate calculus, a propositional-based language, there is a sense in which it would follow from Church's thesis that any well-specified information can be propositionalized. However, this argument implies a rather unsatisfactory way of propositionalizing information. Therefore, I think it is better to let this conjecture stand as an independent claim. It is explicitly accepted by Kosslyn and Pomerantz (1977).
Intelligence program for making inferences uses some variety of a propositional data base.

However, these considerations do not force the conclusion that human inference making is done in terms of abstract propositions. It may be the case that while artificial intelligence programs are capable of such abstraction, humans are not. In fact, there is some evidence that the human system is not capable of such abstraction. It has been shown that semantically irrelevant linguistic details of an argument affect the time and success of making a valid conclusion (e.g., Clark, 1969; Rips & Marcus, 1977). While these data can be interpreted within a propositional model, they do serve to deflate claims about the advantage of propositional representations in abstracting out only truth-relevant information for inference making.

Pro-Imagery Arguments

So far, we have considered arguments against imagery theories and arguments for propositional representations. These two types of arguments define the anti-imagery position. Now I would like to turn to the consideration of the arguments that have been made for imagery theories. By "imagery theories," I mean theories based on the picture metaphor. As noted earlier, this is the only explicit interpretation now available for the concept of an image.

Computational Advantages of Imagery Representations

Kosslyn and Pomerantz (1977) argue that while the same information can be represented via propositions or images, it is easier to perform certain types of computations on images. Therefore, images have a computationally useful function. It is further assumed that if a function is useful, it is likely to be found in the human head. I will be making a similar argument for efficiency considerations later in this article. A basic problem is that claims about one representation being more efficient than another are always based on implicit assumptions about the processes to be used with the two representations. With other process assumptions, these claims may no longer be valid.

Kosslyn and Pomerantz (1977) present an example to make their point. They contrast two formats for representing geographical information: a map with cities on it versus a chart of intercity distances. The first is thought of as similar to an image representation and the second as more similar to a propositional representation. They claim that these are "isomorphic" to each other in all important respects, since the formats contain identical information and either one can be generated from the other (Kosslyn & Pomerantz, 1977, p. 60; however, this is not totally true in that north–south orientations cannot be recovered from the chart.) They claim that different types of computations are appropriate to the two representations:

If we want to know quickly whether three cities fall on a straight line, we consult a map; if we want to know the total distance of an air flight from New York to Los Angeles to Miami, we consult the chart. (Kosslyn & Pomerantz, 1977, p. 61)

However, these claims are only true assuming a particular algorithm on a particular device. There is a relatively simple algorithm for solving the first problem given chart distances: See if two of the distances sum to the third. In some implementations, this algorithm would be more efficient than any maps-scanning operation. Conversely, if it were difficult to add the distances, but one had a string marked off to translate string length into miles, the second problem might be better solved with the map.

It is not possible to decide issues of efficiency of a representation without knowledge of the procedures that will be operating on them. However, later I will argue a point somewhat similar to that of Kosslyn and Pomerantz, which is that efficiency considerations may lead to the need for specialized representations. In contrast to Kosslyn and Pomerantz, I do not think we can decide on the nature of the specialized representations.

Arguments Based on Introspection

Introspections probably provide the most intuitively compelling evidence for the existence of a picturelike image code and for the distinction between this code and a verbal code. Many people find the picturelike quality of their images rather striking. Of course, there is little or no such introspective count of visual imagery in introspective reports. The introspective reports are data that require any other data. However, to suppose that the best evidence for verbal properties of an image is a program could be with the introspective reports. The introspective reports are data that require more introspective reports. The introspective reports are data that require any other data. However, to suppose that the best evidence for verbal properties of an image is a program could be with the introspective reports.
They contrast geographical information it versus a picture. The first is a more representative picture similar to what we have. They claim that each other in all formats contain the one can be brooked & Pomerantz is not totally representations cannot be considered the same. They claim that formats are appropriate.

Rely true assuming a map; if we want a flight from New consult the chart.

Arguments for Imagery From Experimental Demonstration

Most imagery theorists do not base their position entirely on such introspective evidence. Rather, they will marshall a large series of experiments to support their position. Therefore, it is important to consider some of the classes of data presented and to evaluate the ability of these data to decide issues of internal representation. An early class of data concerned the supposed superior mnemonic capacity of the imagery system—better memory for pictures or for imaginable words, better memory under instructions to image, and so on. The culmination of the line of evidence is Paivio’s (1971) book, which provides an extensive review. It appears that reliance on such data is waning. One reason for this waning is that further evidence has indicated that the pictorial material is not superior to verbal material. First of all, verbal material when “deeply processed” can display similarly high levels of memory (Anderson, 1976; Anderson & Bower, 1973; see Craik & Lockhart, 1972, for a general discussion of levels of encoding). Second, the good memory for pictorial material only seems to apply when that pictorial material can be meaningfully interpreted (Anderson & Bower, 1973; Bower & Karlin, 1974; Goldstein & Chance, 1971; Wiseman & Neisser, Note 3)

There is now a second generation of research that seems to support the imagery approach. This research is more concerned with operations on images in immediate memory than the treatment of images in long-term memory. The logic of this research has been succinctly analyzed by Shepard and Podgorny (in press). The attempt is to show that when subjects process objects mentally, the functional consequences for behavior are quite similar to those when subjects process the actual objects. This is evidence that the mental or imaginal representation of an object is the same as the perceptual representation. Shepard and Podgorny note that this equivalence need not imply that either representation is picturlike. However, the frequent interpretation of such an equivalence is that the imaginal representation must be picture-like because (assumption) the perceptual representation is picturalike.

At this level, the logic of these experiments is not at all compelling because many theorists (Anderson & Bower, 1973; Palmer, 1975; Pylyshyn, 1973; Winston, Note 5) would reject the notion of a picturalike representation for the products of perception and would rather assert a propositional representation. In this section, I would like to focus on some of the specific demonstrations and see if they have any additional force beyond this general argument. I will not have room to consider all types of data that have been enumerated in such papers as Kosslyn and Pomerantz (1977), Paivio (1973), and Shepard and Podgorny (in press). However, I will consider a representative subset. This will be enough to make a general point.

Mental Rotations

One of the most influential phenomena uncovered in recent research in cognitive psy-

---

*For instance, if a person reports to us, “My mind is swarming with ideas,” would we want to assume a mental representation of a swarm of bees?
chology has been that of mental rotation (see Cooper & Shepard, in press; Metzler & Shepard, 1974). The basic finding is that the time to decide that one object is a rotation of another object is a monotonic and often linear function of the amount (degrees) of rotation. This is taken as evidence that subjects mentally rotate an image of one object of the pair into congruence with the other object. The continuous nature of the function is taken as evidence that the subject must go through all or many intermediate states in rotating the object. In the Kosslyn and Shwartz (1977) simulation, the image is moved through a series of small changes in orientation.

It is a simple matter to propose a propositional model that mimics this image model. The model would involve a propositional description of an object and of its orientation in space. Just as Kosslyn and Shwartz compute a series of small changes in their image, so a series of small changes can be computed in the propositional representation. Given that it is easy to produce the phenomena in a propositional model, one might wonder why propositional theorists (e.g., Anderson & Bower, 1973) have been so loathe to propose it, and why most find the image account more appealing than the propositional account. Consider Kosslyn and Pomerantz's (1977) criticism of the propositional account:

The imagery account seems somewhat plausible and relatively straightforward. The propositional account seems less satisfactory: Aside from the problem of not knowing how to represent the letters in the first place, it is not clear why rotation is gradual in such a system. It should be especially easy to rotate an image 180° because all the relations could simply be reversed (e.g., right becomes left). To rotate an image 45° should be more difficult, because more complex substitutions must be implemented. Nevertheless, subjects take longer to rotate an object 180° than 45°. It appears that people do not (or cannot) skip from one orientation of an image directly to another, but must proceed gradually. Such a prediction does not follow from basic concepts of propositional representation. (p. 69)

However, it is no less ad hoc to propose that small images suffer a grain limitation. A priori, one might have supposed that smaller images have smaller grain. It is entirely ad hoc to propose a complexity limitation on an image. A property frequently ascribed to images (e.g., Cooper, 1975) is that they can be processed in an unlimited-capacity parallel manner. In contrast, there does exist a propositional model (ACT; Anderson, 1976) that incorporates a complexity limitation on the number of propositions that may be active. This limitation was motivated on grounds entirely independent of Kosslyn's (1975) result. There is nothing intrinsic to an image or to a propositional representation that would lead one to predict a size and complexity. It requires additional, nontrivial assumptions to predict these results without some basis for calling on more arbitrary than the others.

Effects of Image Size and Complexity

Kosslyn (1975) has found that subjects take longer to verify that an imaged object has a certain property if they are instructed to make the image of the object small. Kosslyn argues that there is a "grain" limitation on the mental image and that it is difficult to properly represent the details of a small image. He relates this to the supposed perceptual fact that it is difficult to discriminate details of small objects when presented. Kosslyn has also shown that there is a complexity limitation on an image—that it takes longer to verify that an object has a certain property when it is imaged along with a more complex object.

These results can be accounted for by a propositional model that assumes that a subject activates fewer propositions to represent an object when instructed to imagine it small and that he can activate fewer propositions when he must represent another complex object. However, Kosslyn and Pomerantz (1977) argue that this propositional account seems less satisfactory than the imagery account: Why should people access less information about an object when asked to "image it small?" This seems ad hoc; a propositional model would not lead one to expect such effects. (p. 71)

Judgments of Magnitude

There are experiments (Paivio, 1975) that show judging two mental objects a dimension on which time decreases with the objects on that dimension, except for Paivio's (1975) condition in which two items (e.g., larger. In one context the words; in another a picture of a true-size ratio. Reacting to judgements decreased in e.g. named or depicted increased Paivio's picture condition control to support his claim that a change in the same effect with the presence of projections as he would g. representations. (It is, however, that the size twice as large with which it is with picture-cued. It is more likely that this is evidence representation is in a format rather than a format.

There is some mechanism underlying judgments (Banks & Holyoak, 1977). How to the issue at nothing incompatible or well known that are propositional model. Treatment of many other conventions that do not seem to matter much as to a process defined of
to a propositional representation that would have led one to predict Kosslyn's results of size and complexity. Either representation requires additional, nontrivial process assumptions to predict these results. There is no justifiable basis for calling one set of assumptions more arbitrary than the other.

**Judgments of Magnitude**

There are experiments (e.g., Moyer, 1973; Paivio, 1975) that show that when one is judging two mental objects with respect to a dimension on which they vary, reaction time decreases with the separation of these objects on that dimension. In a model experiment, Paivio (1975) had subjects judge which of two items (e.g., a lamp or a zebra) was larger. In one condition, Paivio presented the words; in another condition, Paivio presented a picture of the objects in their true size ratio. Reaction time to make this judgment decreased in either case as the items named or depicted increased in size disparity. Paivio's picture condition provides the needed control to support his claim that he is getting the same effect with word-cued representations as he would get with picture-cued representations. (It is somewhat disturbing, however, that the size of the effect is about twice as large with word-cued judgments as it is with picture-cued judgments.) Paivio argues that this is evidence that picture information is represented in an analogue image format rather than a discrete propositional format.

There is some dispute about the exact mechanisms underlying these comparative judgments (Banks, Clark, & Lucy, 1975; Holyoak, 1977). However, a more relevant point to the issue at hand is that there is nothing incompatible between the notion of a propositional model and the idea that judgment time can vary with object magnitude. Magnitudes can easily be represented in a propositional model as arguments of propositions. It is easy to provide predicate calculus treatments of many physical and mathematical domains that deal in continuously varying quantities. There is no reason why the distance effects reported could not come from a process defined on these propositional representations. Of course, there is no reason to have expected the effect. However, there is no reason to have expected the effect given an image representation either. The reason to have expected the effect comes from a knowledge of the results obtained in perception. However, as noted earlier, the products of perception can either be represented in a picture format or a propositional format.

**Unlimited-Capacity Image Processing**

A number of experiments (Cooper, 1975; Nielson & Smith, 1973; Smith & Nielson, 1970) lead to the conclusion that images can be matched to perceptual objects as templates or in an unlimited-capacity parallel manner. For instance, Smith and Nielson (1970) find subjects unaffected by complexity in recognizing faces when they have encoded the faces as whole units, but not when they have encoded the faces as verbal strings. This finding has been used as evidence for image over propositional representations. However, there is nothing at all incompatible between propositional representations and parallel processes. The ACT model (Anderson, 1976) is an example of a propositional system in which all basic processes are parallel and some are of effectively unlimited capacity. The spreading activation of Quillian's (1969) TICE model can be conceived of as an unlimited-capacity parallel process (but see Collins & Loftus, 1975).

There are two aspects of this phenomenon that are potentially troublesome, however. The first is the possibility that unlimited-capacity parallel processes can be obtained with encodings of visual information but not with verbal encodings. If one abstract propositional encoding underlies all memory, it would be hard to see why different principles would govern the representation when it encodes information from different sources.

Second, the very phenomenon of truly unlimited-capacity parallel processing is puzzling. I know of no well-understood device that could make recognition discriminations such as found above without an increase in processing time after some bound on object complexity was exceeded. There is sometimes expressed a belief that analogue processes, because of their continuous nature, provide
an infinite amount of information. However, discriminability limitations on a system's ability to respond to small differences imply a finite bound on the amount of information per analogue process (Page, 1965; Rabin, 1964).

It is possible to have a computational system that appears to have unlimited capacity within a certain range. For instance, one can have a large number of parallel processors, only some of which are used on simple problems. The greater difficulty of more complex problems can be masked by recruiting the unused processors. A complexity effect would only be observed for problems that were sufficiently complex to exhaust the available processors.

Template matching is frequently given as an example of unlimited-capacity parallel processing. However, templates have only a certain fineness of discrimination. Faced with distinguishing among objects that require finer discrimination, either a template-matching program will make errors, or more computational resources will have to be allocated to permit a finer grid to be computed.

There are a number of possible responses to reports of unlimited-capacity parallel processing of images. One is to give up the principle of finiteness of human-computing resources. A better response is to assume that the complexity has not been pushed far enough, that we are still in the range where the capacity of a finite set of parallel processes has not been fully exploited. A third response is to quibble with the demonstrations of unlimited parallel processing. It may be that less information is being processed about the more complex objects, but that the tests fail to bring out evidence for this degradation in processing. Consider the experiment by Cooper and Podgorny (1976) on recognition of Attneave polygons: They varied the number of points in the polygons from 6 to 24 and obtained no effect of this measure of complexity on recognition time. As distractors, they used transformations of the target polygons with some of their points perturbed. There were no more errors made to distractors for polygons of greater complexity. This was used as evidence against the possibility of a less careful processing of the more complex shapes. However, the more complex distractors also had more points perturbed. Therefore, one would have to remember a smaller portion of the points from the more complex figures to achieve the same probability of detecting a distractor.

Cooper and Podgorny (1976) had subjects rate similarity of the distractors to the targets. The distractors were, on the average, of equal similarity for the targets of different complexity. Apparently, more points have to be perturbed in a more complex polygon to achieve the same difference in judged similarity. However, to have shown unlimited-capacity processing, they should not have used distractors of equal similarity but rather distractors that had the same number of points perturbed and by the same amount. This would be a test of whether all the information in the target was being processed in the complex figures. To perform the experiment in this manner, distractors of greater similarity would have to be used for more complex stimuli. Cooper and Podgorny found that there were more errors for more similar distractors; so presumably, in the prescribed experiment, they would find that subjects were not performing as well with more complex figures, arguing against unlimited parallel processing.

General Analysis of the Problem of Discriminating Among Representations

I have been arguing that the many examples proposed by image theorists do not provide greater evidence for an image representation than they do for a propositional representation. In each case, it was found that evidence for a particular representation was really evidence for a particular process and that there was no good reason to associate the process with a particular representation (there were important second-order complications with the unlimited-capacity issue). Similar difficulties were also shown with the arguments given by propositional theorists. I would like to formalize the general point that these examples illustrate.

This general point is that it is not possible for behavioral data to uniquely decide issues of internal representation. One cannot just test a representation in the abstract tests of the representation with certain assumptions that use the representation test a representation process that given a set of an image representation that operate on it. One equivalent set of assumptions are that one can be given a propositional representation as an example of parallel processing.

The argument that I this claim has been ch...
of internal representation. The reason is that one cannot just test questions about a representation in the abstract. One must perform tests of the representation in combination with certain assumptions about the processes that use the representation. That is, one must test a representation-process pair. One can show that given a set of assumptions about an image representation and a set of processes that operate on it, one can construct an equivalent set of assumptions about a propositional representation and its processes. Or one can be given a propositional theory and construct an equivalent imagery theory. In fact, it is possible to establish a more general claim: Given any representation-process pair, it is possible to construct other pairs with different representations whose behavior is equivalent to it. These pairs make up for differences in representation by assuming compensating differences in the processes.

The argument that I will give to establish this claim has been changed somewhat from an argument to the same point given in Anderson (1976). It has been changed both to make the claim more general and to try to make it clearer why it is possible to achieve equivalences among different representations.

Suppose one had a theory of internal representation, $g$. This theory would specify a set (probably infinite) of possible internal representations, which we can denote $I_1, I_2, \ldots, I_n$. However, a theory that only specifies internal representations says virtually nothing about behavior and therefore is not testable by itself. One must specify some processes that operate on these internal representations in order for behavior to occur. It is useful to discriminate among three types of processes. There are encoding processes that map external stimuli, $S_i$, into internal representations, $I_i$. These encoding processes will be denoted by the function $E$. The operation of encoding a stimulus can be represented as $E(S_i) = I_i$. Second, there are processes of internal transformation that can be represented by the function $T$. The notation $T(I_i) = I_j$ represents a transformation from internal representation $I_i$ to $I_j$. Finally, there are decoding processes specifying how internal representations are manifested as responses. So if state $I_i$ results in response $R_j$, this will be denoted $R_j = D(I_i)$.

This is a quite general framework for representing cognitive theories, and I do not think it blurs any significant issue involved in the image-propositional controversy. Figure 3a represents a possible scheme of information processing within this framework. We have three stimuli ($S_i, S_j, \text{ and } S_k$) being presented but not necessarily at the same time. By the encoding process $E$, these result in internal representations $I_i, I_j, \text{ and } I_k$. $I_i$ and $I_j$ result by $T$ in an intermediate internal representation $I_m$. To preserve the notation of $T$ mapping one representation into another, we will assume that a conjunction of internal representations can itself be an internal representation. So $T(I_i \& I_j) = I_m$. There are other transformations illustrated that occur at various points of time: $T(I_k) = I_n$ and $T(I_i \& I_k) = I_p$. Finally, the decoding process maps $I_m$ and $I_p$ into response $R$.

If we restrict ourselves to behavioral data, we cannot directly observe the internal processes $E$, $T$, and $D$ nor the internal representations. All we observe is that at various times, the stimuli $S_i$, $S_j$, and $S_k$ arrive and that sometime later response $R$ is emitted. The question of interest is whether behavioral data (i.e., observation of the contingencies between such events and the time of these events) are adequate to constrain a theory of internal representation. Such a theory of representation will be part of a model, $M$, that also specifies the processes that operate on the representation. It will be argued that models with very different theories of representation can perfectly mimic the behavioral predictions of $M$. These alternative models will compensate for differences in the representation by different assumptions about the processes. Therefore, these models are not discriminable from $M$ on the basis of behav-

---

*The hedge "virtually" is inserted because we can assume that whatever its representation, there are certain behaviors of which a system is logically incapable, for instance, solving the halting problem (see Minsky, 1967). Note these "logical" predictions about behavior would be true of all representations and so do not provide a way of discriminating among them.*
ioral data. And, therefore, the representation assumed by $M$ is not discriminable from the very different representations assumed by the other models.

Let $\sigma$ be a specification (e.g., a grammar) of the possible representations $I_1, I_2, \ldots$ under $M$. $\sigma$ by itself is not a very interesting theory of representation. For instance, to assert that our representation was predicate calculus would not be very informative unless we had some idea what events would lead to a representation like given (John, Mary, ball). It would be a very different theory of representation if this formula were used to represent the meaning of “Fred eats pizza” than if it were used to represent “John gives the ball to Mary.” That is, we want a theory of representation to specify how various inputs are represented. A theory of representation should include the encoding process $E$. This forces a connection between our representation and the external world. Let us refer to the pair $(\sigma, E)$ as $M$’s theory of representation. The pair $(T, D)$ will be $M$’s theory of the processes that operate on the representation. What I will show is that given a different theory of representation $(\sigma^*, E^*)$, one can embed that theory in a different model $M^*$ with different processes $(T^*, D^*)$ and have $M^*$ mimic $M$ in all its predictions about behavioral data.

**Preservation of Internal Distinctions**

It is not the case, however, that any theory of representation can be made part of a model $M^*$ that mimics $M$. The theory of representation must satisfy a condition called preservation of internal disti-

![Figure 3](image-url)

(a) A representation of the transformation of information representations in model $M$. (b) A representation of the transformation of information representations in model $M^*$, which mimics $M$.
tion of internal distinctions: One theory of representation \((\delta^*, E^*)\) preserves the internal distinctions of another theory \((\delta, E)\) if there is a one-to-one mapping \(f\) from \(\delta\) to \(\delta^*\) such that \(f\) has a computable inverse, which will be called \(f^{-1}\); and \(f^{-1}\) will map the encoding of \(S\) in the mimicking theory \(E^*(S)\) into the encoding of \(S\) in the target theory \(E(S)\), that is, \(E(S) = f^{-1}(E^*(S))\) for all \(S\). The fact that the mapping is one to one assures that any distinctions among representations in \(\delta\) will be preserved in distinctions among representations in \(\delta^*\). The fact that \(E(S) = f^{-1}(E^*(S))\) assures that \(E\) and \(E^*\) assign corresponding representations to the same stimuli. One consequence of these requirements is that for \(S_1\) and \(S_2\); if \(E(S_1) \equiv E(S_2)\), then \(E^*(S_1) \equiv E^*(S_2)\). That is, the encoding process \(E^*\) does not fail to preserve any of the distinctions among stimuli that are preserved by the process \(E\). It will prove to be necessary that \(f\) have a computable inverse to guarantee that process assumptions \((T^*, D^*)\) of the mimicking model \(M^*\) can be computed. One would not want to propose processes that could not be computed and hence (by Turing’s thesis or Church’s thesis, see Minsky, 1967) were not capable of specification.

While the condition of preservation of internal distinctions excludes some representations from the guarantee of mimicry, it is not so severe as to exclude all interesting possibilities from mimicry. In particular, we could have a propositional model mimic a picture model. The mapping \(f\) in this case would transform the picture into a complete propositional description. (I assume it has been generally conceded that all the information in a picture can be propositionalized.) It is also the case that a propositional model can be mimicked by a dual-code model like that of Figure 2. In this case, \(f\) would map a propositional representation into a dual-code representation that contained the same information.

One might wonder under what conditions the mapping would have a computable inverse. If \(E\) and \(E^*\) are primitive recursive (Minsky, 1967), a mapping \(f\) can be constructed with a computable inverse \(f^{-1}\). One can simply make \(f = E^* \cdot E^{-1}\), where \(E^* \cdot E^{-1}\) denotes the combination of applying first the inverse of \(E\) and then \(E^*\). Similarly, \(f^{-1} = E^* \cdot E^{-1}\). If \(E\) and \(E^*\) are primitive recursive, \(E^{-1}\) and \(E^*\) will be computable; and hence, \(f\) and \(f^{-1}\) will be computable. Note that the above argument establishes that \(f\) and its inverse are computable by showing one way to compute \(f\), that is, \(E^* \cdot E^{-1}\). It does not mean that this is the best or only way to compute \(f\). For instance, in mapping from a picture representation to a propositional representation, there would be no need to map to the external stimulus by \(E^{-1}\) and then to the propositional representation by \(E^*\). A more direct propositional encoding of the picture should be possible.

It should also be noted that for any theory of representation \((\delta, E)\) where \(E\) has a computable inverse, there is at least one other representational theory that will mimic \((\delta, E)\). This is the theory that assigns to each stimulus a representation isomorphic to that stimulus. For this theory, the mapping \(f = E^{-1}\) satisfies the condition of preservation of internal distinctions. Since \(E\) need only be restricted to primitive recursive to have a computable inverse, this means virtually any theory, not just imagery and propositional, can be mimicked.

A final feature to note is that if the mapping \(f\) from \(\delta\) to \(\delta^*\) is a function (i.e., its

---

6 The class of primitive recursive functions is very large but is a subset of the recursive functions that can be computed on a general-purpose computing device equivalent to a Turing machine. One way of thinking of the difference between primitive recursive functions and all computable functions is that it is possible to place time bounds on the amount of computation required to perform a primitive recursive function, whereas some time bounds do not always exist for more general recursive functions:

The class of primitive recursive functions is quite powerful nonetheless. It is difficult to imagine that any process computed by the brain is more powerful than primitive recursive. Therefore, under very minimal assumptions about its nature, \(f\) will have a computable inverse.

7 If \(E\) or \(E^*\) map more than one input into the same value \(I\) or \(I^*\) the inverse of \(E\) or \(E^*\) can map \(I\) or \(I^*\) into any of these inputs. This use of inverse may not be totally conventional but is what is needed for the argument.
inverse is one to one), then there is a potential for mutual mimicry. That is, theories using \( (\delta^*, E^*) \) can be mimicked by theories using \( (\delta, E) \) as well as the converse. I would argue that this mutual mimicry holds between dual-code theory and propositional theory because it seems that one can define a one-to-one function from a dual-code representation to a propositionalization of it.

**Proof of Behavioral Mimicry**

It remains to be established that there exists the potential for mimicry given a one-to-one map \( f \). I will show something stronger— not only is it possible to mimic the observable behavior of \( M \) using \( (\delta, E) \) with a model \( M^* \) using \( (\delta^*, E^*) \), but it is possible to produce a set of internal operations on the representations in \( \delta^* \) isomorphic to the internal operations on \( \delta \). Figure 3b illustrates the point. This is important. Since the mimicking model goes through the exact same steps, not only will it reproduce the same behavior, it will reproduce the same time relationships. That is, it will mimic such things as reaction times.

After all these preliminaries, it only takes a few brief remarks to display processes \( E^* \), \( T^* \), and \( D^* \) that will mimic \( E, T \), and \( D \) just as Figure 3 illustrates. If the encoding operation \( E \) maps the stimulus \( S \) into \( I \), then \( E^* \) maps \( S \) into \( f(I) \). By its construction, the map \( f \) is guaranteed to convey on \( E^* \) this mimicry of \( E \). It is also possible to construct the elements \( T^* \) and \( D^* \) to achieve mimicry of \( T \) and \( D \). That is, if \( T \) maps \( I \) into \( I' \), \( T^* \) will map \( f(I) \) into \( f(I') \); and if \( D \) maps \( I \) into \( R \), \( D^* \) will map \( f(I) \) into \( R \). One \( T^* \) to achieve this can be constructed as \( f \cdot T \cdot f^{-1} \), where this denotes the operation of \( T \) on the internal representation, then the operation of \( T \), and then the operation of \( f \). Since \( f \cdot f^{-1} = 1 \) and \( T \) are computable, so is \( T^* \). If \( T(I) = I' \), then \( T^*(f(I)) = f(T \cdot f^{-1}(f(I))) = f(T(I)) = f(I') \). That is, if \( T \) maps \( I \) into \( I' \), \( T^* \) will map \( f(I) \) into \( f(I') \).

Similarly, we can designate \( D^* \) as \( D \cdot f^{-1} \), \( D^* \) will also be computable.

One might object to the complicated specification of \( T^* \) as \( f \cdot T \cdot f^{-1} \) and of \( D^* \) as \( D \cdot f^{-1} \). However, it does not follow that these are the simplest specifications of \( T^* \) or \( D^* \). This demonstration has the character of an existence proof. I have shown that there exists at least one process pair \( (T^*, D^*) \) that will combine with the representation \( (\delta^*, E^*) \) to yield mimicry of \( M \). For all we know, there is a \( (T^*, D^*) \) that is much simpler than \( (T, D) \) from \( M \). This argument establishes nothing one way or the other about the relative parsimony of the two simplest models for representations \( \delta \) and \( \delta^* \). The point of this section is to establish that there are no purely behavioral criteria for distinguishing between the two theories. Certainly, if it could be shown for some \( T \) that its simplest \( T^* \) mimics was \( f \cdot T \cdot f^{-1} \), there would be grounds for rejecting \( T^* \) on parsimony. However, in actual examples such as the forthcoming illustration, there have always turned out to be simpler \( T^* \) specifications. A more thorough evaluation of parsimony, an unfortunately subjective concept, comes later in the article.

A similar remark can be made about time relationships. To guarantee that \( M^* \) can mimic \( M \), we must assume that the time for each step of \( T^* \) in Figure 3 is identical to the time for the corresponding step of \( T \). For this to be so and because we are only assured of one specification of \( T^* \) as \( f \cdot T \cdot f^{-1} \), we have to assume that the time for \( M^* \) to perform the sequence \( f \cdot T \cdot f^{-1} \) can be made identical to the time for \( M \) to perform \( T \). This can be done by assuming that \( M^* \) can compute \( T \) faster than \( M \) and also can rapidly compute \( f \) and \( f^{-1} \). Thus, by speeding up the operation of \( M^* \), we can always place its operations in time step with \( M \). We might be prevented from such speedup proposals if we had adequate knowledge about possible physiological implementation, but we do not. However, as it seems we can usually construct \( T^* \) more simply than its formally guaranteed specification, the need for such speedup proposals does not seem to arise in practice.

In concluding this argument, I refer the reader back to Figures 3a and 3b, which capture its essence. The figure shows how there can be an isomorphism between the operations on two distinct representations in \( M \) and \( M^* \). Some people have the feeling that if the operations are isomorphic, the distinction between \( M \) and \( M^* \) becomes less I would like to e: It embodies an important fact: At a useful level need be no difference in items that use a wide ran In particular, there need be no behavior of a systemic versus a pictorial.
or $D^*$. This matter of an existent there exists a $D^*$ that will distinguish ($g^*$, $E^*$) from all we know, much simpler than current establishes about the relatively simplest models $M^*$. The point that there are for distinguishing. Certainly, if it is its simplest could be ground yet. However, in the forthcoming it turned out to be more thorough unfortunately in the article. We need about time that $M^*$ can at the time for identical to the step of $T$. For if we are only as $T^*$ as $f \cdot T \cdot f^{-1}$, time to order for $M^*$ to be made to perform $T^*$ and $M^*$ can also rapidly speeding up the replace its opposite. We might be impossible physically, not. More constructively guaranteed speedup procedures.

I refer the the, which can only the how there are operations $M$ and $M^*$, that if the opposite distinction between $M$ and $M^*$ becomes almost meaningless. I would like to endorse this attitude. It embodies an important claim of this article: At a useful level of abstraction, there need be no difference in the behavior of systems that use a wide range of representations. In particular, there need be no difference in the behavior of a system that uses a propositional versus a pictorial representation.

Another counterargument to this argument about representational indeterminacy is that it misses an important distinction. While it may be true that a propositional representation can be modified to accommodate data predicted by a dual-code theory (or vice versa), the counterargument goes, it is still the case that one theory predicted the result and the other theory was modified. Some points need to be made here. First, in practice, it is often not the case that one representation rigorously predicted the result. Rather, it is the case that we "intuit" the result as being implied by the representation.

A second, in those cases of fairly rigorous predictions, the derivation depends critically on process assumptions. Therefore, it is not evidence for the representation per se but for a representation–process pair. It is correct to take this predictive power as evidence for that representation–process pair. However, it certainly seems incorrect to take the outcome as evidence for the representation in abstract, since conjoined with other process assumptions, the representation would lead to very different predictions.

Finally, it is wrong to conclude from a theory's failure to predict a result that the theory needs to do anything more than appropriately modify his theory. Any theory proposed today must be wrong in many aspects. Any worthwhile theory is based on and accounts for a good range of data. Therefore, the appropriate response to inevitable discrepancy usually should be modification of the theory not abandonment. The appropriate move is to abandon a theory rather than modify when an alternate theory is developed with more parsimoniously accounts for data. The parsimony disadvantage of a theory may arise from modifications to account for discrepant results. However, such modifications also may not leave the theory at any parsimony disadvantage.

A natural reaction to this argument is to view it as "so much blind formalism." But considered in the light of the prior discussion about attempts to discriminate between propositional and imagery representation, the argument seems more compelling. That discussion contained frequent evidence for the central claim of this formal argument: Differences in representation can be compensated by differences in the processes that operate on these representations. I would also like to apply this formal argument to the following specific case.

**Example of Equivalence**

This argument has been made in quite abstract terms. The abstractness was necessary to establish the generality of the point about nonidentifiability. However, it is hard to recognize the force of the argument because of its abstractness and also because of the importation of the terminology of recursive functions and the need to create such formal constructions as $T = f \cdot T \cdot f^{-3}$. Therefore, I will give an example of how this result can be mapped into the reality of a contemporary psychology experiment.

---

8 The point of this formal analysis was to show that one cannot decide issues of representation in the abstract. That one must also specify the processes that operate on the representation. It might be asked whether this point can be turned around. Is it possible to decide issues of process in the abstract? Could we get two very different processes to mimic each other by choice of structure in the way it was shown that two structures could mimic each other by choice of process? The situation is not exactly symmetrical in that it is not possible to assign a structure for a process in the way it was possible to assign a process for a structure. A process must specify the structure of its input and output. Nonetheless, there probably is considerable indeterminacy in process assumptions. We can have two processes mimic each other by changing the processes with which they interact. For instance, we can get different processes $T$ and $T^*$ to mimic each other by changing the encoding processes $E$ and $E^*$ and the decoding processes $D$ and $D^*$. Thus, it may not be possible to test one process without specifying the processes with which it interacts.
The task I have chosen is rotation of letters as studied by Cooper and Shepard (1973). I will sketch out an imagery model (representation plus process) of this phenomenon and a propositional representation, then use the above analysis to construct a process that, operating on the propositional representation, will mimic the behavior of the imagery model. I will assume that the imagery representation $\mathbf{s}$ of a letter is a two-dimensional matrix encoding in terms of ones and zeros whether particular squares are filled in. This is basically the stimulus display presented. Therefore, the encoding process for an imagery model is one of identity, that is, $E = I$. We assume that an upright representation of the letter is stored in memory. A test object is presented and encoded into its matrix form, although in the Cooper and Shepard paradigm, the presented object is usually not in an upright form. The subject must decide whether the presented object matches a letter or is a mirror image of a letter (e.g., R vs. R). We assume that the matrix representation of the presented object is rotated until it is in an upright form. This corresponds to $T$, the process of internal transformation. It is easier to tighten our application of the formalism if we assume that the rotation proceeds in discrete steps (as is done by Kosslyn & Schwartz, 1977, in their simulation).

Figure 4. The letter $R$ whose propositional coding appears in Table 1. (The Ps are critical points in the encoding of $R$.)

So, in each discrete time unit, matrix $M_1$ is transformed into matrix $M_2$, where these two are related by a small angle of rotation around an axis through the center of the matrix. We can denote this as $M_2 = T(M_1)$. It should be obvious how to spell out the operation $T$ (e.g., as a computer program). The decoding operation applies when the rotated matrix is in upright orientation. There is a technical difficulty in deciding when the matrix is upright, but I assume as others have that there are features that will allow one to decide that a character is upright before one has recognized it. The recognition operation consists of computing a complete match between the rotated matrix and the stored matrix. This is measured by the number of correspondences in the two matrices. If the stimuli offer a close enough match, a positive response is emitted, and if not, a negative response is emitted. Let $0(M_1, M_2)$ be some scaling of the distance between the filled cells of $M_1$ and the closest filled cells of $M_2$. Then, we have $D$ defined with respect to the prototype letter matrix $M_p$:

$$D(M) = \text{YES} \quad \text{iff} \quad 0(M_1, M_p) \geq C$$

$$= \text{NO} \quad \text{iff} \quad 0(M_1, M_p) < C.$$

The above may or may not correspond to one's favored interpretation of an imagery model for this task. For the present purposes of illustration, it is irrelevant whether it does.

Let us now develop a propositional representation for letters. Figure 4 represents the letter $R$ in standard orientation, and Table 1 gives its propositional code representation. The main proposition, compose($R$, $S_1$, $S_2$, $S_3$, $S_4$, $S_5$, $S_6$) defines $R$ in terms of its component line segments. Other propositions define these line segments according to their shape and terminal points. Others give the length of the segments. Others give the angles between straight lines. Of particular interest are the proposition $\text{AXIS}(R, P_3)$, which identifies $P_3$ as giving the axis of orientation for the letter $R$, and the $\text{VECTOR-ANGLE}$ propositions, which identify the locus of all the crucial points relative to this axis.

To complete our propositional representation, we need to specify $E^*$, the encoding process. This is the operation by which one goes from the external stimulus to a propositional encoding. This process is basically one for which available alogrithms could be simplified. For instance, one could make use of the technique described in the text and the diagram, and then refine that process further, as in the text, in such a way that the result would approach the form of the final analysis. The important point about
from the external stimulus (e.g., a matrix) to a propositional encoding in terms of lines. This process is basically one of picture parsing, for which available algorithms do exist (e.g., Duda & Hart, 1973). Assuming clean displays and a small repertoire of line types, this would be a simple algorithm. The algorithm would become more difficult as the displays deviated from these assumptions. One remark about this picture parser: It could not recognize the stimulus configurations before parsed. Therefore, if Figure 4 were an encoding of a stimulus array, the R in the propositions would be replaced by an arbitrary term or node. The above may or may not correspond to one’s favorite propositional model for letter representation. As with the imagery model, this does not matter for purposes of illustration. No endorsement should be inferred for either model.

Now we have to proceed with the construction of the mimicking processes $T^*$ and $D^*$ for the propositional model. The first thing to note is that we can create the requisite map $f$ between $\theta$ and $\theta^*$ as $E^* \cdot E^{-1}$. Since $E$ was the identity transformation, we have $f = E^*$. Its inverse $f^{-1}$ will be $E^{*^{-1}}$ or a transformation that goes from a propositional encoding like that in Figure 4 to a matrix encoding. Such a function is clearly computable. The other condition required for preservation of internal distinctions is that $E(S) = f^{-1}(E^*(S))$. Since $E$ is the identity and $f^{-1}$ is $E^{*^{-1}}$, this can be rewritten as $S = E^{*^{-1}}(E^*(S))$, which is clearly true.

Now let us consider construction of $T^*$, the process of internal transformation of the representation. We have from the formal results that $T^* = f \cdot T \cdot f^{-1} = E^* \cdot T \cdot E^{*^{-1}}$. However, $T^*$, as a specification of input–output, might seem quite complicated. It might seem to require applying $E^{*^{-1}}$ to map the propositional representation to its matrix form, applying $T$ to do the matrix rotation, and then applying $E^*$ to map from the matrix back to the line drawing. But consider the final effect of this—there will be a change only in the angle in the $\text{VECTOR-ANGLE}$ propositions. This change in angle of orientation could be directly computed. This example illustrates an important point about the formal analyses:

| Table 1 |
| --- | --- |
| $A$ Propositional Encoding of the Letter $R$ in Figure 4 |

| COMPOSE($R$, $S_1$, $S_2$, $S_3$) | ANGLE($S_4$, $S_5$, $90$) |
| STRAIGHT($S_6$, $P_1$, $P_2$) | ANGLE($S_6$, $S_7$, $180$) |
| STRAIGHT($S_8$, $P_3$, $P_4$) | ANGLE($S_8$, $S_9$, $117$) |
| STRAIGHT($S_10$, $P_5$, $P_6$) | AXIS($P_7$, $P_8$) |
| HALF-CIRCLE($S_11$, $P_1$, $P_2$) | VECTOR-ANGLE($P_9$, $90$) |
| LENGTH($S_12$, $10$) | VECTOR-ANGLE($P_{10}$, $63$) |
| LENGTH($S_13$, $20$) | VECTOR-ANGLE($P_{11}$, $0$) |
| LENGTH($S_14$, $10$) | VECTOR-ANGLE($P_{12}$, $270$) |
| LENGTH($S_15$, $20$) | VECTOR-ANGLE($P_{13}$, $315$) |

Note. The $P_i$s correspond to points, and the $S_i$s correspond to line segments.

One should not reason from a complex specification guaranteeing the existence of an input–output relation to the conclusion that the computation of this relation is complex. So, in conclusion, $T^*$ will compute small changes in the angle of orientation in step with the small rotations performed by $T$. To remind the reader of an earlier point: It is no more arbitrary to assume small slow changes in an orientation parameter than to assume small slow changes of rotation.

Now let us consider the construction of $D^* = D \cdot f^{-1} = D \cdot E^{*^{-1}}$. This could be interpreted as a process that transforms the propositional encoding into a matrix format and applies the decision function $D$. However, $D^*$ could be defined directly with respect to the propositional encoding. Some propositional matching algorithm (e.g., Anderson, 1976) could be used to obtain the degree of structural match between the probe and the prototype. There would also have to be a scaling of the match between the lengths and angles in the prototype and the probe. The only difference between a letter and its mirror image, once rotated into congruence, would be the value of the vector-angles. Thus, we can define a function $0^*$ mapping the disparity between $R_p$, the prototype representation, and $R$, a particular representation. We can define $D^*$ as

$$D^*(R) = \text{YES} \text{ iff } 0^*(R_p, R) \geq C^*$$

$$= \text{NO} \text{ iff } 0^*(R_p, R) < C^*$$
For this application, 0* need not be constructed to produce an identical metric on differences between stimuli as 0 produces. The prototypes and their mirror images will be quite different by either 0* and 0, and hence, D* and D will agree as to which are same or different. If the task were to discriminate among very similar stimuli, care would have to be taken to get 0 and 0* to correspond in the metric they imposed. That they can be made to do correspond is a consequence of the general result on equivalence.

This example illustrates how one can generate a propositional model to mimic an imaginal model. An example of the reverse could also be given. I hope the reader will agree that the propositional model produced by this algorithm is plausible relative to the plausibility of the imaginal model from which it was derived. By my own subjective estimate, it seems more plausible.

Possible Responses to Nonidentifiability

There are at least two justifiable responses to this result about nonidentifiability. One is to simply try to develop some model, imaginal or propositional, that is compatible with the data and not worry about distinguishing it from all possible models with other representations. The fact that there is no explicit model, imaginal or propositional, that even comes close to accounting rigorously for the available data should indicate that this is a substantial goal in itself. The second approach is to look for other criteria besides behavioral data for selecting among representations. Personally, I find myself leaning toward the first response. I would like to state a defense of the first response and then turn to a detailed discussion of the second.

There are a number of ways to proceed in research. One is to proceed with a strongly empirical bent, with little guidance from a general theoretical framework. The hope in this approach is that a theory will somehow emerge from the empirical work. Sometimes results on nonidentifiability are used to defend this empirical orientation over a more theoretical orientation. They should not be so used. There are two types of theoretical approaches, and nonidentifiability results only argue against one.

The theoretical approach damaged by nonidentifiability results is that which attempts to discover the “true” theory. This cannot be done if we cannot decide on issues like propositional versus imaginal or serial versus parallel (Townsend, 1974). However, one can proceed in a less ambitious theoretical direction. This is to formulate a more-or-less complete model without a commitment to discriminate it from all other possible models. Nonidentifiability results imply that there are other models that will generate the same predictions. However, the important fact is that many models will not generate the same predictions. A non-unique model may be perfectly capable of being tested and proven wrong. In proving it wrong, one would also be proving wrong all of its equivalent models. It is no mean feat to come up with a model capable of accounting for the existing range of empirical facts. I think it is a fair statement that no current model handles the existing range of results on imagery. Producing such a model seems a more worthwhile endeavor than deciding among the grand contrasts such as imaginal versus propositional.

Other Criteria for Identifiability

Besides Behavioral Data

The argument to this point is that behavioral data do not provide a basis for deciding between imaginal and propositional theories. The range of possible theories is so great that for any imaginal theory, there will be a mimicking propositional theory and vice versa. Past work that has claimed to decide between propositional and imaginal theories rested on implicit “ground rules” about what were the acceptable variations in these theories. However, to date, there are no compelling reasons to accept these ground rules. I have argued that by violating these rules in quite acceptable ways, one could get one type of structure to account for data claimed to be uniquely explainable by another type of structure. Indeed, the representational issue has proven slippery because the ground rules change over time and from researcher to researcher. If it
Physiological Criteria

An obvious additional constraint comes from physiological data. If we could open the brain and observe it operating on pictures or on propositions, it seems that the issue would be settled. There are two problems with this solution, one serious and one not but both worth considering. The nonserious objection is to argue that observations of brain functions have the same problem of interpretation that behavioral data do. Suppose, not just to be bizarre, that we observed an $m \times n$ grid of data encoded on the brain’s surface and that this corresponded to a picture of an object. This observation is a datum that a theory must account for. It would be possible to attribute that datum to some source other than the fact that such a grid was actually neurally encoded. That is, reports of neural observation can be doubted just as introspective reports are doubted. However, such a level of skepticism is clearly unacceptable, whereas the skepticism about the introspective reports is acceptable. It is of interest to note that direct observation always has had a privileged status as a means of determining the state of an object, even though it must be handled with some caution (remember, e.g., the stick that appears crooked in the water). Unless there is good reason for believing otherwise, the working assumption of all men, including scientists, is that there is a one-to-one correspondence between the structure of a visual percept and the structure of the object perceived. The problem with behavioral data arising from introspection is that introspection is not direct observation of the objects under consideration.

The serious reason for challenging physiological data is that they do not provide anything like direct observation of the mental objects. Use of it tends to require more perilous chains of inferences than use of behavioral data. No neuroscientist has found anything like direct evidence for propositions or images. This is not to say that such data are impossible to achieve in principle, only that it is unlikely. This is also not to say that the more indirect data do not provide constraints on a psychological theory. (See Fodor, 1975, for a discussion of the difficulty in reducing the concepts of cognitive psychology to physiological concepts.) However, such indirect data cannot decide between image or proposition.

The principal physiological data cited for an imaginal theory concern hemispheric asymmetry in information processing. It is claimed that the right hemisphere is better adapted for spatial tasks and that the left hemisphere is better adapted for linguistic and analytic tasks. Part of the evidence for this hemispheric differentiation comes from experiments showing selective loss of abilities with damage to one of the two hemispheres (Blakemore, Iversen, & Zangwill, 1972; Kimura, 1963; Meier & French, 1965). Evidence for specificity of function is also found in split-brain patients (Gazzaniga & Sperry, 1967). Also, behavioral experiments (e.g., Giffen, Bradshaw, & Nettleton, 1972; Klatsky & Atkinson, 1971) have shown differences in the speed with which tasks can be performed that are presented to the right visual field (and hence directly to the left hemisphere) versus those presented to the left visual field (and hence directly to the right hemisphere). Tasks that involve an important verbal component are performed better when presented to the right visual field, while visual tasks do better in the left visual field.

These studies on hemispheric specialization really provide very little evidence on the form of information representation. One could propose that all information has a propositional form, but that propositions encoding visual information are stored in the right hemisphere and propositions encoding verbal in the left. Another possibility, suggested by Anderson and Bower (1973), is that rather than having the data differentially stored, one could have procedures differentially stored. That is, procedures for performing verbal tasks would be in the left hemisphere and procedures for
spatial tasks in the right hemisphere. Both types of procedures could take propositional information as their data.

A recent experiment by Patterson and Bradshaw (1975) is particularly interesting. They found evidence that the right hemisphere performs gestalt, templatelike operations on visual stimuli, whereas the left hemisphere is responsible for more analytic operations on visual stimuli. In an experiment that involved easy discriminations between a test face and a memorized face, they found a left visual field advantage. A right visual field advantage was found in a task requiring a difficult discrimination. In both cases, the memorized faces were learned under identical procedures. Consequently, it does not so much seem that there is differential storage of visual information in the right hemisphere as that there is differential ability to perform certain types of operations. If so, the data on hemispheric asymmetries lose all ability to discriminate among types of representation.

**Parsimony and Plausibility**

I would like to put together the criteria of parsimony and plausibility because I believe that plausibility is just an extension of the parsimony—the parsimony of the theory when integrated with our general sense about nature and human cognitive functioning. These criteria have some promise of discriminating among theories of representation. As I will argue, it may prove possible to decide on grounds of parsimony and plausibility whether there are two distinct representations, one typically used for pictorial information and one typically used for verbal information, or whether there is just one representation for all information. That is, I think we may be able to decide between a dual-code model versus a single-code abstract model. However, I am not optimistic that parsimony and plausibility will be of much help in establishing the nature of the representation. That is, even if we can decide in favor of a dual-code model, there seems little hope for deciding that one code is verbal and one pictorial. Similarly, if we decide for an abstract code, I see little hope of establishing that the code is propositional. So perhaps we might be able to establish whether a distinction should be made between codes but not be able to identify the character of the internal codes.

Propositional representations offer a potential advantage in terms of the parsimony of the theories that are formulated in terms of them. Because both verbal and pictorial information are abstracted into a common representation, it is only necessary to propose one set of psychological laws governing the processing of that information. This potential parsimony of propositional theories also offers a means for their rejection on the basis of an extensive research program. If it can be shown that different laws govern the processing of information depending on its source (verbal or visual), this would be evidence against a propositional theory. Any single piece of evidence of this sort would not be devastating for a propositional theory. The earlier theorem guarantees that a propositional theory can always be made to account for such perturbations. For instance, one could suppose that the processes that operated on a propositional representation depended on what was encoded—verbal or visual information. However, any extensive use of such clearly ad hoc explanations would pose a fatal strain on the propositional theory in terms of its plausibility and parsimony.

There are some data relevant to whether similar laws hold for pictorial and verbal information. The generation of research that attempted to show superior memory for pictorial information exemplifies this logic. However, as noted earlier, there is now reason for doubting the force of this work. A current candidate that is more promising is the work of Cooper and others discussed earlier, which indicates that there may be different complexity functions describing matching of verbal and nonverbal information. If this research were interpreted as Cooper be a serious strain on the parsimony of the propositional single abstract trace.

On the other side of the argument my working hypothesis is that visual versus Anderson & Paulson, there is the interference apply to the verbal descriptions vs. absolute size of the interfere quite similar. It is as is and pictorial information with each other. Similar reported by Pezdek (Not there have been frequent interference of presenting in the same modalities (Atwood, Salthouse, 1974, 1975; Bower, 1973). However, interpreted as resulting from the fatality of the content in condition rather than anything per se. For instance coding of two arbitrary modality may overlap color, shape, and size; not be the potential for two objects in different sible explanation was Pomerantz (1977). Of fact that these modalities over short retention in peripheral perceptual just are likely to be lost retardations provide the content is presented in the is likely to overlap.

A question that has considerable research is whether information have retention characteristic (a review). It has been held that subject strat respect to rehearsal, a rather extent any such.

1. I need a simulation on other issues. It Newell, 1973) that
REPRESENTATIONS FOR MENTAL IMAGERY

were interpreted as Cooper advocates, it would be a serious strain on the plausibility–parsimony of the propositional model with its single abstract trace.

On the other side of the fence, I can report my work looking for interference phenomena with visual versus verbal material (Anderson & Paulson, 1978). Here, evidence is found that the same principles of interference apply to the two sets of material (verbal descriptions vs. faces) and that the absolute size of the interference effects may be quite similar. It is also found that verbal and pictorial information mutually interfere with each other. Similar results have been reported by Pezdek (Note 6). It is true that there have been frequent reports of greater interference of presenting two sets of information in the same modality rather than different modalities (Atwood, 1971; Kroll, 1975; Salthouse, 1974, 1975; but see Anderson & Bower, 1973). However, these results can be interpreted as resulting from the greater similarity of the content in the same-modality condition rather than use of the same modality per se. For instance, a propositional encoding of two arbitrary objects in the visual modality may overlap in such features as color, shape, and size; whereas there would not be the potential for such overlap between two objects in different modalities. This possible explanation was noted by Kosslyn and Pomerantz (1977). Of relevance here is the fact that these modality effects are strongest over short retention intervals. Encodings of peripheral perceptual properties of a stimulus are likely to be lost rapidly. Perceptual properties provide the content on which two stimuli presented in the same modality are most likely to overlap.

A question that has been subject to considerable research is whether visual and auditory information have the same short-term retention characteristics (see Kroll, 1975, for a review). It has become apparent in this field that subject strategies, particularly with respect to rehearsal, are sufficiently complex to prevent any simple answer to this question. I fear a similar fate may await research on other issues. It has been argued (e.g., Newell, 1973) that the human cognition is so complex and interactive that it is not possible to address simple issues of the system, that it is necessary to build and test complex models that embody a combination of many assumptions. It is fear of this possibility that causes me to doubt whether plausibility and parsimony will yield any answer to questions about the underlying representation.

Efficiency and Optimality

It seems a reasonable assumption that the human system processes information in a way to maximize the efficiency and optimality of its performance. This means that a constraint on any theory is that it not propose that the system is processing information inefficiently. It unfortunately is not always a trivial matter to decide how the efficiency of the system should be measured in absence of detailed knowledge of its physiological implementation. Anderson and Kline (Note 4) proposed that the efficiency of a system can be measured in terms of the efficiency of its computer simulation. If some such proposal were accepted, it would be possible to place considerable constraint on a theory. However, it is unlikely that there is going to be general consensus in the field about such a definition of efficiency. Nonetheless, it seems possible to apply very general (non-implementation-specific) notions of efficiency to impose some constraints on theories of mental representation.

One very general consideration leads to an interesting conclusion about mental representations. It is the case that well-designed systems tend to have special representations for the kinds of information they have to process frequently. These representations are designed to facilitate the kind of computations useful for this kind of information. For instance, we know from physiological evidence that visual and auditory information are given very different encodings at initial neural levels.

Another good example comes from an advanced computer language like INTERLISP (Teitelman, 1976). INTERLISP has list structures that are useful for encoding symbolic structures (like propositions), arrays that are useful for encoding dimensionally organized information (like pictures), and even some
string capabilities useful for encoding sequential information (like verbal input). The array and string capabilities are additions to the original LISP (McCarthy et al., 1962), which only had list structures. These additions were forced by the practical needs of doing various types of information-processing operations in LISP. It is the case that information encoded in arrays or strings can be represented in lists but at a severe cost to efficiency of processing.

This INTERLISP example leads one inescapably to a tricode theory. That is, it seems clear that the human must process three kinds of information: visual–spatial, verbal–sequential, and abstract–propositional. The INTERLISP example reinforces a belief that the kinds of information representations optimal for these three domains are different. It would seem that there would be a strong survival advantage pushing in the direction of these three separate codes with the potential for intertranslation among them. Personally, I find this a priori argument for a tricode theory quite compelling. It is interesting, however, that it leads to the opposite a priori conclusion than parsimony. On the a priori parsimony criterion, we gave the advantage to a pure propositional theory.

Consideration of efficiency points in the direction of useful empirical and theoretical studies. Consider the complexity results reported by Gushko and Cooper (in press). Their data would seem to indicate that matching of verbal but not of pictorial material is affected by complexity. This is a study aimed right at the question of efficiency of operating on different types of representation. It is perfectly possible to mimic these results with a single propositional representation given the earlier formal argument. However, to do so would require proposing that when matching verbal information, the system holds back on using its potential for unlimited parallel processing. Besides its already mentioned lack of parsimony, this assumption seems outrageously inefficient.

Even if spatial processing were not unlimited-capacity parallel, it seems possible to establish that a propositional model would have to be made inefficient to mimic a dual-code model, and a dual-code model would have to be made inefficient to mimic a propositional model. Suppose it were the case, as it is in the INTERLISP analogy, that certain computations could be made more rapidly or spatially organized information. We could mimic these computations with a propositional representation but this would require the propositional computations to proceed much faster when processing spatial information than nonspatial information. Similarly, we could mimic computations (e.g., abstract inferences) appropriate for propositional operations with operations on spatial representations. However, this might require the spatial operations to occur faster on abstract than on spatial information. Either way, the parameters describing rate of information processing would not be invariant across material. We would be forced to propose that sometimes operations were not performed as rapidly as they could be.

These considerations lead to recommendations for more research of the Cooper variety, where one looks at processes on different information, trying to argue that the processes cannot be performed by the same mechanisms proceeding at the same rate on the same information representations. Rather, there are specialized representations with specialized processes for special types of information. Whether in fact this will be the conclusion of such research is still up in the air. It may turn out that there really are remarkable invariances in processing rates across types of information.

It is worth noting that if efficiency considerations do indicate the existence of multiple representations, they will do little to indicate the character of the differences among the representations. As noted in this article's discussion of Kosslyn and Pomerantz (1977) on this point, we cannot decide whether a particular representation is efficient without knowing the kinds of processes that will operate on it. However, we may be able to use efficiency considerations to decide that a distinction among representations needs to be made. Therefore, the conclusion on this score is similar to that from parsimony: It may be possible to decide that there are different rep-
ODE model would to mimic a propo-
were the case, as
tion, that certain
more rapidly on
vation. We could
with a propo-
this would require
ations to proceed
ig spatial informa-
imation. Similarly,
ons (e.g., abstract
propositional op-
spatial representa-
tion) require the spatial
on abstract than
ay, the para-
formation processes
nt across material.
use that sometimes
ed as rapidly as
al to recommenda-
he Cooper variety,
es on different in-
he processes
mechanisms
on the same in-
ther, there are
with special-
ion. be the conclusion
air. It may
are remarkable in-
ter types of
if efficiency con-
existence of multi-
ill do little to
differences among
ed in this article’s
omerantz (1977)
ide whether a
icient without
esses that will op-
y be able to use
 decide that a
dictions needs to be
on this score
em: It may be
ere are different rep-
resentations without deciding how they are
different.

Conclusions

The following is a summary of the conclusions
of the article:
1. The picture metaphor is the only current
explicit interpretation of the image theory.
2. The frequent criticisms of the
picture metaphor are not valid. One can have
a viable dual-code model involving picture and
verbal representations (see Figure 2).
3. The arguments for the necessity of a
propositional representation are far from
compelling. The best have to do with the utility
of such a representation for inference making.
4. The arguments for imaginal representa-
tions based on introspections, computational
considerations, empirical results, and
physiological considerations are not convincing.
5. It is not possible to decide between
imaginal and propositional representations
strictly on the basis of behavioral data.
6. The criteria of parsimony and efficiency
may allow a decision about whether there are
different types of internal representation or
just one abstract representation. This would
be based on a research program that investiga-
gated whether verbal and visual information
displayed similar properties.

There are a number of reactions to the
possibility that we may not be able to identify
the nature of the information representation.
Frequent among my colleagues is disbelief
and dismay—this would imply that cognitive
psychology is not possible. However, I think
the implication of this possibility is not that
cognitive psychology should be abandoned;
rather, it should undergo a slight change. Our
goal should be to develop some model capa-
ble of accounting for human intelligence, that
is, predicting behavior in a wide variety of
situations where human intelligence is mani-
ifested. The fact that the model may be in-
distinguishable scientifically from other quite
different models need not be a source of un-
happiness. In fact, it is possible to take com-
fort in such equivalences. If a particular model
is equivalent to many other models, we can
be more confident in its basic truth. Even if
the physical implementation directly described
in our model proves false of the human brain,
there are many other ways the model could
be true. The possibility of equivalent models
also offers computational advantages. Just as
is the case in the wave and particle models
of light, one version of the model might be
useful for certain computations and another
version of the model for other computations.
It is also worthwhile to note with respect to
this example from physics that scientists
were able to make progress without unique iden-
tifiability.

The function of science is to discover what
is the case, not to prescribe what should be
the case. If equivalence and nondeterminacy
seem to be the case, we should not be timid
about acknowledging that possibility.

Reference Notes

1. Pylyshyn, Z. W. The symbolic nature of mental
representations. In S. Kanef (Chair), Objectives
and methodologies in artificial intelligence. Sym-
posium, Canberra, Australia, May 1974.
2. Baylor, G. W. A treatise on the mind’s eye (Tech.
Rep.). Montreal: University of Montreal, Insti-
tute of Psychology, July 1971.
3. Wiseman, S. & Neisser, U. Perceptual organiza-
tion as a determinant of visual recognition memory.
Paper presented at the meeting of the Eastern
4. Anderson, J. R. & Keine, P. Design of a produc-
tion system. Paper presented at the Workshop on
Pattern-Directed Inference Systems, Honolulu, May
60–65.)
5. Winston, P. H. Learning structural descriptions
from examples (Project AI-TR-231). Cambridge:
Massachusetts Institute of Technology, Artificial
Intelligence Laboratory, 1970.
6. Peshek, K. Cross-modality integration of sentence
and picture memory. Paper presented at the meet-
ing of the Psychonomic Society, Denver, Colo-

References

Anderson, J. R. Language, memory, and thought.
Anderson, J. R., & Bower, G. H. Human associative
Anderson, J. R., & Paulson, R. Interference in mem-
ory for pictorial information. Cognitive Psychology,
Atwood, G. An experimental study of visual imagina-
tion and memory. Cognitive Psychology, 1971, 2,
290–299.


REPRESENTATIONS FOR MENTAL IMAGERY


A. Spatial comprehension in verification tasks. Perception & memory.


C. Archives of Neuropsychology, 1974.


E. Specialization of meaning in memory. 1974.


L. Inferences with nature on the papers of science. (Ed.), Visual...