Journal Title: approaches to cognition: contrasts and controversies

Volume: 
Issue: 
Month/Year: 1986
Pages: 37-78

Article Author: Palmer

Article Title: the information processing approach to cognition

Note: Palmer

12/22/2010 2:00 PM
(Please update within 24 hours)

Call #: BF311 .A64 1986

Location: evans

Not Wanted Date: 06/20/2011

Status: Faculty
Phone: 845-2503
E-mail: tya@psyc.tamu.edu

Name: Yamauchi, Takashi

Pickup at Evans

Address:
Mail Stop 4235 Dept. of Psychology, TAMU
College Station, TX 77843
Stephen E. Palmer  
*University of California, Berkeley*

Ruth Kimchi  
*University of California, San Diego*

Of the many alternative approaches available for understanding human thought, one that has dominated psychological investigation for the last two decades is information processing (IP). For better or worse, the IP approach has had an enormous impact on modern cognitive research, leaving its distinctive mark on both the kinds of theories that have been proposed and the kinds of experiments that have been performed to test them. Its influence has been so profound that some writers have argued that IP has achieved the status of a “Kuhnian paradigm” for cognitive psychology (Lachman, 1979). It is unclear whether or not this claim is really justified, but it has even been suggested that IP has achieved the status of a paradigm in cognitive psychology.

Whenever an approach so dominates a scientific field, it is important to understand—or at least to try to understand—its foundations. The assumptions that underlie its use. These must be scrutinized for their strength, plausibility, empirical support, utility, and potential limitations. Only then can one begin to see how the approach is related to others, how it has taken the field to its present state, and where it may be headed in the future. The goal of such an enterprise is essentially to provide a practical scientific approach to capture the activities and intuitions of scientists accurately and succinctly. If the practitioners agree, the approach succeeds in capturing the nature of their beliefs and their work. It provides a replacement for vague intuitions with well-defined constructs as the foundation of research.

We believe that the time has come to examine the foundations of information processing in psychology. There has been some work along these
come primarily from philosophers and computer scientists rather than from IP psychologists themselves. From IP-oriented philosophers have come formulations of a new philosophical doctrine—called functionalism—offered as a possible solution to the age-old mind/body problem (e.g., Dennett, 1978; Fodor, 1968; Putnam, 1960). From computer scientists have come related proposals that the operation of the human mind can be simulated, or perhaps even duplicated, on modern digital computers (e.g., Newell, 1980; Newell & Simon, 1972, 1976). The trouble is that many cognitive psychologists who consider themselves IP practitioners, ourselves included, find that some of the assumptions made in these arguments are too strong or of the wrong type (or both) to accurately reflect the nature of IP in psychology.

The present chapter represents our attempt to present a principled description of the IP approach as it is practiced within psychology. We try to formulate the assumptions underlying IP in terms that are based as explicitly as possible on how IP theories are constructed and tested by most IP psychologists. The accounts given by philosophers and computer scientists are just too far removed from what psychologists actually do to be certain that their views accurately reflect our own. Perhaps it will turn out that they do—although we argue otherwise—but this certainly is a matter that requires and deserves more serious attention than it has yet received. We must not assume that their views are the same as ours but rather determine whether they are or not. One of our principle aims, then, is to analyze the nature of IP with an eye toward clarifying its relation to these other proposals.

In the first half of this chapter we attempt to give a fairly broad characterization of the information-processing approach to cognition and the assumptions on which it is based. In the second half we consider the relation of IP psychology to other approaches to cognition, discussing how they agree and how they conflict.

FIVE ASSUMPTIONS OF INFORMATION PROCESSING

We take as our starting point the proposition that the intuitive basis of the IP approach is a theoretical analogy between mental activity and a program running on a computer. Whatever deeper roots IP psychology might have in communication theory, mathematical logic, or formal linguistics, the idea that the mind works like some sort of computer program is certainly the principal reason for IP's current popularity. The analogy runs roughly as follows. Certain information from the environment (the "input") is available to the mind through sensory systems, much as input information is available to a computer program through peripheral devices such as terminals, card readers, and the like. Some of this information is then manipulated in more or less complex ways by mental operations, much as a computer program manipulates information according to the rules it embodies. Among these are transform, store, and match information memories of past situations, from a combination of these. As a result of this information processing in a different form (the "output") the activity of its peripheral output may be generated. Interestingly, this general approach was generally available.

Is there anything more to IP psychology than there is and try to specify what is most structured our discussion around what held by IP psychologists and the construct theories of cognition. Our present discussion as a preview of the analysis.

1. Informational Description: Each "informational event," each operation (what it starts with), the operation (what it starts with), the output of the initial event, and the output of the final event

2. Recursive Decomposition: Each event at one level of description is decomposing it into (1) a number of informational events, and (2) the way the events specify how the information "flow".

3. Flow Continuity: All inputs must be available in the output.

4. Flow Dynamics: No output information is available and success that input.

5. Physical Embodiment: In the world of IP, being described as an information system, the system (here called representatives) are embodied in changes of state.

We do not pretend that this list is complete, but they are certainly enough. Some held set of "core beliefs." In this chapter, we make assumptions, elaborate on their implications. Unfortunately, space prevents even the equally important methodological discussion.
r scientists rather than from IP philosophers have come formulation—offered as a possible (e.g., Dennett, 1978; Fodor, have come related proposals that ed, or perhaps even duplicated, 1980; Newell & Simon, 1972, logists who consider themselves one of the assumptions made in attempts to (or both) to accurately reflect present a principled description of cognition. We try to formulate the used as explicitly as possible on most IP psychologists. The acientists are just too far removed ibelieve that their views accurately do—although we argue otherwise and deserves more serious in the others. One of our principle costs toward clarifying its relation to give a fairly broad characterizatin and the assumptions on the relation of IP psychology to lay agree and how they conflict.

ATION PROCESSING

that the intuitive basis of the IP activity and a program running olology might have in communication, the idea that the mind certainly the principal reason for ity as follows. Certain in unformable to the mind through sensory to a computer program through ers, and the like. Some of this complex ways by mental operurs information according to the rules it embodies. Among these mental operations are ones that select, transform, store, and match information arising from the present situation, from memories of past situations, from plans for future situations, or (usually) some combination of these. As a result of such operations, the mind produces information in a different form (the “output”) that is expressed as overt behavior, in much the same way that a computer program outputs information through the activity of its peripheral output devices such as terminals, line printers, and so forth. Interestingly, this general proposal about mental operations was made as early as 1943 by Kenneth Craik, long before modern digital computers were generally available.

Is there anything more to IP theory than this loose analogy? We believe that there is and try to specify what it is in the remainder of this section. We have structured our discussion around five assumptions that are almost universally held by IP psychologists and are fundamental to their theories about how to construct theories of cognition. We list them here without background or discussion as a preview of the analysis we are about to present:

1. Informational Description: Mental events can be functionally described as “informational events,” each of which consists of three parts: the input information (what it starts with), the operation performed on the input (what gets done to the input), and the output information (what it ends up with).

2. Recursive Decomposition: Any complex (i.e., nonprimitive) informational event at one level of description can be specified more fully at a lower level by decomposing it into (1) a number of components, each of which is itself an informational event, and (2) the temporal ordering relations among them that specify how the information “flows” through the system of components.

3. Flow Continuity: All input information required to perform each operation must be available in the output of the operations that flow into it.

4. Flow Dynamics: No output can be produced by an operation until its input information is available and sufficient additional time has elapsed for it to process that input.

5. Physical Embodiment: In the dynamic physical system whose behavior is being described as an informational event, information is embodied in states of the system (here called representations) and operations that use this information are embodied in changes of state (here called processes).

We do not pretend that this list exhausts the assumptions underlying IP psychology, but they are certainly among the most important ones and form a widely held set of “core beliefs.” In the following discussion we try to justify these assumptions, elaborate on their significance, and analyze at least some of their implications. Unfortunately, space does not permit any corresponding discussion of the equally important methodological and empirical aspects of IP psychology.
Assumption 1: Informational Description

Saying that the mind is like a computer program really just means that we can describe both of them in essentially the same way. We don't want to restrict our formulation of the analogy to modern digital computers because many mental processes—especially ones at the sensory and motor ends—seem to operate a lot more like "analog" machines than digital ones. Furthermore, we want to avoid using the term computational altogether because of its technical meaning in the theory of computability (cf. Johnson–Laird, 1983; Newell, 1980), which we do not want to presuppose. Therefore, we state the first assumption in terms of describing mental processes as something we call informational events, which we define as follows:

(1) Mental events can be functionally described as "informational events," each of which consists of three parts: the input information (what it starts with), the operation performed on the input (what gets done to the input), and the output information (what it ends up with).

By "mental events" we mean to include not only conscious experiences but all internal happenings that influence behavior, many of which will not produce any conscious experiences at all. For instance, native speakers seem to be able to parse sentences quite well without any conscious experience of the corresponding mental processes, and yet one will almost certainly need to suppose that there are some internal events of this sort to explain how people understand language. Naturally, we do not mean to exclude events of which people are conscious either. The phrase "mental events" in our first assumption is really just a placeholder for whatever events turn out to be necessary to account for what things people are able to do and how they do them. The concepts of "information" and "operations" also need to be defined, of course, but we postpone this lengthier discussion to a later section.

The sort of description proposed in the first assumption can be represented by a "black box" diagram like the one shown in Fig. 3.1A. It is assumed that the output is determined by the input together with the operation performed on it, and so one really needs just the first two of the three parts—input plus operation—to specify what is going on. If the operation is complicated enough that one cannot describe it directly, the operation can be expressed implicitly by specifying the mapping from input to output. When the operation so defined is ascribed to the human mind, we call this specification a "mental mapping."

Functional Description. Another aspect of the first assumption that needs to be discussed is the notion of a "functional" description. The intent here is to single out a domain of discourse for IP theories of mind that appropriately reflects the kind of accounts they offer. In this context, functional descriptions are to be distinguished from both physical and phenomenological descriptions.

We claim that it is possible to construct theories at these levels and that there are important psychological theories that primarily concern the relationships, and secondarily with how these are organized.
We don’t want to restrict our computers because many mental processes—seem to operate at least is. Furthermore, we want to avoid any of its technical meaning in the (1983: Newell, 1980), which we do the first assumption in terms of the use of informational events, which s “informational events,” each of which will not produce any conscious experiences but all of which are conscious experiences. The concept of consciousness is really just a place to account for what things the concepts of “information” and “content” but we postpone this lengthier assumption can be represented by the diagram in Fig. 3.1A. It is assumed that the operation performed on it, the three parts—input plus operation is complicated enough that n can be expressed implicitly by the operation. The operation is a complex informational operation that maps input stimulation to output behavior (A). This operation can be decomposed into an information flow diagram of several simpler component operations (B), each of which can be further decomposed into still simpler operations. The resulting recursive structure of the information flow can be represented by a hierarchical graph (D) in which the nodes at a given level correspond to the component operations of the flow diagram at that level and the links connecting them represent the decomposition relations between different levels of description.

We claim that it is possible to construct theories relevant to psychology at any of these levels and that there are important relationships among them. However, it is not clear how the middle level of functional descriptions, and secondarily how this level relates to the other two levels. Theories at the physical level are concerned with the nature of material substances and events that take place within brains. Descriptions at this level are...
given in terms of causal sequences of biochemical events involved in neural firings and interactions. Such descriptions are ultimately reducible, at least in theory, to quantal happenings among countless quarks, or whatever the most microscopic level of physical reality happens to be. Very few, if any, theories in psychology are proposed solely within the physical domain.

Theories at the phenomenological level are concerned with systematic and principled descriptions of conscious experiences in terms of other conscious experiences. Such experiential descriptions are different in kind from any sort of physical description, a fact that has led philosophers to debate issues about the relation between mind and body for centuries. Phenomenological accounts of mental events were once the primary theoretical goal within psychology, when the “introspective method” reigned supreme. Because the subjective experiences on which they are based are not publicly observable entities, however, true phenomenological theories fell from favor with the advent of behaviorism (Watson, 1913). Since then phenomenological description has played an important, but decidedly supporting, role in psychology. This is true despite the fact that many psychological theories—particularly in the field of perception—still have as their goal an account of phenomenal experience. The crucial difference is that, since Watson, psychological theories have seldom been advanced in which the account is given solely in introspective terms. Outside of psychology proper, however, important phenomenological theories have been offered by philosophers such as Heidegger (1962) and Merleau-Ponty (1962).

At this point it is appropriate to give a clear definition of the functional level. Unfortunately, we do not know of a good one—other than the one implied by the information-processing approach—and so we can only appeal to intuitive notions here. Theories at the functional level are concerned directly with neither material substances nor subjective experiences, but rather with how the brain or mind works or behaves within the context of the environment. The presumption is that this functional level of description is considerably more abstract and general than the physical level (in the sense that many physically quite different objects can have the same function), and yet this functional description is still tied to physical reality in principled ways. Most psychological theories have at least one foot firmly in this functional level. Even physiological theories, which might seem to be exclusively physical, are usually about the relation between some physical structure and its function, as when researchers attempt to say what a given cell “signals” in the environment (e.g., Hubel & Wiesel’s (1968) “edge detectors” and “bar detectors”). Prominent examples of psychological theories couched exclusively or primarily within the functional domain include Skinner’s (1953) theory of organismic behavior, Piaget’s (1950) theory of cognitive development, Freud’s (1933) theory of personality, and the many theories that came out of the “functionalist” school, as well as all modern IP theories.

Information and Operations. What distinguishes IP theories from most other functional theories in psychology is its further claim that the appropriate
type of functional description is characterized in terms of information. We cannot here present rigorous yet operational definitions of the entire theoretical approach since aspects of the whole IP enterprise have not been fully explicated. It is for this reason that the entire theoretical approach should not be taken out of context. In science often prove to be the case.

The situation is especially problematic at first glance, it might seem as though the classic Shannon’s (1948) mathematical formulation of information is not what the first assumption and not, we will see, either. To see why, one needs only a closer formulation of information: It is easy to see that expresses the reduction in entropy of a message transmitted through a channel where the mathematical details it is easy to see that expresses the reduction in entropy of a message transmitted through a channel where the mathematical details it is easy to see that expresses the reduction in entropy of a message transmitted through a channel where the mathematical details it is easy to see that expresses the reduction in entropy of a message transmitted through a channel where the mathematical details it is easy to see that expresses the reduction in entropy of a message transmitted through a channel where the mathematical details it is easy to see that expresses the reduction in entropy of a message transmitted through a channel where the mathematical details it is easy to see that expresses the reduction in entropy of a message transmitted through a channel where the mathematical details it is easy to see that expresses the reduction in entropy of a message transmitted through a channel where the mathematical details it is easy to see that expresses the reduction in entropy of a message transmitted through a channel where the mathematical details it is easy to see that expresses the reduction in entropy of a message transmitted through a channel where the mathematical details

What IP psychologists mean by the “information content of the ‘messages’ and that to which it refers.” As many commentators of information theory have also noted (e.g., Dretske, 1979), the “information content” refers only to the amount of information, not to the actual “information itself”. We have started: We know that information is informative, but we still need to determine what information content refers to. This is not the case. The term “information content” refers to the present state of affairs we are using the terms, in that they describe mental events. We would like a belief among IP psychologists that do not depend on at least some of the IP paradigm.

The key to understanding these beliefs is to understand the nature of information itself. This might be illustrated by the following examples in which two physical
type of functional description is informational; that is, mental events are to be characterized in terms of information and operations that relate information. We cannot here present rigorous yet uncontroversial definitions of information and operations because they are among the least understood and most problematic aspects of the whole IP enterprise. This may seem surprising given how heavily the entire theoretical approach seems to rest on them, but the most basic concepts in science often prove to be the most difficult to analyze explicitly.

The situation is especially peculiar with respect to "information" in that, at first glance, it might seem as though there is a readily available formalization in Shannon's (1948) mathematical theory of information. But Shannon's conception of information is not what we meant when we talked about information in the first assumption and not, we suspect, what any other IP theorist means by it either. To see why, one needs only to consider the general nature of Shannon's formulation of information: It is a unidimensional quantity (measured in "bits") that expresses the reduction in uncertainty by a receiver about a source via a message transmitted through a noisy communication channel. Even without the mathematical details it is easy to see that this is not what the first assumption is about. As many commentators on the mathematical theory of information have also noted (e.g., Dretske, 1979; Garner, 1962), Shannon's conception of "information" refers only to the amount of something as measured in bits, whereas "information itself" is about the something that gets measured in these bits. What IP psychologists mean by information is far more closely related to the content of the "messages" and the whole communication context that surrounds them than it is to the "amount of reduction in uncertainty." Shannon's theory is really about the informativeness or surprise value of messages rather than about their actual information content. This means that we are almost back to where we started: We know that information is not the same as Shannon's measure of informativeness, but we still don't know what it is or how it is related to his formulation. This is a rather embarrassing situation, but one that accurately reflects the present state of affairs in IP psychology.

As we are using the terms, information is an abstract construct in theoretical descriptions of mental events. We have used it in this way to reflect the pervasive belief among IP psychologists that IP theories are abstract, functional entities that do not depend on at least certain physical characteristics of the events being described. But we still need to define this level of abstraction and to say how it arises in the IP paradigm.

The key to understanding the abstractness of the IP level lies in the abstract nature of information itself. This abstraction can be intuitively demonstrated by examples in which two physically very different signals can carry the same

---

1We say "most" because there are some examples of functional theories based on informational descriptions that are, nevertheless, not instances of IP theories. The most notable example is James J. Gibson's (1966, 1979) theory of ecological optics in which informational descriptions play the central role, but which does not conform to the additional assumptions of IP.
information about some referent state of affairs. For example, Paul Revere and his coconspirator agreed that information about the arrival of British troops was to be carried by the number of lights hung in the church window ("one if by land and two if by sea"). However, the same information could have been equally well transmitted by the opposite arrangement ("one if by sea and two if by land"), or by the color of a single light ("red if by land and blue if by sea"), or by the number of cannon shots fired, or, indeed, by any other pair of signals that Revere could have distinguished from his vantage point on the opposite shore, provided that he and the signal-sender had agreed on the signals and their corresponding interpretations in advance. 2 There is an important sense in which all these alternative signals—one light, two lights, red light, blue light—would have been informationally equivalent under the corresponding scenarios, despite their wide variety as different physical events. The reason is that they all "stand for" or "represent" the same referent event. This informational equivalence is, by definition, more abstract and general than mere physical equivalence. It is a form of functional equivalence because it is concerned with the extent to which different events could be substituted for each other and still "work in the same way" or "cause the same outcome." It is in precisely this sense that IP descriptions are more abstract than the physical events they describe. The abstract functional nature of IP descriptions thus lies in the nature of information itself, specifically in the abstract equivalence relation defined by substitution.

Mapping Theories. At this point in our development of IP theory the whole operation performed by the human mind is conceived as a single, complex function that maps input information to output information. Informational theories of mind can be and have been proposed at this highly abstract level. We call such theories "informational mapping theories" or simply "mapping theories." 3 They specify what the mapping is in a systematic and well-defined way without claiming to specify how the mapping theories make no psychological or computational black box"; they merely specify that from inputs to outputs.

Despite this limited goal, mapping theories make no psychological or computational black box"; they merely specify that the mapping only for some modest descriptive purposes. Chomsky's (1965) transformational grammar (1957, 1975, 1977) transformational grammar. Johnson-Laird's (1983) predicate-argument psychology. Leeuwenberg's (1971, 1978) systems model. and Prazmowska's (1979) "computational theory of motion perception" theory of motion perception, together set the scope of the mapping theory. The mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating the mapping theory of naming the capital is defined simply by enumerating

3 Whereas in this example Revere and his partner had to agree explicitly on the informational correspondence between events and signals (and hence the interpretation of the signals by Revere), the same cannot literally be true for informational theories of mind, because there are no parties to do the agreeing. Instead, the role of agreement in this example must be played by some other method of arriving at a conventional interpretation: namely, some process of selection, such as evolution in cases of innately given interpretations (e.g., simple unlearned reflexes) or learning in cases of organism-acquired interpretations. In informational theories of mind, interpretations of signals carrying information about the world must be achieved through some process that provides feedback about the appropriateness of the chosen interpretation. Thus the informational correspondence is "agreed" upon by what works: which interpretation is evolutionarily successful or which one leads to desired outcomes in the learning situation. The end result is the same as if an agreement had been made between the environment and the organism. The signals carry information about the environmental events that is largely independent of the specific physical nature of those signals.

3 Mapping theories are very similar to what Marr (1982) has called computational theories. We find his label unfortunate because it strongly suggests that the mapping is accomplished by some sort of computation. If the mapping theory really makes no claim about how the mapping is accom-
For example, Paul Revere and the arrival of British troops was much window ("one if by land and two if by sea"), or by any other pair of signals that point on the opposite shore, or on the signals and their corresponding scenarios, despite the reason is that they all have informational equivalence in, or physical equivalence. It is a frame with the extent to which the term will always "work in the same way", this sense that IP describes, they describe. The abstract nature of information itself, defined by substitution.

Development of IP theory as a single, complex information. Informational theories are highly abstract level. We call or simply "mapping theorems and well-defined way agree explicitly on the informational (interpretation of the signals by Revere), mind, because there are no parties to do just be played by some other method of selection, such as evolution in mind reflexes or learning in cases of mind, interpretations of signals carry a process that provides feedback about informational correspondence is "agreed" successful or which one leads to desired one as if an agreement had been made by information about the environmental state of these signals.

has called computational theories. We mapping is accomplished by some sort in about how the mapping is accomplished without claiming to specify how this mapping is accomplished. In other words, mapping theories make no psychological claims about what goes on "inside the mental black box"; they merely describe the result of the mapping performed from inputs to outputs.

Despite this limited goal, mapping theories of large domains can be notoriously difficult to construct and are correspondingly rare. In part this is because precise specification of the mapping often requires extensive use of formal mathematical tools such as algebra, geometry, formal grammars, predicate calculus, computer programming, and the like. Another difficulty lies in the vast scope of a mapping theory for the entire range of human performance. Therefore, it is not surprising that the available examples attempt to define the input-output mapping only for some modest subdomain of human mental abilities: e.g., Chomsky’s (1965) transformational grammar theory of linguistic syntax, Horn’s (1975, 1977) differential geometric theory of perceiving shape from shading, Johnson–Laird’s (1983) predicate calculus theory of syllogistic reasoning, Leuwenberg’s (1971, 1978) symbolic coding theory of shape perception, Longuet–Higgins and Prazdnik’s (1980) theory of optical flow, Marr and Poggio’s (1979) "computational theory" of stereoscopic vision, and Ullman’s (1979) theory of motion perception, to name some of the most prominent examples. If the scope of the mapping theory is sufficiently narrow, of course, then it can be defined simply by enumerating the input–output correspondence, e.g., the mapping theory of naming the capital letters of the alphabet in a particular type font can be specified completely by a list of 26 ordered pairs. Such simple mapping theories are seldom stated explicitly because they are so intuitively obvious.

Because mapping theories do not attempt to specify how the mapping from inputs to outputs is accomplished, scientific interest in them centers on their formal rigor, predictive accuracy, and ecological validity. Unfortunately, it is often unclear just how accurately the specified mapping predicts the mental mapping because the theorist makes some form of the "competence assumption": an assumption that the theory specifies the mapping independently of any "performance" or "resource" limitations (cf. Chomsky, 1965). This strategy can be justified on the grounds that if one does not care how the mapping gets done, the theory need not—indeed, should not—take such considerations into account. However, the fact that such limitations usually do affect the observed mapping weakens the empirical implications of the mapping theory and, therefore, makes it hard to test. The problem is that it is difficult to discriminate cases in which the inaccuracies arise simply because the competence assumption is inappropriate versus cases in which they arise because of fundamental inadequacies in the mapping theory itself.

plished, calling it "computational" seems to prejudge an important issue that lies outside the mapping theory itself. We discuss these issues in more depth in the later section on "Computationalism."
Mapping theories are also often judged by criteria of ecological validity as well. The reason is that if a theory concerns only what mapping is performed, it is important that the mapping be one that applies, at least in principle, to a wide range of realistic stimulus situations. If not, it is unclear what the importance of the theory is for understanding human behavior. Unfortunately, the criteria for deciding how valid a theory is ecologically are seldom well defined, and judgments usually rest on vague intuitions about the superficial similarity between the conditions to which the theory applies and the naturally occurring conditions in the ecology of the organism. Ecological criteria are usually taken to be less critical in evaluating theories that attempt to specify how the mapping occurs. Such theories are taken to be primarily about the internal mechanisms they postulate, and so application to simple (and seemingly ecologically invalid) situations are valued for the additional scientific rigor they often permit in empirical tests.

To the preceding list of mapping theories one might wish to append some important related theoretical efforts that are similar in kind but miss one of the critical attributes of rigor, accuracy, and ecological validity. James J. Gibson's (1950, 1966, 1979) "ecological optics" theory of spatial perception was clearly in the spirit of a mapping theory by our definition—i.e., he tried to specify the mapping from proximal stimuli to perceptual responses—and had strong claims to ecological validity, but only infrequently did it achieve the requisite level of rigor to specify the actual mapping (e.g., Gibson, Olum, & Rosenblatt, 1955). Naturally, when a mapping theory is not well enough defined to determine the mapping unambiguously, its accuracy is correspondingly difficult to evaluate, and this has been a problem with many of Gibson's proposals. Other possible candidates for mapping theories come from "normative" theories, often adopted wholesale from other disciplines. They are usually rigorous but often turn out not to give a very good account of human performance, even taking the competence assumption into account. Examples of this type include mathematical logic as a theory of human reasoning (e.g., Henle, 1978; Inhelder & Piaget, 1958; see Johnson–Laird, 1983, for a critique), mathematical information theory as a theory of human performance (e.g., Fitts & Posner, 1967; Garner, 1962), and Bayesian probability theory as a theory of human inference (e.g., Edwards, 1965, 1968; see Kahneman & Tversky, 1973, for a critique).

4Given the indeterminacy mentioned previously in determining whether inaccuracies in the mappings are due to errors introduced by the competence assumption or more fundamental errors in the theory, there is a certain unavoidable element of personal judgment in deciding whether a given theory is an instance of a "good" mapping theory or belongs in this category of "inadequate" normative theories. There seems to be a correlation with the theory's age: Older mapping theories are more likely to be thought inadequate, probably just because they have been more thoroughly tested. Chomsky's (1965) transformational grammar theory, for instance, would now be thought by many to be an inadequate "normative" theory rather than a psychologically interesting mapping theory.

The relatively small number of theories that have be taken as evidence that they not propose internal mechanisms and that they are testable. The constancy of the mapping eliminates or makes ineffectual errors. Their testability derives from the empirical reality at both the input and output ends to human behavior. The importance of this is confirmed by inspecting the foregone conclusions failed as psychological theories, and how they do not produce important research findings. For example, the ways in which Bayesian probability theory (cf. Edwards, 1965) produces theories of human performance have been shown by the next test.

Notice that there is nothing about an approach. They are certainly correct in that the relationship is a weak one. Gibson, Johnson–Laird, and others are calling mapping theories, yea, mapping theories, approach. Some of his followers' answers about their opposition (e.g., Johnson–Laird, 1983). What these theorists primarily objection is that a psychological theorist should decompose theoretical descriptions of internal events that purport to be a component to this second critical assumption.

Assumption 2: Recursive Decomposition

Perhaps the most central assumption is that no event can be described more fully at a lower level of description. Decomposing it into simpler information about the assumption of recursive decomposition is a more fundamental error in the theory, there is a certain unavoidable element of personal judgment in deciding whether a given theory is an instance of a "good" mapping theory or belongs in this category of "inadequate" normative theories. There seems to be a correlation with the theory's age: Older mapping theories are more likely to be thought inadequate, probably just because they have been more thoroughly tested. Chomsky's (1965) transformational grammar theory, for instance, would now be thought by many to be an inadequate "normative" theory rather than a psychologically interesting mapping theory.

Informational theories that map (or "process models") because...
criteria of ecological validity as to what mapping is performed, it is clear that the importance of the criteria for ecological validity is seldom well defined, and judgments about them are usually taken to be less than what they actually are. Unfortunately, the criteria for ecological validity are naturally occurring conditions in the real world. The internal mechanisms they engage are usually taken to be less rigid than the actual mechanisms they engage. One might wish to append some similar in kind but miss one of the ecological validity. James J. Gibson's theory of spatial perception was clearly valid—i.e., he tried to specify the responses—and had strong claims. He defined the level of rigidity difficult to evaluate, Gibson's proposals. Other possible theories, often adopted and turn out not to be at all sufficient, even taking the competence of mathematical logic as a test (Inhelder & Piaget, 1958; see Edwards, 1978; Inhelder & Piaget, 1958; for a critique).  

The relatively small number of currently viable mapping theories should not be taken as evidence that they are unimportant. Indeed, their advocates have argued that they are absolutely critical to the enterprise of understanding the mind and that the lack of such theories is largely responsible for psychology's slow progress (cf. Chomsky, 1965; Gibson, 1966, 1979; Marr, 1982). IP psychologists are interested in them—or should be—because, even though they do not propose internal mechanisms, they strongly constrain possible mechanisms, and they are testable. The constraints arise from the fact that knowledge about the mapping eliminates or makes implausible an enormous set of possible mechanisms. Their testability derives from the fact that they are directly tied to empirical reality at both the input and output ends and so can be compared directly to human behavior. The importance of mapping theories in psychology is confirmed by inspecting the foregoing list. Even in cases where they have clearly failed as psychological theories, testing them to find out when and how they go wrong has produced important results that have strongly affected further theorizing. For example, the ways in which formal logic (cf. Johnson—Laird, 1983) and Bayesian probability theory (cf. Kahneman & Tversky, 1973) have failed as theories of human performance have led theorists to more accurate formulations. Notice that there is nothing about mapping theories that is unique to the IP approach. They are certainly consistent with it (cf. Marr, 1982), but the relationship is a weak one. Gibson, for example, was a strong advocate of what we are calling mapping theories, yet he was openly opposed to and critical of the IP approach. Some of his followers in ecological psychology are even more adamant about their opposition (e.g., Shaw & Bransford, 1977; Turvey, 1977). What these theorists primarily object to is IP's further assumption that a psychological theorist should decompose the mental mapping by postulating a structure of internal events that purport to describe how it is achieved. We now turn our attention to this second critical assumption of the IP approach.

Assumption 2: Recursive Decomposition

Perhaps the most central assumption of IP theories is that any unitary informational event can be described more fully at a more specific (or "lower") level by decomposing it into simpler informational events. We state this conjecture as the assumption of recursive decomposition:

(2) Any complex (nonprimitive) informational event at one level of description can be specified more fully at a lower level by decomposing it into (1) a number of components, each of which is itself an informational event, and (2) the temporal ordering relations among them that specify how the information "flows" through the system of components.

Informational theories that make use of this assumption are true IP theories (or "process models") because they make claims about what is inside the
"black box" of the mind. This level of theory corresponds to what Marr (1982) has called the "algorithmic" level. Unlike mapping theories, which only specify what the mental mapping is, IP theories try to specify something about how that mapping is accomplished. Such IP theories are usually represented graphically by drawings such as those shown in Fig. 3.1B and 1C. We call these IP flow diagrams. For instance, Fig. 3.1B shows a possible flow diagram for the unitary operation depicted in Fig. 3.1A. It is typical of first-level IP flow diagrams proposed as theories of mental structures involved in cognitive tasks (e.g., Atkinson & Shiffrin, 1968). Each of the components at this level has its own "primitive" description in terms of the three functional parts of an informational event: (1) its input information, (2) the operation that gets performed, and (3) the output information that results.

A flow diagram is only an incomplete representation of an IP theory, however, because it does not show precisely what input-to-output mapping is accomplished by the operations depicted in the flow diagram. What is needed to complete the theory is a "mini-mapping theory" for each of the internal informational events postulated in the diagram or, equivalently, a direct description of the operation. In some cases the mapping theory of the components is so trivial that it does not really need to be spelled out. However, in many other cases the mapping for the hypothesized components is very complex and does need to be explicitly defined, although theorists often fail to do so. The primary difference between these "mini" mapping theories for components within an IP theory and "pure" mapping theories that stand alone at the highest (undecomposed) level is that the former are only indirectly tied to empirical reality, whereas the latter are directly tied to it. For this reason the IP theorist has substantially more freedom in postulating "internal inputs" and "internal outputs" for hypothetical informational events inside the head than the pure mapping theorist has for defining the nature of "external inputs" and "external outputs."

The decomposition assumption states that operations can be broken down into a flow diagram. Because this does not exclude operations that resulted from previously decomposing a higher level operation, it implies that the decomposition of operations into flow diagrams can be performed recursively. For instance, Fig. 3.1C shows how one operation in the structure shown in Fig. 3.1B, short-term memory, might be further decomposed into a flow diagram of still simpler operations (Baddeley, 1976). Note that Fig. 3.1B could be similarly developed in our diagram.

The recursive nature of decomposition has led us to refer to as "height" ("high level" versus "low level") and the "general" versus "specific" cell just enunciated enumerates the components we use this conception to clarify programs that simulate them.

The Role of Primitives. It has been argued that the decomposition assumption could, in principle, be abandoned. We have included an implicit стори concept of a "complex (nonprimitive)" component that can be further decomposed. By itself it may not be decomposed. It is also possible for this concept to be further decomposed. We have avoided being too specific about this because there are at least two conditions under which to stop theorizing at such a "low" level: (1) the conditions under which to stop theorizing about it. Many IP psychologists prefer to work at a level that is well below the "hardware" level, theorizing at such a "low" level, the computational strategy is to define a set of simple "primitives": choose some set of simple "primitives" and stop when these are reached. We define a plausible set of primitives in terms of a universal Turing machine (Newell, 1980), and in terms of just such primitives.

---

"We purposely avoid calling them "flowcharts" because of the more restricted, technical meaning this term has in computer science as a device for specifying "transfer of control" in a serial computer. The IP flow diagrams we use here are intended to be a more general representation of how information flows over time through a system of processing components, more or less like a "time-lapse photograph" of the dynamic IP system in action. There are many difficult and technical issues involved in specifying just what these diagrams mean, but we ignore them here and rely on an intuitive understanding of them, fully realizing that this will need to be specified further at some later time.

"The distinction between "computational" and "computationally" aligned with the disciplines of computer scientists have proposed theories that are "computational," and there does seem to be a tendency for psychologists to favor the computational on
operations (Baddeley, 1976). Naturally, any of the other operations shown in Fig. 3.1B could be similarly decomposed, although we have not shown this in our diagram.

The recursive nature of decomposition implies that a tree graph can be constructed to represent the hierarchical embedding of flow diagram components, as shown in Fig. 3.1D. The vertical dimension reflects what we have been referring to as “height” (“high level” versus “low level” operations) or “specificity” (“general” versus “specific” descriptions), and the horizontal dimension simply enumerates the components that belong to each level of flow diagram. Later we use this conception to clarify the relation between IP theories and computer programs that simulate them.

**The Role of Primitives.** It is clear that without some stopping rule the decomposition assumption could, in principle, be applied recursively ad infinitum. We have included an implicit stopping rule in our second assumption through the concept of a “complex (nonprimitive) informational event,” because only these can be further decomposed. By implication, then, a primitive informational event cannot be further decomposed. We have intentionally left this stopping condition vague because there are at least two plausible and well-used strategies for defining such primitives, one based on “software” considerations and the other on “hardware” considerations. They broadly characterize two styles of theorizing that coexist relatively happily within the IP approach, a “computational” or “software” one based on choosing primitives for computational plausibility and a “physiological” or “hardware” one based on choosing primitives for their neurological plausibility.6 There is, in addition, a third strategy for dealing with the conditions under which to stop decomposing, and that is simply not to worry about it. Many IP psychologists, perhaps even most of them, are quite satisfied to work at a level that is well above any ultimate “primitives” and leave theorizing at such a “low” level to other theorists.

The computational strategy is to base the stopping rule on “software primitives”; choose some set of simple, well understood, primitive IP operations a priori and stop when these are reached. For instance, Newell and Simon (1972) define a plausible set of primitive IP operations which are equivalent in power to a universal Turing machine (Newell, 1980), and IP theories have been proposed in terms of just such primitives (e.g., J. R. Anderson, 1976, 1983; Just &

---

6The distinction between “computational” and “physiological” approaches is not precisely aligned with the disciplines of computer science and psychology, respectively. Some computer scientists have proposed theories that are essentially “physiological” in our sense (e.g., Feldman, 1981; Hinton & Anderson, 1981; Marr & Poggio, 1976, 1979). Conversely, some psychologists have proposed theories that are “computational” in our sense (e.g., J. R. Anderson, 1976, 1983). Still, there does seem to be a tendency for psychologists to favor the physiological strategy and computer scientists to favor the computational one.
Carpenter, 1977; Newell & Simon, 1972). This approach has the logical advantage of definiteness: one does indeed know just what the conditions for stopping are. However, it has the disadvantage of potential inapplicability: one does not know beforehand whether these conditions will ever be reached. In choosing a given set of primitive operations it is important to know, for example, that they are logically sufficient to capture the full range of mental capabilities one wishes to describe. Using a set that has the same computational power as a universal Turing machine is obviously a good place to start, but as we will discuss later (see the section “On Computationalism”), there is no way of knowing a priori that even this will turn out to be sufficient. If the primitives chosen are insufficient, any theory based on them will necessarily fail. Even assuming that the primitives are logically sufficient, they should also be the psychologically correct primitives. Different sets of primitives can yield quite different IP analyses of the same behavior, so choosing the wrong set in advance will produce the wrong theory.

The physiological strategy is based on “hardware primitives”: stop when the hypothesized IP operations are functional descriptions of known physical components in the device being described. This approach can be illustrated by analogy to “black box” problems in physics. The student presented with the box knows beforehand that the electrical components inside the mysterious container are things like resistors, capacitors, transistors, and the like, each of whose functional descriptions he or she supposedly knows well. When a “theory” of the contents of the box is specified as a circuit in which each hypothetical component has the functional description of a resistor, capacitor, transistor, or whatever, then the “bottom” has been reached. Naturally, the student cannot really know until the box is opened that it does not contain some far more complex circuit—such as a microcomputer that simulates the hypothetical simple circuit—but as long as the functional characteristics of the physical components are known in advance, there must be at least one correct answer in terms of the functional descriptions of these hardware primitives.

The strategy of hardware primitives is necessarily applicable because the functional definition, be applicable to describe, one is also logically sufficient to account for all the levels of the workings that is how the device actually works, as far as the device is to work, however. One is that the levels of the “lower” (more specific and detailed) primitives, the hardware primitives in this way, rather than the software primitives found in the software that programs them. If this is also true, one will unnecessarily increase the complexity of the system.

A potentially more serious problem is that we are not yet in the same position as the student, in that we do not really know what to expect in terms of the relationship to mental function. The mind is complex. However, the important structure of the nervous system such as molecular processes at the level of the complex Hebbian cell assemblies. It is known that there is currently good agreement on the physiological structure, all of which is at the level of the physical unit of processing. However, there are also other peripheral sensory systems—such as the visual system, and spatial vision (De Valois & Di Lollo, 1988)—that are known to be capable of such operations.

Still, many IP theorists do see the mind as, implicitly, in that they couch their theories among processing elements at the level of functional procedures (e.g., Feldman, 1981; Rumelhart & McClelland, 1988; Rumelhart, 1981; Palmer, 1988). Physiologically oriented theories, which are based on the physiological processes such as sensation and perception, are known. However, the corresponding notion for the processes is not entirely clear, although it is known that the corresponding notion for the processes is not entirely clear. In particular, the physiological oriented theories to explore the role of the brain in higher level cognitive processes (Rumelhart & Anderson, 1981, for a good study).

Complexity Reduction. The IP approach into well-formed flow diagrams...
The strategy of hardware primitives has much to recommend it. It is necessarily applicable because the functional description of the hardware must, by definition, be applicable to describing its behavior. By the same token, it must also be logically sufficient to account for the capabilities of the device, because that is how the device actually does it. There are several potential problems, however. One is that the level of hardware primitives may be considerably "lower" (more specific and detailed) than that of software primitives. For example, the hardware primitives in modern digital computers are much lower level than the software primitives found in the languages that people typically use to program them. If this is also true of the mind/brain, using hardware primitives will unnecessarily increase the complexity of the "bottom level" IP description.

A potentially more serious problem is that cognitive psychologists probably are not yet in the same position as a physics student with a "black box problem" in that we do not really know what the brain's critical physical units are with respect to mental function. The most obvious candidate is, of course, the neuron. However, the important structures might ultimately turn out to be much smaller, such as molecular processes at synaptic membranes, or much larger, such as complex Hebbian cell assemblies (Hebb, 1949). There are at least a few cases in which there is currently good agreement between an IP description and a known physiological structure, all of which do currently point to the neuron as the basic physical unit of processing. However, most of these examples come from fairly peripheral sensory systems—such as color vision (De Valois & De Valois, 1975) and spatial vision (De Valois & De Valois, 1980)—and the relevant unit of processing might conceivably be quite different for more complex cognitive operations.

Still, many IP theorists do seem to use a "hardware primitives" rule, at least implicitly, in that they couch their theories in terms of excitation and inhibition among processing elements more or less like individual neurons or neural pathways (e.g., Feldman, 1981; Hinton, 1981; Marr, 1982; McClelland & Rumelhart, 1981; Palmer, 1983; Palmer & Bucher, 1981; Posner, 1978; Rumelhart & McClelland, 1982). It is perhaps not too surprising that these physiologically oriented theories tend to be of relatively peripheral mental operations such as sensation and perception, processes about which at least something is known of the corresponding neural hardware. Computationally oriented theories tend to be of more complex, central processes like memory, language understanding, and problem solving, processes about which relatively little of importance is yet known with respect to the hardware involved. The principle exception to this generalization comes from recent attempts by some physiologically oriented theorists to explore the potential capabilities of quasineural devices for higher level cognitive processes like memory and categorization (see Hinton & Anderson, 1981, for a good sample).

**Complexity Reduction.** The rationale for decomposing mental operations into well-formed flow diagrams is to specify the nature of a single **complex**
operation in terms of information flow among a number of simpler operations. In principle, at least, the decomposition should “factor out” some portion of the complexity implicit in a mental operation when it is considered as a unitary informational event by making it explicit in the flow relations among a number of simpler operations. This is what we mean by “complexity reduction”: With each lower level of description the internal complexity of the component operations should decrease. Generally speaking, this reduction comes at the cost of more components and more complex flow relations among them. However, the goal is to make complexity explicit, so the net effect of decomposition is to reduce the unwanted commodity: implicit complexity. In effect, as one proceeds down the decomposition tree (see Fig. 3.1D), more and more of the complexity is accounted for by additional arrows and boxes (the part made explicit in the flow diagram) and less and less by the “mini” mapping theories of the boxes themselves (the part still implicit within the operations).

The assumption that decomposition reduces implicit complexity probably should have the same status within IP as the previous two, but it does not for several reasons. The primary problem is that being rigorous about it would require well-defined measures of IP complexity both for mapping theories (i.e., the internal nature of the unitary operations) and for IP flow diagrams (i.e., the ordering relations among these unitary operations). No such measures of either kind are currently in use or, to our knowledge, have ever been suggested. Most IP practitioners probably do believe, at least implicitly, that something like this complexity reduction assumption applies to IP theories because they feel that they understand more about what is going on after an operation has been decomposed than they did when it was a unitary, unarticulated event. This intuition depends heavily on the initial unitary operation being more complex than the several operations into which it is decomposed. Without any well-defined measures of informational complexity, however, it is difficult to tell for any given IP theory whether such beliefs are really justified.

Emergent Properties. One effect of decomposition not captured by the notion of complexity reduction is that the resulting component operations are not only quantitatively simpler than the initial one, but qualitatively different from it. For instance, what would be described as a unitary operation at a high level as a memory search operation (“look for a target, T, in list L”) is radically recast in terms of the information flow among its component operations: data retrieval (“get the next element, E, from list L”), pairwise comparison (“compute the similarity, S, of E to T”), decision (“is S greater than some critical value, C?”), and conditional control (“if so, return positive; otherwise, go to start”). Notice that although each lower level operation does something quite specific that is easily described on its own, none of them does anything like “search sequentially through a list.” Only when they are configured into an appropriate flow diagram do they, together, perform a search operation. “Search” does not really exist in any of the lower level operations; they are put together in the proper way to give a “search” appropriate to speak of a “search” property, as in the description where what is happening is a search.

Thus we see that higher level operations, that lower level decomposition of a system specified by the flow relations among its components gives rise to these properties. This is true in physical systems where we can look at systems of elements with different states, gaseous, liquid, and solid states of the same molecules, because no single molecule gaseous, liquid, or solid (cf. P. 149). These elements depend on the relations among many of their subcomponents. In phonic: The properties of the output cannot be predicted from those of its smaller physical components. We perceive of emergent properties in the mind, but no magic involved, just the combining of different ways that produce different properties.

Assumption 3: Flow Continuity

Information-processing theories, like many formal languages, are, of course, because the order of the components imposes certain important constraints on the operation of IP, concerning the syntax:

(3) All input information required to perform the output of the operations that make up the theory.

This assumption is perhaps seen as a rather obvious one. However, it is really just a corollary to the decomposition assumption. The decomposition assumption states that the input flow must be broken into other “boxes” that lead directly to output. Decomposition is not sufficient for the output unless all of the operations are “well formed,” and the theory implies the desirability of it. However, it could not actually carry out the decomposition unless it could not actually carry out the output. This requires that each operation in the IP theory must be supplemented by certain (but unspecified) constraints on all of the suboperations (described earlier) of each hypothesis. As we have seen, each hypothesis must give, at best, a rather vague,
number of simpler operations. In “factor out” some portion of the ten it is considered as a unitary flow relations among a number of “complexity reduction”: With complexity of the component operation reduction comes at the cost of alterations among them. However, the net effect of decomposition is to complexity. In effect, as one proceeds more and more of the complexity decrease (the part made explicit in the ‘i’ mapping theories of the boxes operations).

The implicit complexity probably previous two, but it does not for at being rigorous about it would ity both for mapping theories (i.e., and for IP flow diagrams (i.e., the ions). No such measures of either, have ever been suggested. Most implicitly, that something like this IP theories because they feel that after an operation has been dearticulated event. This intuition being more complex than the d. Without any well-defined means is difficult to tell for any given IP ‘d.

Composition not captured by the noting component operations are not but qualitatively different from it. A unitary operation at a high level as a ‘T, in list L’ is radically recast in component operations: data retrieval pairwise comparison (“compute the later than some critical value, C?”), etc., otherwise, go to start”). Notice something quite specific that is does anything like “search sequenconfigured into an appropriate flow operation. “Search” does not really exist in any of the lower level components individually; it emerges only when they are put together in the proper flow relations. Strictly speaking, then, it is appropriate to speak of a “search” taking place only at the higher level of description where what is happening is conceived as a unitary event.

Thus we see that higher level IP descriptions sometimes contain emergent properties that lower level descriptions do not. It is the organization of the system specified by the flow relations among the lower level components that gives rise to these properties. There is nothing mysterious in this. It is equally true in physical systems where systemic “macro” properties arise only in large systems of elements with different “micro” properties. As a simple example, the gaseous, liquid, and solid states of matter arise only in aggregations of many molecules, because no single molecule, by itself, has the properties of being gaseous, liquid, or solid (cf. Putnam, 1975; Searle, 1983); such properties depend on the relations among many molecules. Even better physical analogies can be found in complex human artifacts like stereos, automobile engines, and telephones. The properties of the object as a whole are qualitatively different from those of its smaller physical components. It seems entirely appropriate to conceive of emergent properties in IP systems in essentially the same way; there is no magic involved, just the configural interaction of different subsystems in ways that produce different properties at the systemic level.

Assumption 3: Flow Continuity

Information-processing theories postulate decompositions of mapping theories into psychologically meaningful components. Not just any decomposition will do, of course, because the ordering or “flow” of information among components imposes certain important constraints. These constitute the third assumption of IP, concerning the syntax of IP flow diagrams:

(3) All input information required to perform each operation must be available in the output of the operations that flow into it.

This assumption is perhaps so obvious that it almost goes without saying; it is really just a corollary to the decomposition assumption. In terms of flow diagrams, it states that the input for each “box” consists of the output of all the other “boxes” that lead directly to it by forward-going arrows. If this information is not sufficient for the operation to occur, then the flow diagram is not “well formed,” and the theory it represents is logically deficient in the sense that it could not actually carry out the operation it purports to describe. To determine whether such flow constraints have actually been met, the IP flow diagram of an IP theory must be supplemented with a “mini” mapping theory (in the sense described earlier) of each hypothesized operation. In reality, most IP theorists give, at best, a rather vague, verbal description of the input–output charac-
teristics of the components, and it is often hard to determine whether the flow-continuity constraint has been met from such verbal statements. As Newell and Simon (1963) have argued forcefully, this problem can be solved by supporting the theory with a computer simulation, because the program necessarily specifies a "mini" mapping theory for each operation in the flow diagram. Unfortunately, simulations are seldom actually done, and some residual amount of "handwaving" invariably remains in any verbally stated theory. (Later we discuss the relation between simulation programs and IP theories in greater depth.) Flow diagrams that meet the flow-continuity requirement, insofar as this can be determined, constitute IP theories of the unitary informational event at some lower level of description.

Assumption 4: Flow Dynamics

Some additional assumption is needed to specify the temporal properties of information flow within the system. Here we try to specify some of the most general constraints in terms of the assumption of flow dynamics:

(4) No output can be produced by an operation until its input information is available to it and sufficient time has elapsed for it to process this input.

The dynamics of information flow are particularly important in psychology because they often play a central role in empirical tests of the theory. We have attempted to capture only the constraints that (1) processing cannot begin until at least some input information is available and (2) that every operation takes some amount of time, no matter how small.8 Beyond these two notions, flow dynamics are pretty much up to the theorist and the constraints imposed by the proposed flow diagram.

The most frequently made additional assumption about the time course of processing is that each operation in the flow diagram constitutes a discrete stage (Sternberg, 1969a). In stage theories, each operation has a specific duration (plus or minus random variability) that depends on parameters determined by the input information. Before the end of this time interval the operation has no output, and at the end its output is assumed to be fully available as input to the next operation. This notion of discrete stage theories gained immense popularity with the success of Sternberg's (1966) formulation of the additive factors model and another, stage models have done so as well.

Other conceptions of flow dynamics make things more complicated by providing more flexible and realistic conceptions of the same theory. Some theories provided and proposed a "chain") as an alternative. McClelland's model of this type that he applied to each operation begins to produce output, and certainty in the results is relied upon from preceding operations. Information flow actually provided does stage theory is currently under construction. Conceptions of flow dynamics have, however, not been that helpful. Even if stage processing does turn out to be much harder to pin down (see Townsend, 1971; 1972; Townsend & Ashby, 1983) for many different versions of serial, parallel, and, depending on which additional assumptions are made predictions that are not of models, and vice versa.

This points out a problem in tests of models that may be far more general than this particular: the test produce rigorous tests of large and small (versus parallel) because pairs of models with insufficiently different predictions may well be that rigorous tests required for the classes of models than are usually provided.

Assumption 5: Physical Elements

Earlier we argued that information processing that exist in the domain of IP does not allow for the possibility of "instantaneous" processing, we presume that it would be used in preference to exponential models.
3. INFORMATION

To determine whether the flow-verbal statements. As Newell and Simon can be solved by supporting the program necessarily specifies the flow diagram. Unfortunately, the residual amount of “handwaved” theory. (Later we discuss the theories in greater depth.) Flowment, insofar as this can be determinational event at some lower

Specify the temporal properties of flow dynamics:

1. Until its input information is ready to process this input.

2. Particularly important in psychological tests of the theory. We have (1) processing cannot begin until at least every operation takes some time; these two notions, flow dynamics constraints imposed by the proposed assumption about the time course of the operation constitutes a discrete stage duration has a specific duration (plus parameters determined by the input) and the operation has no output, and available as input to the next operated immemorized immense popularity with the increasing success of Sternberg’s (1966, 1969b) work on memory scanning and his formulation of the additive factors method (Sternberg, 1969a). In one form or another, stage models have dominated IP theories of flow dynamics ever since.

Other conceptions of flow dynamics are not ruled out, of course; they just make things more complicated. Norman and Bobrow (1975) suggested that a more flexible and realistic conception of information flow was needed than stage theories provided and proposed the hypothesis of “continuously available output” as an alternative. McClelland (1979) developed a specific mathematical theory of this type that he aptly called “cascade processing.” In cascade theory, each operation begins to produce some output almost as soon as it gets some input, and certainty in the result increases over time as more and more input is received from preceding operations. Whether this more complex conception of information flow actually provides a more accurate model of IP dynamics than does stage theory is currently unclear, but the formulation of alternatives to stage conceptions of flow dynamics has been an important theoretical development. Even if stage processing does turn out to be correct, the crucial evidence will undoubtedly come from explicitly testing its assumptions against those of well-defined alternatives (e.g., Meyer, Yantis, Osman, & Smith 1984; Miller, 1982).

Assumption 5: Physical Embodiment

Among the most important issues related to flow dynamics is whether a given pair of operations are executed sequentially (serial processing) or simultaneously (parallel processing). This is specified in the information flow diagram by the obvious conventions: a “chain” of arrows from each operation to the next (serial) or several arrows diverging from a point and leading to several operations at once (parallel). Although this distinction is quite clear theoretically, it turns out to be much harder to pin down experimentally than was initially suspected (see Townsend, 1971, 1972; Townsend & Ashby, 1983). The problem is that many different versions of serial and parallel process models can be constructed and, depending on which additional assumptions are made, some serial models make predictions that are not empirically distinguishable from some parallel models, and vice versa.

This points out a problem in testing IP models, a problem that we suspect may be far more general than this particular example. It may be quite difficult to produce rigorous tests of large classes of alternative IP models (e.g., serial versus parallel) because pairs of models in the different classes make the same or insufficiently different predictions when they are examined in detail. The moral may well be that rigorous tests require comparisons between much more detailed classes of models than are usually considered.
mechanical, electronic, optical, and biochemical devices and to say something about how abstract information and operations relate to the physical reality of this material world. Clearly they need some physical "vehicles" in the dynamic real-world event that is being described in terms of information and operations. We make explicit the intuitions that information and operations are carried by some physical "medium" or "substrate" in the fifth assumption of the IP approach, that of physical embodiment (or "implementation").

(5) In the dynamic physical system whose behavior is being described as an informational event, information is carried by states of the system (here called representations) and operations that use this information are carried out by changes in state (here called processes).

Notice that as we are using the terms, representation refers to the physical system that "carries" (or "contains" or "embraces" or "instantiates") information, and processes refer to the physical events that "carry out" (or "perform" or "embodies" or "instantiates") the operations. Thus, "information" and "operations" exist in the formal domain of IP descriptions, whereas "representations" and "processes" exist in the physical domain of objects and events in the world when these are viewed as information processing. Thus, we do not mean to imply that representations and processes are merely physical objects and events, but physical objects and events under an informational and operational description.

We have not yet said anything about what makes these systemic states count as representations that carry information about some other state of affairs. This is an important metatheoretical question within the IP framework, and some work has been done explicitly on it (e.g., Bobrow, 1975; Newell, 1980; Palmer, 1978; Rumelhart & Norman, 1984). The prevailing notion is that representations are defined by (1) being used as a surrogate for some referent world and (2) preserving the abstract informational structure of that referent world. These two aspects can be easily demonstrated in how a standard road map acts as a representation of the roads, towns, and spatial layout of the geographical region it depicts. To use the map as a surrogate of the region, one needs to establish a correspondence (or mapping) from geographical objects to map objects and geographical relations to map relations. In an actual road map, this correspondence is spelled out in the "key" and the various labels attached to map-objects. If a representation is to work as a surrogate, however, it also has to be reasonably accurate. This is where preserving informational structure comes in. Structure is preserved when the truth of statements about the referent world is preserved by the truth of the corresponding statements about its representation. For instance, true statements about the lengths of roads, directions of roads, distances between cities, and so forth correspond to true statements about the corresponding map entities: the lengths of lines, directions between lines, and distances between small circles.

Together, these two aspects of surrogate of the world.

Unfortunately, space does not permit a full discussion of the issues involved in deciding whether a particular mapping is a perceptual and cognitive procedure to be represented for a more complete model of theories and to Rumelhart and Norman (1984) and specific assumptions that have been made.

Flow Diagrams versus Programs

Proposed that IP theories are well-mapped onto the mapping theories of the operational environment. This view of IP is related to the way programs are IP theories of the open system model (Newell & Simon, 1963). Superficially, because they are both in the same domain, they are definitively related, we see them as different theories are descriptions where the theory is described.

According to the present view of the theory by virtue of the fact that the mini-mapping theories of its corresponding program is associated with each program, the program a psychological theory is said to be compatible with) many different theories. In fact, programs are often hierarchically embedded flow diagrams (e.g., as shown in Fig. 3.1). The mental program which flow diagram correspondent to is the program.

The important theoretical issue with these flow diagrams is when there is a level of description where the flow of the sequence of elementary logical operations when it runs the simulation in its program and this level of description to be preserved; it is the obvious level of flow diagram to be preserved. Unfortunately, there is the flow diagram that is not preserved and operations specific to the higher-level simulationist used: Lisp, Fortran, Pascal, and so on. Even this level of description is no more automatically assumed that certain languages are more efficient at working in the mind than others (see Newell & Simon, 1972).
3. INFORMATION

Ical devices and to say something as relate to the physical reality of physical "vehicles" in the dynamic ms of information and operations. tion and operations are carried by in the fifth assumption of the IP "implementation").

behavior is being described as an a states of the system there called formation are carried out by changes

presentation refers to the physical mbodies" or "instantiates") infor- events that "carry out" (or "per- erations. Thus, "information" and P descriptions, whereas "represent- nal domain of objects and events in ation processing. Thus, we do not es are merely physical objects and re an informational and operational

makers these systemic states count at some other state of affairs. This is in the IP framework, and some work 1975; Newell, 1980; Palmer, 1978; notions are that representations are some referent world and (2) preserving referent world. These two aspects road map acts as a representation of geographical region it depicts. To use hds to establish a correspondence (or objects and geographical relations to correspondence is spelled out in the t map-objects. If a representation is to be reasonably accurate. This is where a Structure is preserved when the id is preserved by the truth of the ation. For instance, true statements hds, distances between cities, and so the corresponding map entities the nd distances between small circles.

Together, these two aspects of representation allow a model to be used as a surrogate of the world.

Unfortunately, space does not permit us to present even a small subset of the issues involved in deciding what sorts of internal representations people use in perceptual and cognitive processes. The reader is referred to Palmer’s (1978) presentation for a more complete analysis of the nature of representation in IP theories and to Rumelhart and Norman’s (1984) discussion for a survey of issues and specific assumptions that have been made in recent IP theories.

Flow Diagrams versus Programs as Psychological Theories. We have proposed that IP theories are well-formed information flow diagrams plus mini-mapping theories of the operations within them. We now want to consider how this view of IP is related to the well-known claim that running computer simulation programs are IP theories of the mental processes that they simulate (e.g., Newell & Simon, 1963). Superficially, at least, they seem to be compatible, because they are both in the same line of theoretical analogy. Although they are definitely related, we see them as distinct claims, at least in the sense that IP theories are descriptions whereas running computer programs are events to be described.

According to the present view, a running simulation program is only an IP theory by virtue of the fact that it too can be described by a flow diagram plus mini-mapping theories of its components. If there were one unique flow diagram associated with each program, then there would be no difficulty in calling the program a psychological theory. However, any program can be described by (or is compatible with) many different flow diagrams at different levels of specificity. In fact, programs are often written by constructing a sequence of hierarchically embedded flow diagrams at more and more specific levels of detail (e.g., as shown in Fig. 3.1). Therefore, important problems arise in deciding which flow diagram corresponds to the theory allegedly embodied in the program.

The important theoretical issue for the simulation theorist concerns which flow diagrams he or she takes to be psychologically meaningful. For instance, there is a level of description (i.e., a flow diagram) that corresponds to the sequence of elementary logical operations the digital computer actually executes when it runs the simulation in "machine language." Almost nobody would take this level of description to be psychologically meaningful, and yet it is the most obvious level of flow diagram to identify with "the program." At a higher level, there is the flow diagram that reflects the statement-by-statement sequence of operations specific to the higher level programming language that the simulationist used: Lisp, Fortran, Pascal, APL, or whatever. It is very unlikely that even this level of description is psychologically meaningful, although it has been claimed that certain languages are much closer to the elementary IP operations of the mind than others (see Newell & Simon, 1963).
At some still higher level of description the operations represented by the flow diagram become plausible components of a psychologically reasonable IP theory of the mind. A simulationist might well want to claim, for example, that high-level flow diagram components like searching through a network structure from several nodes simultaneously (as in "spreading activation" theories) or matching one feature list against another for similarity actually happen in the mental process simulated by the program. Even in a high-level programming language, such components usually involve large chunks of "code" that include many ad hoc details, such as the programming tricks required to mimic parallel processes on serial machines. Clearly the latter are not part of the IP theory the simulationist had in mind, whereas the former, large-scale components of the higher level flow diagram are part of the theory. Thus, decisions about which components are psychologically real determine what level of description of the program constitutes the IP theory it embodies. Because there may be higher level flow diagrams that the theorist also wants to claim are psychologically meaningful, the theoretical interpretation of the program corresponds to drawing a line across the hierarchy of flow diagrams to separate those components that are psychologically meaningful ("above" the line) from those that are implementational details ("below" the line). The reason for wanting to keep the higher level flow diagrams as part of the psychological theory, of course, is that they may well be correct even though lower level ones are wrong.

This way of viewing simulation programs also makes clear one of their principle drawbacks as cognitive theories: The constraints of writing a complete, runnable program require the theorist to specify much more than he or she actually needs to specify for the IP theory. This includes everything in the hierarchy of flow diagrams that exists "below the line." As anyone who has ever written a simulation program soon realizes, enormous amounts of time and energy must be expended in figuring out how to get the machine to do what is required, even though many of the details of how this gets done are not really part of the theory. The payoff for this additional work is the assurance that one's theory is actually capable of performing the task simulated and that one really has mini-mapping theories of each component operation in the flow diagram.

In summary, we are arguing that a simulation program implies an IP theory but is ambiguous about just what that theory is. Once the theorist has identified the level of flow diagram that separates meaningful theoretical statements from mere programming details, the theory attached to the simulation becomes clear. We view this analysis as a clarification of the program-as-theory idea rather than as a contradiction. We agree wholeheartedly with much of the spirit and motivation behind the program-as-theory movement, such as specifying vague verbal theories more precisely and allowing their adequacy to be tested rigorously (Newell & Simon, 1963, 1972). Our objection is merely that simulation theorists should be more careful than they often have been in specifying the relation between their program and the IP theory of the mental processes it simulates.

RELATIONS TO OTHER PERSPECTIVES

We stated at the outset that one of the objectives of this chapter was to clarify its relation to other perspectives. Implicitly at least some of what we have described as IP theory, we are in a position to view as different approaches to cognitive theory in general, behaviorism, and ecology. We then noted our views of cognition that are closest ("innate AI"). A synthesis, and Turing's initial intuition that IP is different from behaviorist views turns out to be correspondingly careful. IP is not so weak as to make fewer substantive claims about related proposals.

Cognitivism

It is important to be clear at the outset that, in the current literature, psychology, cognitivism refers to something different from what was assumed that behavior can only be understood in terms of "cognitive" (or "mentalistic") states, memories, images, plans, and the like. Cognitivism is a viewpoint expressed in Wason and Police (1966) and formulated vigorously by Tolman and others. Prior to this time, psychological research did not have defined, but it was one of a number of behaviorist alternative to bring this in focus.

Information processing is certainly central to many other perspectives, but its theories are based on such hypotheses that go beyond simple cognitivism in many ways. For example, are we to envisage a form that such theories should take here in terms of informational discontinuity, flow dynamics, and physical systems from other cognitive theories. Indeed, we do not ascribe to one or more of these a looser (or at least different) set of

Palmers and Kimchi (1972)
operations represented by the flow
psychologically reasonable IP theory
to claim, for example, that high-
through a network structure from
activation’ theories) or match-
actually happen in the mental
high-level programming language.
so of “code” that include many ad-
quired to mimic parallel processes
part of the IP theory the simulta-
ne-scale components of the higher
us, decisions about which compo-
level of description of the pro-
se there may be higher level flow
are psychologically meaningful,
responds to drawing a line across
se components that are psychologi-
that are implementational details
keep the higher level flow dia-
ourse, is that they may well be
is also makes clear one of their
constraints of writing a complete,
more than he or she actually
ides everything in the hierarchy of
As anyone who has ever written a
amounts of time and energy must be
bine to do what is required, even
one are not really part of the theory.
ance that one’s theory is actually
that one really has mini-mapping
flow diagram.
ulation program implies an IP theory
is. Once the theorist has identified
meaningful theoretical statements from
ed to the simulation becomes clear.
program-as-theory idea rather than
y with much of the spirit and the
ment, such as specifying vague-
ing their adequacy to be tested
our objection is merely that simul-
y often have been in specifying the
relation between their program as it runs on a computer from their psychological
theory of the mental processes it simulates.

RELATIONS TO OTHER VIEWS OF COGNITION

We stated at the outset that one of our primary goals in trying to specify IP theory
was to clarify its relationship to other views of cognition. Having now stated ex-
plicitly at least some of what we believe are the principal assumptions underlying
IP theory, we are in a position to contrast it with some other noteworthy ap-
proaches to cognitive theory in the history of psychology: cognitivism, behav-
orism, and ecologism. We then turn our attention to some more contemporary
views of cognition that are closely related to IP: computationalism (or “weak AI”),
functionalism, and Turing-machine functionalism (or “strong AI”). Our ini-
tial intuition that IP is different from and weaker than these other contempo-
rary views turns out to be correct when the underlying issues are examined
carefully. IP is not so weak as to be meaningless or unfalsifiable, but it does
make fewer substantive claims about the nature of human mentality than these
related proposals.

Cognitivism

It is important to be clear at the outset that IP psychology, as described here, is
not the same as “cognitivism,” but a specific brand of it. As the term is used in
psychology, cognitivism refers to a very broad theoretical position in which it is
assumed that behavior can only be properly understood by postulating internal
“cognitive” states such as percepts, attitudes, beliefs, goals, memories, images, plans, and the like. This explicitly cognitivist stance was
formulated vigorously by Tolman (1932) in response to the explicitly anticogni-
tive viewpoint expressed in Watsonian behaviorism (Watson, 1913, 1925).
Prior to this time, psychological theory was certainly “cognitive” in the sense
we have defined, but it was only implicitly so. It took the challenge of the
behaviorist alternative to bring the cognitivist viewpoint into focus.

Information processing is certainly an explicitly cognitivist approach in that
its theories are based on such hypothetical internal states. However, IP goes well
beyond simple cognitivism in making further assumptions about the specific
form that such theories should take. It is these further assumptions—described
here in terms of informational description, recursive decomposition, flow contin-
unity, flow dynamics, and physical embodiment—that distinguish IP theories
from other cognitive theories. Indeed, there are many cognitivists who simply do
not ascribe to one or more of these assumptions, preferring to work within a
looser (or at least different) set of theoretical constraints that are nevertheless
quite "cognitive." Historically prominent examples include Freud, Piaget, Vygotski, Wertheimer, Kohler, Koffka, Bartlett, and, of course, Tolman. Thus, IP is correctly viewed as a proper subset of cognitivism: All IP psychologists are cognitivists, but not all cognitivists are IP psychologists.

Behaviorism

The relation between IP and behaviorism is far more complex, because (1) there are several different brands of behaviorism that need to be distinguished and (2) each brand includes somewhat different proposals with which IP is in agreement on some and opposed on others. Perhaps because IP arose historically as a reaction against behaviorism, there is a tendency to see them as diametrically opposed. As is often the case, however, the new approach has much more in common with the one it supplants than is initially acknowledged. This is clearly true of the relation between IP and behaviorism.

As initially proposed by Watson (1913, 1925), the behaviorist approach broke with the then-traditional introspective approach by identifying publicly observable behavior as the central concern of psychology rather than private mental experiences. This general proposal contains at least two quite different aspects that need to be distinguished, however: methodological behaviorism and theoretical behaviorism. The methodological proposal is that because subjective experiences are not scientifically observable, behavior is the proper object of study in psychology as an objective natural science. With this IP practitioners invariably agree, at least in the sense that the data on which theoretical issues in IP are decided are objective measures of overt behavior.

There is a more controversial extension of methodological behaviorism: namely, that because conscious experience is inherently unscientific, consciousness should play no role whatsoever in scientific psychology. On these grounds, radical behaviorists avoided all questions about "purely mental events" such as imagery and thought. Practitioners of the IP approach have rejected this extreme position in a number of different ways. First, many IP psychologists use personal introspection as a source of ideas and hypotheses about cognitive events. Of course, these must then be subjected to more rigorous evaluation by measuring observable behavior in others to be scientifically respectable, but this is a standard procedure for much IP work in cognitive psychology. Second, the behaviors that IP psychologists measure are often overt reports of subjective experiences, such as ratings of perceptual or conceptual similarity (e.g., Shepard & Chipman, 1970) or verbal "thinking aloud" protocols while solving a complex problem (e.g., Newell & Simon, 1972). Thus, subjective experiences are considered important enough to warrant explanation, but only insofar as they can be made public by behavioral criteria. Third, IP psychologists are decidedly more interested in and optimistic about the scientific study of "purely mental events" such as imagery, thought, and consciousness.

Far from banishing them as inherently unscientific, IP have come to view them as valid and useful tools of scientific methods. During the past two decades, great deal of progress on the topic (see Cooper, 1982) and thought processes has been accomplished within the framework using objective measures such as these "mental events" in observing the literature. The resurgence of interest in consciousness and the things are conscious whereas others (e.g., 1978), whether a stimulus can affect consciousness (Dixon, 1971; Marcel, 1983, b), and account for consciousness (John, 1983) is deemed sensible only to the extent that the behavioral measures, thus conforming in important ways to behaviorism.

Theoretical behaviorism concerns psychology. "Radical behaviorism" is behavior that are not directly observable in behaviorism, this volume. These could not be directly observed in behaviorism, and are not generally opposed. Indeed, it is this principle that justifies the use of observable internal events in specific terms, because this is the mechanism by which behaviorism is generated in IP theories. (We shall return to this point in more detail.) These theoretical constructs have been unchallenged for long, even as they led to revisionism in behaviorism and Tolman's "purposive behaviorism." The behaviorism movement has been termed "mediating" stimuli and responses were taken to be "internal surrogates." We see a line of reasoning—strikingly similar to the one which the "primitives" were taken to be the case in the S-R theories and responses, both external and internal. It is important to note that this theoretical precursor to the S-R theories is formally equivalent to the S-R theories (Suppes, 1969). There are important differences between S-R theories and IP theories, however, and the vast repertoire of concepts for specifying the nature of the hypothesis.
3. INFORMATION

Far from banishing them as inherently unscientific, most IP psychologists have come to view them as valid and important topics that will yield to appropriate scientific methods. During the past decade or two, IP psychologists have made a great deal of progress on the topics of imagery (e.g., Kosslyn, 1980; Shepard & Cooper, 1982) and thought processes (e.g., Newell & Simon, 1972). This has been accomplished within the framework of methodological behaviorism by using objective measures such as reaction times and protocol analysis to anchor these "mental events" in observable behavior. There has lately been a resurgence of interest in consciousness itself, focusing on issues like why some things are conscious whereas others are not (Mandler, 1975; Shallice, 1972, 1978), whether a stimulus can affect IP events without itself becoming conscious (Dixon, 1971; Marcel, 1983a, b), and what sort of IP architecture is needed to account for consciousness (Johnson-Laird, 1983). Still, these projects are deemed sensible only to the extent that they can be grounded in objective behavioral measures, thus conforming in the end to the central tenet of methodological behaviorism.

Theoretical behaviorism concerns the nature of explanation in scientific psychology. "Radical behaviorism" is a theoretical view in which all accounts of behavior are couched in strictly "external" terms of environmental histories (cf. Riley, Brown, & Yoerg, this volume). Internal "mentalistic" constructs that could not be directly observed in behavior were rejected out of hand as improper theoretical objects. It is to this part of radical behaviorism that IP is fundamentally opposed. Indeed, it is this proposal that is rejected by all cognitivists, not just IP psychologists. Within the IP approach the behaviorist rejection of unobservable internal events is specifically opposed to the decomposition assumption, because this is the mechanism by which hypothetical "internal events" are generated in IP theories. (We shortly consider this issue of "unobservables" in more detail.) These theoretical strictures of behaviorism were too extreme to go unchallenged, even among otherwise devout behaviorists, and eventually they led to revisionist movements such as "neobehaviorism" and Tolman's "purposive behaviorism."

The neobehaviorist movement was initiated by Hull's (1952) introduction of "mediating" stimuli and responses within the organism (e.g., S→R→S→R) that were taken to be "internal surrogates" of observable S-R connections. Here we see a line of reasoning — strikingly similar to IP's decomposition assumption — in which the "primitives" were taken to be minimal associations between stimuli and responses, both external and internal. For this reason, neobehaviorism is an important theoretical precursor to IP. It has even been shown that such mediated S-R theories are formally equivalent to a particular class of finite automata (Suppes, 1969). There are important pragmatic differences between mediated S-R theories and IP theories, however, because the computer analogy brought with it a vast repertoire of concepts from computer science that could be used to specify the nature of the hypothesized internal events. These IP constructs—
including information structures such as lists, arrays, matrices, and so forth, plus operations such as encoding, storing, retrieving, transforming, comparing, deciding, branching, looping, and the like—are richer and more powerful tools for theorizing about mental events than were the neobehaviorists’ simple associations among “covert stimuli” and “covert responses.” Still, neobehaviorism accomplished the important step of lifting radical behaviorism’s absolute ban on theories that appealed to internal processes in explaining behavior.

Tolman’s (1932) “purposive” behaviorism went a step further in allowing openly “cognitive” constructs back into the theoretical arena. He argued convincingly for the importance of goals, hypotheses, plans, cognitive maps, and the like in understanding the behavior of rats as well as man. As we noted previously, Tolman’s theoretical views are actually “cognitivist” rather than “behaviorist,” although he remained a firm believer in the methodological tenets of behaviorism. We do not discuss the relation between Tolman and IP further here because an excellent discussion of this topic is presented elsewhere in this volume (Riley, Brown, & Yoerg).

The Role of Unobservables. We have claimed that the primary difference between the IP and radical behaviorist approaches to theories of mind is the former’s willingness—even eagerness—to postulate internal mental structures in accounting for observable behavior versus the latter’s unwillingness to do so. If the IP approach is preferable to this radical form of behaviorism, the use of unobservable constructs in psychological theory must be justified. We do so by appealing to the well-documented use of unobservable entities in other sciences and by examining a few successful cases within psychology itself.

There is a long and important history to the use of unobservables in the natural sciences, and it includes some of the most profound discoveries of the last several centuries. Examples abound of scientists proposing initially unobservable constructs on the basis of purely “behavioral” research and later having their theories confirmed by more “direct” observation. Biologists such as Mendel proposed the existence of genes that carry hereditary traits on the basis of measured regularities in the characteristics of offspring, and this happened long before DNA was actually observed within cell nuclei. Physicists like Rutherford deduced the internal structure of atoms from measuring the scatter that resulted when they were bombarded by X-rays, again, many years before the existence of these subatomic particles was confirmed more directly. There have also been many important cases in which unobservables were postulated without any presupposition of direct observation, e.g., Newton’s hypothesis of gravitational attraction and Darwin’s concept of natural selection. More recently physicists have suggested that gravity may, in fact, be carried by an elementary particle, but this conjecture antedated considerably their success as “unobservable” constructs in physics. No one has yet supposed that the process of natural selection will itself turn out to be a directly observable “thing.”

From such examples we can conclude that success of unobservables in science. The construct must provide a framework for organizing data, and be experimentally testable. They must be logically and behaviorally coherent. And they must subsequently be confirmed by the sort of results found in IP theories.

The criterion of adequacy describes observable effects; it relates several different sets of observable effects. The rationale for this justification of unobservables is not developed quite nicely by neobehaviorists like hunger and thirst in their theories, we will restate their arguments here. As with the objection to the presence of unobservables was actually continued and expanded by IP theory, it can be continued and expanded by IP theory. The role of unobservables in an IP-based theory is not limited to the role of unobservables in a theory of color vision (cf. De Vries) or in the theory of spatial vision (cf. De Vries). It has achieved the status of, say, an accumulation of knowledge. Each has provided a substantially new experimental test results in its domain.

Successful prediction is the test of a theory. It was repeatedly successful predictions of the existence of genes and physicists predicted the existence of quarks, confirmation of a theory’s predictions. The laws of nature arise from the existence phenomena. The major examples of experimental progress in psychology is generally true of the duplex theory, which is different in more complex cognitive, language understanding: Each case with a few results that seem to start
ays, matrices, and so forth, plus
transforming, comparing, de-
ther and more powerful tools for
obehaviorists’ simple associa-
to an explanation of behav-
went a step further in allowing
eoretical arena. He argued con-
oses, plans, cognitive maps, and
as well as man. As we noted
ally “cognitivist” rather than
believer in the methodological
relation between Tolman and IP
this topic is presented elsewhere
med that the primary difference
ches to theories of mind is the
date internal mental structures in
ater’s unwillingness to do so. If
of behaviorism, the use of
must be justified. We do so by
ervable entities in other sciences
psychology itself.
se of unobservables in the natural
 profound discoveries of the last
proposing initially unobservable
research and later having their
ion. Biologists such as Mendel
ditary traits on the basis of mea-
, and this happened long
uclei. Physicists like Rutherford
asuring the scatter that resulted
ays before the existence of
directly. There have also been
were postulated without any pre-
on’s hypothesis of gravitational
ction. More recently physicists
ried by an elementary particle.
success as “unobservable” con-
t the process of natural selection
thing.”

From such examples we can discern at least two important factors in the
success of unobservables in scientific theories. The first is adequate description:
The construct must provide a succinct and parsimonious account of results
known at the time the theory is proposed. The second is successful prediction:
They must figure centrally in generating novel hypotheses about the existence of
new phenomena in substantially different circumstances, and these predictions
must subsequently be confirmed. We submit that unobservable constructs of the
sort found in IP theories are at least potentially adequate on both counts.

The criterion of adequate description is, of course, the reason for postulating
unobservable constructs in the first place: It “makes sense” of some set of data
or relates several different sets of data in ways not previously considered. The
rationale for this justification of unobservables in psychological theory was de-
developed quite nicely by neobehaviorists in defending the use of internal variables
like hunger and thirst in their theories (cf. Tolman, 1932) and there is no need to
restate their arguments here. As we said earlier, the move away from the stricture
against unobservables was actually started by the neobehaviorists and merely
continued and expanded by IP psychologists. There is certainly no problem in
principle with IP constructs providing adequate descriptions of well-documented
phenomena, and there are several noteworthy successes, such as the two-stage
theory of color vision (cf. De Valois & De Valois, 1975) and the spatial frequency
theory of spatial vision (cf. De Valois & De Valois, 1980). Neither of these
has achieved the status of, say, subatomic particles or gravity in physics, but
each has provided a substantial framework for describing a large number of
experimental results in its domain.

Successful prediction is the true hallmark of success for any scientific theory.
It was repeated successful predictions that really convinced biologists of the
existence of genes and physicists of the reality of subatomic particles. Repeated
confirmation of a theory’s predictions leads eventually to such a dense cluster of
results around it that the theory displaces specific results in textbooks. There are
at least a few good examples in psychology. For instance, color vision is usually
presented in texts on perception almost exclusively by the theory with just a few
examples of experimental procedures to give a feeling for the research. The same
is generally true of the duplex theory of pitch perception. The situation is quite
different in more complex cognitive domains, such as semantic memory and
language understanding: Each of several competing theories is presented along
with a few results that seem to support it. and it is the results, not the theory, that

---

9 We are considering that psychophysical theories generally fall within the IP framework. Al-
though much of this work predated the IP movement in cognitive psychology, our position is justified
on the grounds that most psychophysically derived theories conform to all five assumptions afore-
mentioned and to the whole spirit of the IP enterprise. Indeed, it was this kind of connection between
IP and sensory processes that Lindsay and Norman (1972) developed so skillfully in their seminal
textbook, Human information processing.
are emphasized as characterizing the domain in question. Such theories are "brittle" in the sense that they break down (or simply do not apply) when seemingly minor changes in procedure are introduced, and they often fail to make successful new predictions or succeed only under certain unusual conditions. There is no reason to suppose, however, that robust IP theories cannot someday be achieved in these areas as they have been in sensory domains; the problems are just a great deal more numerous and complex.

There is a third possible criterion of success that applies to at least some unobservable constructs proposed to science: potential observability. A construct is potentially observable if it is, in principle, capable of being measured directly, even if the means for doing so are not yet at hand. Not all unobservable constructs have this status, of course, as exemplified by gravity and natural selection. In psychology, a distinction has been made between "intervening variables," which do not imply potential observability, and "hypothe
tical constructs," which do (MacCorquodale & Meehl, 1948). Proposed IP components are clearly meant to be stronger than intervening variables, which merely redescribe empirical relations succinctly in somewhat different form. Because intervening variables are not meant to have any reality "inside the black box" (i.e., they claim only to specify what the mapping is, not how it is accomplished), they only play a role within the IP approach as part of mapping theories. IP theories, however, do make substantive claims about what goes on inside the head, so the internal operations they propose must be some species of hypothetical construct. This does not mean, however, that all such internal operations imply the existence of specific, isolable, and identifiable neurological mechanisms that embody them.

We believe that IP operations are potentially observable in the brain in just the same sense that programs are potentially observable in Turing machines. If the program is "hardwired" into the machine, it is physically observable in the wiring of its circuits. If it is implemented on a general-purpose machine, it is functionally observable in the behavior of the machine, but not physically observable in the sense of dedicated hardware components. In the latter case, the hardware constitutes the architecture on which the programs run, but this architecture only weakly constrains the programs that can run. In the case of the brain, peripheral operations seem to take place on "dedicated hardware" where there is some distinctly physical reality to its functional components. More central "cognitive" operations may well take place on a more general-purpose architecture. Most IP theorists who try to account for complex cognitive events find it hard to imagine that the brain has enough distinct neural circuits for all the necessary mental processes to be realized in dedicated hardware.

In summary, there is no good reason to avoid unobservable constructs in psychological theory and several good reasons to use them as needed. The natural sciences provide ample precedent for this, and psychology itself includes some successful examples. The IP theories are, in principle, capable of making new prediction. Some may even be true and applicable to the brain, although this is not a necessary condition.

Ecologism

Ecological psychology has been a major influence on psychology (e.g., Gibson, 1950, 1966, 1979). Ecological psychology is a highly influential approach, partly because ecological psychology provides a conceptual framework for the study of perception in terms of negating the IP approach as a cognitive psychology approach and examining the relationship between organisms and their environments. The central theme of ecological psychology is the idea that perception is an active process. Gibson's view of perception was that it is an "active" process, which involves an organism actively interacting with its environment. This is in contrast to the "passive" view of perception, where the organism passively receives stimuli from the environment.

One way to understand the importance of the active role of the organism within which it was initially conceived is to consider the active role of the organism. This is in contrast to the passive role of the organism, which simply reacts to stimuli from the environment. The active role of the organism within which it was initially conceived is to actively interact with the environment, which is a key feature of ecological psychology. This is in contrast to the passive role of the organism, which simply passively receives stimuli from the environment.
n in question. Such theories are (or simply do not apply) when introduced, and they often fail to only under certain unusual conditions. That robust IP theories cannot have been in sensory domains; the complex system that applies to at least some potential observability. A construct capable of being measured directly, hand. Not all unobservable confirmed by gravity and natural selection, the “intervening variables,” observable, and “hypothetical” (Eehl, 1948). Proposed IP components are unobservable, which merely somewhat different form. Because my reality “inside the black box” mapping is, not how it is accomplishing as part of mapping theories claims about what goes on inside it. There must be some species of hypothesis that all such internal operations include identifiable neurological mechanisms observable in the brain in just the way observable in Turing machines. If the it is physically observable in the general-purpose machine, it is a machine, but not physically observable components. In the latter case, the one programs run, but this architecture can run on it. In the case of the “dedicated hardware” where its functional components. More often place on a more general-purpose count for complex cognitive events in distinct neural circuits for all the dedicated hardware.

avoid unobservable constructs in ones to use them as needed. The bias, and psychology itself includes some successful examples. The kinds of unobservable constructs that appear in IP theories are, in principle, capable of both adequate description and successful prediction. Some may even be potentially observable in the “hardware” of the brain, although this is not a necessary claim.

Ecologism

Ecological psychology has been proposed, at least in part, as an alternative to IP psychology (e.g., Gibson, 1979; Shaw & Bransford, 1977). It is not easy to summarize the ecological approach, partly because it is not yet well defined, and partly because ecological psychologists themselves capture their approach partly in terms of negating the IP approach as they see it. Here we briefly discuss the ecological approach and examine its relation to the foregoing view of IP.

Generally speaking, ecological psychologists reject the computer program metaphor and the appeal to representations and processes as explanatory concepts in theorizing about perception and cognition. Rather, they view the organism from an evolutionary perspective—i.e., as a system whose primary purpose is self-survival and adaptation—and organisms in their environments as mutually constraining systems.

Ecological psychology is primarily derived from J. J. Gibson’s approach to perception (1950, 1966, 1979). Gibson viewed perception as a biologically adaptive activity and emphasized the dynamic interaction between organisms and their environments. The centrality of the role that the environment plays in Gibson’s view of perception was aptly and succinctly captured by Mace (1977): “Ask not what’s inside your head but what your head’s inside of.”

Gibson asserted that there is information in complex patterns of stimulation available to the organism, corresponding to the important distal objects and events in its environment. This significant information is to be found in “higher order” variables that remain invariant over the visual transformations taking place as the organism actively explores the world around it. In his theory of “ecological optics” Gibson tried to identify and specify these invariants. According to his theory the active perceptual system “picks up” this information without the involvement of logical processes using specific knowledge about the world. This approach led him to the notion of “direct” perception.

One way to understand the idea of direct perception is in the historical context within which it was initially proposed. Gibson believed that perception was direct in the sense that, contrary to the Helmholtzian notion of “unconscious inference,” it does not involve epistemic mediation. As we understand it, this argument does not necessarily deny that representations and processes play a role in perception, but only that it does not make use of explicit forms of knowledge and logical inference processes of the sort associated with thought, memory, and
problem solving. We do not see this view of direct perception as incompatible with the basic tenets of IP. As evidence for this, it is entirely possible to construct a "data driven" IP theory of perception that does not rely on explicitly repre- sented "facts" or logical "deductions": Marr's (1982) theory is one prominent example, at least up to the point at which patterns are identified as instances of known types. Such IP theories can easily be seen as consistent with this weak interpretation of direct perception.

There is, however, a stronger interpretation of direct perception that seems to strike at the heart of the IP approach by rejecting the decomposition assumption. The goal of ecological optics is to specify the mapping from proximal stimuli to perceptual responses, and its contention is that this can be done by correlating physical invariants in the dynamic optic array with perceptual responses. Thus Gibson's theory is not concerned with the mechanisms that enable an organism to do this, except that the perceptual system detects the invariants by some sort of physiological "resonation" (Gibson, 1966). It is in this sense—namely, the lack of interest in specifying the nature of internal mechanisms—that Gibson rejects, if not the decomposition assumption itself, certainly the motivation behind it. Ullman (1980) presents an extended critical analysis of this view, and the interested reader is referred to his paper and the many commentaries that follow it for more information.

From the IP point of view, Gibson's theory is an important attempt to formulate a mapping theory of perception, because he was occupied almost entirely with what the mapping is to the exclusion of how it might be achieved. Thus, although his theory is incomplete from an IP standpoint, it is entirely compatible with the IP framework, because an ecological mapping theory of perception is the logical starting point for an IP theory of perception (see Marr, 1982). From Gibson's point of view, however, IP theories of perception are not compatible with his ideas, because he believed that psychological theories of internal mechanisms were irrelevant to understanding perception. Recent followers of Gibson's ecological approach claim that Gibson did recognize that information pickup involves processes, but from his brief allusions to resonating, optimizing, and

10We refer here to "explicit" knowledge and "logical" inference to distinguish the classical position of "unconscious inference" from some modern theories that seem to make use of knowledge and inference of a very different sort. One interesting example is Marr and Poggio's (1976) model for computing global stereopsis: The excitatory and inhibitory connections among the binocular processing units actually embody knowledge of the world (that the visual world is generally continuous in depth, at least over local regions) and of light (that because most surfaces are opaque, each point on an image generally comes from just one depth plane). Even so, this model does not seem to be in the spirit of Helmholtzian "unconscious inference," but more like a sort of "resonance" process of which Gibson might have approved. The difference between classical "unconscious inference" and this type of computational theory seems to be that the knowledge in the former case is explicit, whereas in the latter case it is implicit, and that the rules of inference in the former case conform to the standard logical conception, whereas in the latter case they do not.

the like, they infer that these processes are carried out by a general-purpose device (Turvey, Kugler, & Shaw, 1982). It should be clear at this point that the IP approach is not incompatible with the basic tenets of IP; it only requires differentiating between the common practice of IP psychology and IP approaches. It is not many mapping theories in IP that are composed exclusively on the internal constraints that support the mapping. In fact, one may argue that the most recent attempts among IP theories (Marr, 1981; Marr, 1982; Shepard, 1981) to a summary, there is a clear distinction between the two approaches. According to the IP view, conflicting as they are often asked.

Computationalism

As we have seen from the foregoing, the key to ecologism, the most controversial of the many theories of perception, is that perceptions are decomposable. For a higher-level feature to be decomposable, it must be something one could prove that a lower-level feature is decomposable. This problem has been studied by the field of computability (see Johnson-Laird, 1983, for a discussion). From this work by Church (1936), we can conclude that if a function (9) is such that if the mental mapping is decomposable, then it can indeed be decomposed. This means that the mechanical mapping is decomposable according to the computational view of perception.

Probably the best known definition of computation is that of Church (1936): A function is computable if it is definable by a Turing machine (now called a "Turing machine")—that is, it can be defined by a Turing machine, which a function is computable if it is definable by a Turing machine (now called a "Turing machine")—that is, it can be defined by a Turing machine, which a function is computable if it is definable by a Turing machine (now called a "Turing machine")—that is, it can be defined by a Turing machine, which a function is computable if it is definable by a Turing machine (now called a "Turing machine")—that is, it can be defined by a Turing machine.
3. INFORMATION

information as incompatible with the like, they infer that these processes have to be different from those suggested by a general-purpose device such as the modern digital computer (Carello, Turvey, Kugler, & Shaw, 1982). This assertion may constitute an argument against the computer program metaphor, but not necessarily against the IP approach as presented here.

It should be clear at this point that we do not see ecological psychology as incompatible with the basic tenets of IP. Why then, do ecological psychologists often present their view as a reaction against IP? The answer to this question requires differentiating between theoretical principles of IP on the one hand and common practice of IP psychologists on the other. As noted previously there are not many mapping theories in psychology, and most current IP models focus almost exclusively on the internal structure (representations and processes) proposed to accomplish the mapping without any attempt to specify the ecological constraints that support the mapping in the first place. There have been some recent attempts among IP theorists to remedy this situation (e.g., Bregman, 1981; Marr, 1982; Shepard, 1981), but it is generally a criticism well taken. In summary, there is a clear difference in emphasis between the ecological and IP approaches as typically practiced, even though they are not theoretically as conflicting as they are often assumed to be.

Computationalism

As we have seen from the foregoing discussion of cognitivism, behaviorism, and ecologism, the most controversial assumption of the IP approach is that mental events are decomposable. For this reason, it would certainly be reassuring if somehow one could prove that this were true. As it turns out, a closely related problem has been studied by mathematical logicians in the “theory of computability” (see Johnson–Laird, 1983, for a psychologically motivated introduction). From this work by Church, Turing, and others, we know at least something about certain conditions under which an operation is decomposable. It turns out that if the mental mapping is an example of what is called a “computable function,” then it can indeed be decomposed into a set of primitive functions that are combined according to well-specified rules, just as would be required for the decomposition assumption to be unequivocally true.

Probably the best known definition of a computable function is due to Turing (1936): A function is computable if it is the input–output mapping of a machine (now called a “Turing machine”) that reads binary symbols from and writes binary symbols to an indefinitely long tape. The sequence of logical steps by which such a function is computed in such a machine is called an algorithm, and there are many different algorithms that can result in the same computable function. Indeed, it turns out that there are many different sets of “primitive functions” and rules for combining them that are, in principle, capable of gener-
ating the set of all computable functions; Turing machines, recursive functions, lambda calculus (Church), Post canonical systems, Markov processes, and others (cf. Newell, 1980). So if the mental mapping is actually computable, then the job of IP psychology amounts to discovering what the mind's primitive IP functions are and how they are composed into the larger, more complex algorithms of the human mind (see previous section on "The Role of Primitives"). Mental architecture is probably quite different from that of standard universal Turing machines (or modern digital computers), and a psychologically more plausible formulation has been described by Newell (1980).

We can now say that the decomposition assumption would be true if the mental mapping is a computable function. But is it? There are many functions that are not, so perhaps it is one of these. Unfortunately, no way is currently known to determine this analytically because there is no well-defined method for determining whether an arbitrary function is computable or not. The only alternative is to try to show that it is by finding an algorithm that "simulates" the mind. The goal is to program a computer so that it passes Turing's test: to behave so much like a person that experts cannot distinguish its behavior from that of a real person (Turing, 1950). Because such a program would, by definition, be a well-defined composition of primitive IP operations, its existence would prove by demonstration that mental decomposition is possible. Because an infinite number of other algorithms will produce the same computable function, there would still be the problem of finding the right decomposition, but at least we would know that such a decomposition exists. Psychologists would presumably play a central role in determining which of several alternative decompositions is correct.

This line of reasoning suggests an approach to cognitive theory that we call "computationalism" (or "simulationism"). Its proponents assume that the mental mapping is indeed a computable function, at least as a working hypothesis, and proceed to try to attain the goal presumed possible. Their method is to write computer programs (which are, by definition, computable) that try to simulate human behavior. Searle (1980) has called this sort of approach "weak AI"; the programmer views the modern computer as an important tool for discovering things about the nature of mentality but does not assume that the programmed machine actually has mental states. Later we contrast this view with "strong AI" in which the computer is presumed actually to have the mental states that it simulates.

The primary difference between computationalism and IP concerns the logical relation between the assumptions of computability and decomposability. As we have said, computability implies decomposability, but does decomposability imply computability? It turns out that it does not, because the mental mapping, as a whole, may not be computable, yet it still may be possible to decompose it meaningfully into an IP theory. Perhaps the easiest way to see this is to suppose that there is one small part that is isolated in its own "black box." Then it is clear that the system could be a meaningful overall input--output mapping without any such uncomputable component necessarily preclude successful analysis.

It is not obvious what this sort of process is, and Searle (1983) have recently argued against it like this might be true. They argue that the subset of mental operations may indeed be analyzable by the "network"--and this part of the mind could be computable. However, they also argue that the problem we have yet fundamental skills, presupposed their fundamental skills, presupposed what they call "the Background of the mind," and this analysis. In truth, their reasons for not having yet had much impact on the debate are critical for the present discussion: it may turn out to be right and computationalism to have such a potentially valid assumption is a useful approach to solving at least some of the problems that depend on how much decomposability of the mind as a result.

The possibility of decomposability of the mind as a result of a relation between IP and weakness. Compositionalists start with certain premises that constitute the program code of the computer, which may or may not be computable. They then compose it into flow diagrams. The diagram may not be in itself computable. They then proceed to decompose, but not a necessary consequence of the whole system. If AI is less likely than IP to be computable.

We see, then, that IP psychologists who support the mind will ultimately be engaged in a digital computer, because, strictly speaking, different machines will be able to simulate it, but they may also change the very nature of the mind.

483, KIMCHI
ring machines, recursive functions, systems, Markov processes, and other things actually computable, then the ring what the mind's primitive IP into the larger, more complex action on "The Role of Primitives".

ent from that of standard universal uters), and a psychologically more Newell (1980).

n assumption would be true if the But is it? There are many functions Unfortunately, no way is currently

there is no well-defined method for computable or not. The only alter-

an algorithm that "simulates" the

that it passes Turing's test: to behave

istinguish its behavior from that of a

rogram would, by definition, be a perations, its existence would prove

on is possible. Because an infinite

he same computable function, there

ight decomposition, but at least we

ysts. Psychologists would presumably several alternative decompositions is

uch to cognitive theory that we call 

. Its proponents assume that the men-

on, at least as a working hypothesis, ed possible. Their method is to write

ion, computable) that try to simulate his sort of approach "weak AI": The

as an important tool for discovering does not assume that the programmed we contrast this view with "strong

ually to have the mental states that it

ationalism and IP concerns the logical utability and decomposability. As we sosability, but does decomposability is not, because the mental mapping, as still may be possible to decompose it the easiest way to see this is to suppose

that there is one small part that is not computable, and that this part can be isolated in its own "black box" as an uncomputable primitive within the system. Then it is clear that the system could be decomposed in this way, even though its overall input-output mapping would not be computable. Indeed, there might be many such uncomputable components in the mind, but their existence does not necessarily preclude successful decomposition.

It is not obvious what this sort of analysis would be like, but Dreyfus (1979) and Searle (1983) have recently offered arguments suggesting that something like this might be true. They allow that some of human knowledge and mental operations may indeed be analyzable into component pieces—what Searle calls "the Network"—and this part may well be compatible with computationalism. However, they also argue that there will necessarily be some residue of complex yet fundamental skills, presuppositions, and stances toward the physical world—what they call "the Background"—that will defy all attempts at computational analysis. In truth, their reasons for claiming this are not entirely clear and have not yet had much impact on the beliefs of most computationalists. However, the critical point for the present discussion is that even if Dreyfus and Searle did turn out to be right and computationalism wrong, mental decomposition would still be a potentially valid assumption for cognitive psychology and IP a potentially useful approach to solving at least some of its problems. Its usefulness would depend on how much decomposition were possible before the intractable "Background" were reached, and how much would have been learned about the nature of mind as a result.

The possibility of decomposition without strict computability clarifies the relation between IP and weak AI (computationalism). Classical computationalists start with certain precisely stated primitive functions and show how to compose them into more complex functions. The resulting complex function is necessarily computable by definition. IP psychologists start with a complex function, which may or may not be computable, and proceed by trying to decompose it into flow diagrams. The components of these flow diagrams may or may not be themselves computable. Thus, computability is a sufficient condition for decomposition, but not a necessary one. Weak AI therefore places stronger constraints on a theory of mind than does IP. By the same token, of course, weak AI is less likely than IP to be correct.

We see, then, that IP psychologists are not necessarily committed to believing that the mind will ultimately be simulable on an appropriately programmed digital computer, because, strictly speaking, IP does not imply computability. This conclusion is consistent with the fact that serious doubts exist among a substantial subset of IP psychologists that mental events will ever be fully simulated on the sort of digital computers that are currently available. Perhaps someday different machines will be developed that will change such opinions, but they may also change the very definition of computability.
Functionalism

Functionalism was proposed by philosophers as a possible solution to the mind/body (or mind/brain) problem (e.g., Fodor, 1965; Putnam, 1960). It assumes that the necessary and sufficient conditions for mentality are to be found in the functional organization of the brain rather than in its particular physical embodiment. This implies that any device that has the same functional organization as the human brain literally has mental states in exactly the same sense that people do. This includes all aspects of mentality, including intentionality, consciousness, and internal experiences.

The initial intuitive basis for this claim comes, once again, from the computer analogy. There is an important sense in which the same program (software) can be run on many different digital computers (hardware). The analogous implication for the mind/body problem is that the same "mind" might potentially be instantiated on a wide variety of physically different devices. The essence of functionalism, then, is that systemic organization defines the necessary and sufficient conditions for having mental states, and brains are merely the one material form in which this particular organization has been realized so far. Nothing in principle prevents other sorts of objects from having the same mental states people do, provided that they have the right sort of organization.

It is certainly true that many IP psychologists seem to reject the spirit of functionalism by taking certain aspects of neurology very seriously in their theorizing. For instance, several psychological IP theories have been proposed that rest importantly on neurological or quasineurological constructs such as excitation and inhibition (e.g., Hinton & Anderson, 1981; Posner, 1978), aftereffects due to priming (e.g., Posner, 1978), and even hemispheric specialization (e.g., Friedman & Polson, 1981). On closer examination, however, these proposals do not really conflict logically with functionalism. If true, they would only imply that brain properties like activation, inhibition, aftereffects, and so forth play some functional role in the nature of mind, and that for another physical object to have the same mental states it would presumably need to have the same functional properties. These might easily be realized in nonbiological objects such as a computer. Thus, even if such theories were correct in detail, they would not strictly imply that it is the physical properties of the central nervous system that are responsible for their role in mental events. However, these examples do demonstrate rather convincingly that many IP psychologists do not really "buy the functionalist line" that the hardware of the brain is uninformative about the nature of mind.

A stronger argument against functionalism can be made on strictly logical grounds. Functionalism, and the principles on which it is based, require only that some event "mental" can be captured in events that are not captured, and that these events will be recognized for something literally having them. The most obvious candidates for these things literally to have mental states are the devices that are used to capture this aspect of mentality. For instance, such devices are specified functional organization as a computer or a computer program will be able to characterize certain aspects of it that are scientists have proposed such theories (e.g., Shallice, 1978).

The important point in understanding the validity of the IP approach is that IP theory captures all aspects of mental events. Most psychologists hope that IP theory will provide a complete account of the mind, but not many would feel that it has yet achieved such a result. In this case, the "computer" turn out to be merely a necessary part of the system of mental states that people describe. It will be realized in the physical world, however, if it turned out this way, it would provide a stronger argument against functionalism because it would force the specific biological processes out of the mind. Functionalism, however, it should be clear that the evidence in support of Searle's matrix theory of consciousness may well turn out to be entirely within the realm of functionalism.

Turing-Machine Functionalism: A machine that can be programmed to perform a particular task is essentially a functional system that is capable of performing that task. Indeed, this would be true of any system, including one made of electronic

\[\text{footnote}{11}\text{The modern philosophical doctrine called functionalism should not be confused with the older psychological school also known as functionalism, which is associated with the work of James, Hall, Cattell, Dewey, Woodworth, and others. There is little, if any, connection between them.}
A stronger argument against the claim that IP is equivalent to functionalism can be made on strictly logical grounds. It turns out that IP is weaker than functionalism, and the principle reason is that because IP theories are descriptions, they require only that some, not necessarily all, aspects of what makes an event "mental" can be captured by the flow diagrams. There may be certain aspects that are not captured, and these may be considered necessary conditions for something literally having mental states as opposed to merely simulating them. The most obvious candidate for such a mental aspect is consciousness: Qualitatively different phenomenal experiences might be necessary for something literally to have mental states, and yet IP theories might not be able to capture this aspect of mentality. One could be an IP theorist and a true dualist at the same time, for instance. Such a person would believe that, whereas IP flow diagrams specify the functional organization of the brain, internal experiences are required for it literally to have a mind, and these are quite independent of the specified functional organization. It is also possible, of course, that IP theories will be able to characterize consciousness in terms of flow diagrams, at least those aspects of it that are scientifically approachable, and several IP psychologists have proposed such theories (see Johnson-Laird, 1983; Marcel, 1983b; Shallice, 1978).

The important point in understanding the relation of IP to functionalism is that the validity of the IP approach does not stand or fall in terms of whether it captures all aspects of mental events, whereas functionalism does. Naturally, IP psychologists hope that IP theories will turn out to account for everything, but not many would feel that it had failed if it turned out to leave some mysteries unsolved. In this case, the "correct" flow diagram theory of the mind would turn out to be merely a necessary, but not sufficient, condition for having the sort of mental states that people do. It would be devastating to a functionalist, however, if it turned out this way. For example, Searle (1983, Epilogue) has argued against functionalism by suggesting that consciousness is inherent to specific biological processes occurring in brains, but this is not necessarily an argument against the IP approach as described earlier. In defense of functionalism, however, it should be noted that there is presently no scientific evidence in support of Searle's materialist, biological view, only his own intuitions. Consciousness may well turn out to be a problem whose solution will be found entirely within the realm of functional organization.

Turing-Machine Functionalism (or "Strong AI"). If a properly programmed computer has the same functional organization (i.e., implements the same algorithm) as the human mind, then it follows from functionalism that the machine would literally have a mind in exactly the same sense that people do. Indeed, this would be true of any machine implementing the same algorithm, including one made of electronic, mechanical, optical elements, or whatever.
This conjunction of computationalism and functionalism is usually called "Turing-machine functionalism" (e.g., Block, 1978; Block & Fodor, 1980; Fodor, 1980) and sometimes "strong AI" (Searle, 1980). Note that it would not be sufficient for the computer merely to duplicate the input–output mapping function of the human mind because many different algorithms (and, therefore, different functional organizations) can produce the same mapping. Functionalism proposes that it is the functional organization that must be the same, and this is a much stronger constraint than just that the mapping be the same. For example, if one proposed an IP theory of consciousness that required that two processes run in parallel, a strictly serial computer that merely simulated parallelism by switching back and forth between them would not have the organization specified by the theory, even though it would exactly mimic its input–output function. Therefore, it need not be itself conscious but might well only mimic consciousness. It isn’t just what gets done that matters, it’s how it gets done. In any case, it should be clear that IP as characterized here is not equivalent to Turing-machine functionalism, because it does not strictly imply either computationalism or functionalism, much less their conjunction.

At this point we can summarize the foregoing discussion by putting it all into a single framework involving just two issues and the positions taken on them by each approach. The first issue concerns the analyzability of mental processes into functional components. The strong position is that mental processes are computable (computationalism and Turing-machine functionalism), and the weak position is that they are merely decomposable (functionalism and IP). Radical behaviorism and Gibsonian ecologism take the more pessimistic position that the question of analyzability itself is entirely misguided.

The second issue concerns the logical implications of functional organization for the nature of the human mind. The strong position here is that the functional organization is both necessary and sufficient for having mental states (functionalism and Turing-machine functionalism) and the weak position that it is merely necessary (computationalism and IP). Again, it is possible to hold the more pessimistic position that functional organization is logically unrelated to issues concerning the human mind (radical behaviorism and ecologism). Thus, the four meaningful conjunctions of the two positions on these two issues define the four principle modern views shown in Table 3.1. Of these, Turing-machine functionalism is the strongest and IP the weakest. Note that, by the same token, IP is the most general view and is compatible with all three of the other approaches.

Summary and Conclusion

We have presented a metatheory in terms of five central assumptions: informational description and representation, the notion of physical embodiment—further refined as a constructive reduction, together these five assumptions of mental processes in terms of component operations and the temporal relations of the flow diagram are taken to include a mapping theory of each process.

We then contrasted this view with a second opposing view: behaviorism, ecologism, and ID. The Turing-machine functionalism (stronger) and IP (weaker) views can be captured, at least schematically, by the four major issues: the extent to which mental states and processes and the logical implications of mental states. According to the theory (with) computationalism, functionalism.

We believe that IP is currently the strongest of cognitive theories, but this is obviously just our current opinion and discussion in the hope that it will be stronger than has previously been available. We also believe that we can see one or more ways, being clear about the ultimate goal of understanding cognitive processes.

ACKNOWLEDGMENTS

We are grateful to the many people who have helped us in this work: Phil Johnson–Laird, Tony Marcel, Stafford Shapiro, and the people of this volume, Terry Knapp and Lynn Nadel.
ntionalism is usually called "Turing" (1978; Block & Fodor, 1980; Fodor, 1980). Note that it would not be the input-output mapping function different algorithms (and, therefore, fuzc the same mapping). Functionalization that must be the same, that the mapping be the same. For consciousness that required that two inputs that merely simulated parallelism would not have the organization could exactly mimic its input-output process but might well only mimic in matters, it's how it gets done. In characterized here is not equivalent to state not strictly imply either computation or conjunction.

Opening discussion by putting it all into and the positions taken on them by analyzability of mental processes into is that mental processes are combine functionalism, and the weak (functionalism and IP). Radical the more pessimistic position that the misguided.

Implications of functional organization ng position here is that the functional critical for having mental states (fun- m) and the weak position that it is P). Again, it is possible to hold the organization is logically unrelated to behaviorism and ecologism). Thus, positions on these two issues define Table 3.1. Of these, Turing-machine most. Note that, by the same token, natable with all three of the other crucial distinction to me. The example that

<table>
<thead>
<tr>
<th>TABLE 3.1</th>
<th>Logical implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Necessary &amp; Sufficient</td>
<td>Necessary</td>
</tr>
<tr>
<td>Computable</td>
<td>Turing-Machine Functionalism</td>
</tr>
<tr>
<td>Decomposable</td>
<td>Functionalism</td>
</tr>
<tr>
<td>(Irrelevant)</td>
<td>-</td>
</tr>
</tbody>
</table>

Summary and Conclusion

We have presented a metatheoretical view of the IP approach to cognition in terms of five central assumptions that underlie it. The first two assumptions—informational description and recursive decomposition—are the most critical. The third and fourth assumptions—flow continuity and flow dynamics—further specify the nature of the decomposition assumption, and the final assumption—physical embodiment—further specifies the informational description assumption. Together these five assumptions form the basis for constructing IP theories of mental processes in terms of information flow diagrams that represent the component operations and the temporal relations among them. Unless the operations of the flow diagram are taken to be primitives, the IP theory should also include a mapping theory of each operation.

We then contrasted this view of IP to other prominent approaches to cognitive theory: behaviorism, ecologism, computationalism (weak AI) functionalism, and Turing-machine functionalism (strong AI). Many of the differences among these views can be captured, at least schematically, by their stands on two fundamental issues: the extent to which mental processes are analyzable into functional components and the logical implications of this analysis in accounting for the nature of mental states. According to this analysis, IP is weaker than (but compatible with) computationalism, functionalism, and Turing-machine functionalism.

We believe that IP is currently the most viable theoretical approach to cognition, but this is obviously just our own opinion. We offer the foregoing analysis and discussion in the hope that it gives a clearer conception of the IP approach than has previously been available. Even if IP turns out to be fatally flawed in one or more ways, being clear about the underlying issues can only help in the ultimate goal of understanding cognition.

ACKNOWLEDGMENTS

We are grateful to the many people who read and criticized earlier drafts of this chapter: Phil Johnson-Laird, Tony Marcel, Saul Sternberg, and Richard Young plus the editors of this volume, Terry Knapp and Lynn Robertson. The initial draft was written while Steve
Palmer was visiting the MRC Applied Psychology Unit in Cambridge, England, and he wishes to thank Alan Baddeley for the opportunity to work there and Phil Johnson-Laird for the many hours of discussion that helped shape some of the ideas contained in this chapter. Preparation of the chapter was facilitated by a grant from the Alfred P. Sloan Foundation to the University of California, Berkeley, and by Grant BNS-8319630 from the National Science Foundation to the first author.

REFERENCES


action between hypothetical constructs and 107.

ceiling: Ask not what’s inside your head, 30.

transfond (Eds.). Perceiving, acting, and key.

phon. Experiments on visual masking and 7.

phon: An approach to the relation between emotive Psychology, 15, 238–300.


dels of human information processing: In 300:


as empirical inquiry: Symbols and search: 300.

d and resource-limited processes. Cognitive 300.


anization: A transformational approach. In 300.


in the perceived pointing of ambiguous 300.

Perception and Performance, 7, 88–114. 300.


Putnam (Ed.). Mind, language and reality: 300.


