

4

The Representation of Knowledge in Memory¹

David E. Rumelhart

University of California at San Diego

Andrew Ortony

University of Illinois at Urbana-Champaign

INTRODUCTION

While originating from the senses, knowledge is not a blind record of sensory inputs. Normal people are not tape recorders, or video recorders; rather, they seem to process and reprocess information, imposing on it and producing from it knowledge which has structure. The human memory system is a vast repository of such knowledge. Some of this knowledge seems to be in the form of specific memories of particular events which we have experienced; some of it seems to be in the form of more general abstractions no longer tied to any particular time, place, or source. It is one of the tasks of a theory of the representation of knowledge to provide a characterization of the way in which knowledge is structured so that progress may be made toward answering other important questions: how is memory organized so as to usually permit relevant information to be accessed when required? how is old knowledge employed in the acquisition of new? how does our current knowledge state modulate our actions? No theory we know of provides a completely satisfactory answer to questions such as these. Nevertheless, substantial strides toward answering them have been made during the past five or ten years. It is, we believe, a tribute to these strides that a paper on the representation of knowledge can be found in a volume such as this.

¹The paper originally presented at the Conference by Rumelhart and discussed by Ortony has been replaced by this joint effort. Hence the absence of formal and open discussion.

The progress to which we refer can be regarded as the focal point of a new emerging discipline called Cognitive Science. The research of an increasing number of people working in artificial intelligence, cognitive psychology, and linguistics on problems concerned with the representation of meaning and the structural and processing aspects of knowledge, reveals a substantial convergence of opinion on the essential components of systems for representing knowledge. Recent papers by Bobrow and Norman (1975), Minsky (1975), Norman, Rumelhart, and LNR (1975), Rumelhart (1975), Schank and Abelson (1975), and Winograd (1975) all attest to this convergence. While differing from one another, sometimes in important ways, there is nevertheless agreement on the broad outline stated or implied by these authors. In this paper we will sketch this broad outline and develop some of the arguments in favor of the general approach. Although many aspects of the ideas we will develop differ from those of earlier approaches, it is not our purpose to critically review any of them, nor to attempt any explicit comparison between our approach and that of others. Rather, we have drawn upon what we consider the best aspects of each of these other developments and constructed what seems to us the most reasonable composite view. To the degree that technical detail is required we have formulated our ideas in the context of the Active Structural Networks of Norman, Rumelhart, and LNR (1975) and as a synthesis of the work of Rumelhart and Levin (1975), and of Rumelhart (1975).

While we will not be comparing our account of knowledge representation, which we will express in terms of "schemata" (singular — "schema"), with the accounts of others, it is appropriate that we indicate the concepts with which proponents of similar views are associated. The theory proposed by Minsky (1975) is based on what he called "frames." Winograd (1975), Charniak (1975), and other workers in Artificial Intelligence have largely followed Minsky's usage. Bobrow and Norman (1975) and Norman (1975) have used the term *schema* much as it is used in this paper. Norman, Rumelhart, and LNR (1975) have used the term *definition* where we will use *schema*. Schank and Abelson (1975) and Schank *et al.* (1975) use the term *script* to refer to one class of schemata and the term *plan* to refer to a class of somewhat more abstract schemata. Rumelhart (1975) also uses the term *schema* to refer to a set of abstract schemata similar to Schank and Abelson's *plans*.

SCHEMATA

A central theme in work of the kind referenced above is the postulation of interacting knowledge structures, which, as indicated already, we shall call "schemata." The term finds its way into modern psychology from the writings of Bartlett (1932) and it is to him that most workers acknowledge their debt. It is interesting to note, however, that in his *Critique of Pure Reason*, Kant (1787)

utilizes a notion of schema ours than is even Bartlett's.

Schemata are data structures in memory. They exist for general events, sequences of events, atomic. A schema contains information that is believed to generate a question. Schemata, in some sense, are like scripts. Although it oversimplifies the notion of a schema as analogous to a script, it is corresponding to the script instance of the concept that is related to a particular event.

There are, we believe, at least three ways to combine to make them powerful: (1) schemata have various levels of abstraction; (2) schemata represent generalizations; and (3) schemata are related to other notions. In the remainder of this paper we will be illustrating by example aspects of it so many notions are inter-

Variables in Schemata

Just as a play has roles and performances so schemata are bound by, different aspects of context of linguistics these advocates case grammars fol-

² Kant says:

[This] representation of a triangle . . . never attain that universal . . . angles . . . The schema of a triangle . . . The concept 'dog' delineate the figure of a foot . . . any single determinate figure represent in *concreto*, actual

He goes on: . . . the image of imagination; the *schema* of a triangle, and, as it were, a monogram, in accordance with which, imagination is connected with the concept

utilizes a notion of schemata that in many ways appears to be more similar to ours than is even Bartlett's.²

Schemata are data structures for representing the generic concepts stored in memory. They exist for generalized concepts underlying objects, situations, events, sequences of events, actions, and sequences of actions. Schemata are not atomic. A schema contains, as part of its specification, the network of interrelations that is believed to generally hold among the constituents of the concept in question. Schemata, in some sense, represent stereotypes of these concepts. Although it oversimplifies the matter somewhat, it may be useful to think of a schema as analogous to a play with the internal structure of the schema corresponding to the script of the play. A schema is related to a particular instance of the concept that it represents in much the same way that a play is related to a particular enactment of that play.

There are, we believe, at least four essential characteristics of schemata, which combine to make them powerful for representing knowledge in memory. These are: (1) schemata have variables; (2) schemata can embed one within the other; (3) schemata represent generic concepts which, taken all together, vary in their levels of abstraction; and (4) schemata represent knowledge, rather than definitions. In the remainder of this section we discuss these four features in turn, illustrating by example aspects of each. The section on variables is long because in it so many notions are introduced for the first time.

Variables in Schemata

Just as a play has roles that may be filled by different actors on different performances so schemata have variables that may become associated with, or bound by, different aspects of our environment on different occasions. In the context of linguistics these variables have been called "cases" by those who advocate case grammars following Fillmore (1968). We might, for example, have

² Kant says:

[This] representation of a universal procedure of imagination in providing an image

No image could ever be adequate to the concept of a triangle in general. It would never attain that universality of the concept which renders it valid of all triangles... The schema of the triangle... is a rule of synthesis of the imagination... The concept 'dog' signifies a rule according to which my imagination can delineate the figure of a four-footed animal in a general manner, without limitation to any single determinate figure such as experience, or any possible image that I can represent *in concreto*, actually presents.

He goes on: ... the *image* is a product of the empirical faculty of reproductive imagination; the *schema* of sensible concepts, such as of figures in space, is a product and, as it were, a monogram, of pure *a priori* imagination, through which, and in accordance with which, images themselves first become possible. These images can be connected with the concept only by means of the schema to which they belong.

a schema for GIVE that would have three variables: a giver, a gift, and a recipient. On different occasions the variables in the GIVE schema will take on different values. These values are determined by aspects of the environment, that is, by contextual and situational factors, as well as the to-be-comprehended stimulus. Thus, the environment provides referents for the mental conceptualizations which become associated with the variables in the schema. But, while the variables may be bound by different aspects of the environment on different occasions, still the relationships internal to the GIVE schema will remain constant. In particular, the giver will somehow cause the recipient to get the gift, and in normal cases this is true regardless of the identity of the giver or recipient, or the nature of the gift. We say "in the normal case" advisedly, for in certain cases that will be discussed later, exceptions can be found to these generalizations.

Figure 1 illustrates the Active Structural Network representation for the GIVE schema discussed above. (For a detailed discussion of active structural networks see Norman, Rumelhart, & LNR, 1975). The uncircled upper case term represents the name of the schema. The pointer labeled "iswhen" points to the internal structure of the schema. The variables X, Y, and Z are pointed to by arrows labeled "giver," "recipient," and "gift," respectively. The encircled terms represent subschemata. The arrows pointing from the subschemata show how the variables of the schema relate to those of the subschemata. Thus, the "giver" of the GIVE schema is the "agent" of the CAUSE subschema. Note also that the GET subschema plays the role of the "caused event" for the CAUSE subschema. As we shall see, actual schemata are rather more complex. The representation given here is solely for purposes of illustration.

Just as certain characteristics of the actors are specified by the playwright (e.g., sex, age, appearance), so too a schema contains, as part of its specification,

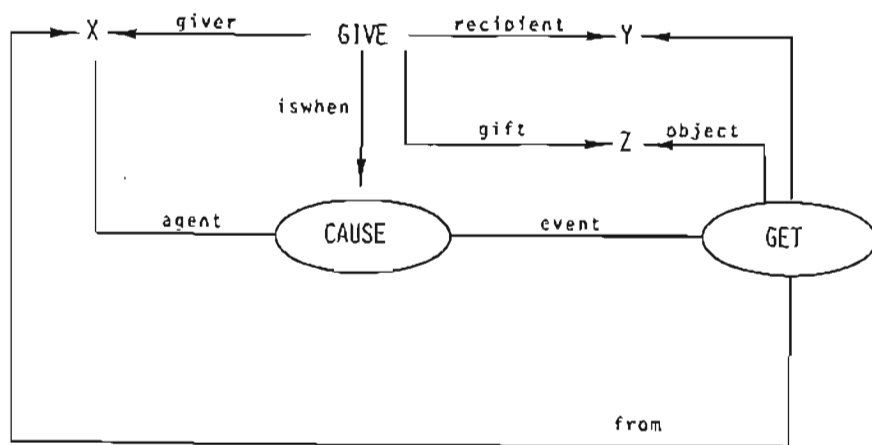


FIG. 1. Diagrammatic representation of a GIVE schema.

information about the type variables of the schema. The example, have specifications must be capable of willful action variables can take serve two objects might realistically be insufficient information they some of the variables.

The following example is binding occurs. Suppose we something. We can imagine a the breaker, the object, and expect the breaker to be an the method to be some action believed to be sufficient to be such a schema might be represented

Consider, now, the following:

- (1) John broke the window
- (2) The ball broke the window
- (3) John broke the bubble

In each case, using our analogy, an enactment of the BREAK

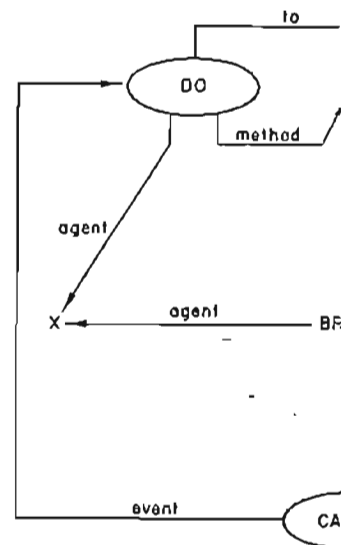


FIG. 2. Diagrammatic representation of a BREAK schema.

information about the *types* of objects that may be bound to the various variables of the schema. Thus, in our GIVE example above, we might, for example, have specifications within the schema to the effect that the "giver" must be capable of willful action (animalé?). Such constraints on the values that variables can take serve two important functions: (1) they tell what sorts of objects might realistically be bound to each variable; and (2) when there is insufficient information they can allow good guesses to be made about at least some of the variables.

The following example is intended to clarify the process whereby variable binding occurs. Suppose we have a schema for the concept of someone breaking something. We can imagine at least three variables associated with the schema: the breaker, the object, and the method whereby the object is broken. We might expect the breaker to be an agentive force, the object to be rigid or brittle, and the method to be some action of which the breaker is capable and which is believed to be sufficient to break the object in question. Figure 2 illustrates how such a schema might be represented as an active structural network.

Consider, now, the following sentences:

- (1) John broke the window.
- (2) The ball broke the window.
- (3) John broke the bubble.

In each case, using our analogy of a play, we can say that the sentence describes an enactment of the BREAK play. Nevertheless, we get quite different images of

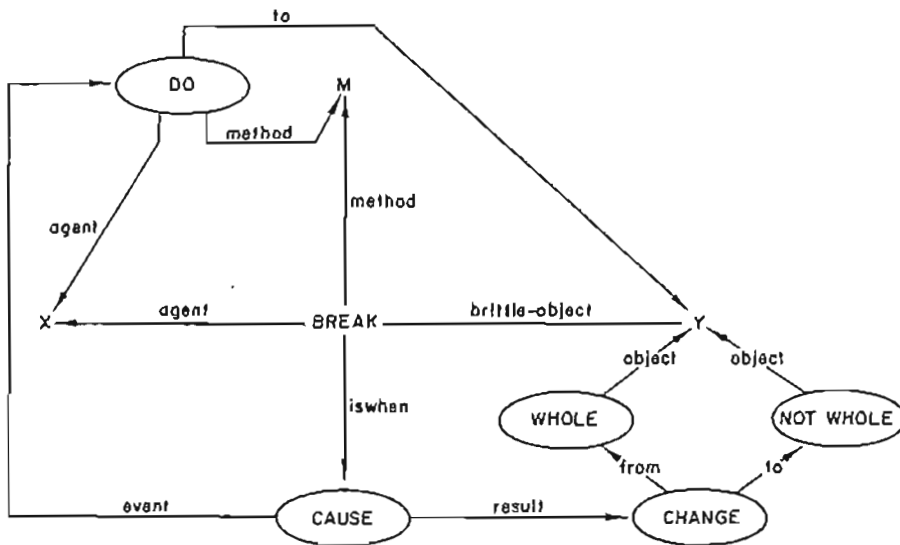


FIG. 2. Diagrammatic representaiton of a BREAK schema.

the relationships between the subjects and objects of these sentences. That is, in spite of the surface similarities among the sentences, the roles are assigned so as to produce quite different enactments of the play.

Compare first sentences (1) and (2). Suppose, perhaps, because the word "break" appears in the sentence, that we have been led to consider our BREAK schema as a possible account for these sentences. We must now somehow associate the other concepts referred to in the sentences with the variables of the schema. The variable constraints can help us do that. In both cases "the window" will probably be taken as the "object" and associated with variable Y (in the figure). The window clearly meets the criterion of being a rigid object. In sentence (1), John will also easily be determined to fit as the "breaker," X , since John is presumably the name of a person and since people are stereotypical agentive forces. However, in sentence (2) "the ball" is not easily bound to variable X because it is not easily considered an agentive force. Thus, we must make a guess about the identity of variable X . We know that X can be bound to an unspecified person or "someone." But what about "the ball." There is another variable in the break schema that can be used to account for "the ball," namely, the method. Thus, we have "someone caused the window to become broken by using the ball." But what action did this "someone" perform? We are given no direct information and must again make a plausible guess. We know from the variable constraints that whatever it was, it must have been sufficient to cause a window to break. We could leave it at that, or we could look back to the ball schema, or we could search our memories for other cases of objects like balls breaking objects like windows and see what sorts of activities were involved. Were we to carry out this extra step of inference, we probably would determine that the ball (perhaps a baseball) was somehow propelled through the window. In this way, although the particular method was nowhere stated, the variable constraints within the schema have enabled a probable value to be assigned to one of the variables. This assignment of inferred values to variables we refer to as the assignment of *default values* (Minsky, 1975). Note, however, that these default values need not be fixed independently of the values of the other variables. Instead, they are filled contingently, the value being assigned to a particular variable (such as the method in our example) depending on the value of other variables (such as the object in our example).

Now contrast sentence (1) with sentence (3). Binding the X and Y variables (the breaker and object, respectively) causes no problems. In both cases John is the breaker. In one case the object is "the window" and in the other it is "the bubble." However, we probably get quite different images of the method involved in these two cases. This difference presumably results from the knowledge we have about what sorts of activities are sufficient to break an object of the strength of a window as opposed to what sorts of activities are sufficient to break something of the rigidity of a bubble. This information could have already

been abstracted and directly discovered by consulting windows and bubbles having

To summarize, schemata constraints. These serve two values to variables by specifying in the schema. Second, when basis of the current input or default assignments. Once at the environment, from memory *instantiated*. It will transpire that of activating subschemata or modifying the original assignment schemata is akin to Craik and

Before leaving our discussion mentioned that variable constraints a variable *cannot ever accept* of variable constraints as a particular variable can take are more typical than other information has to be represented that judgments about values these values were rated as prototypical values is argued tends to require the giver to institutions can give in much are less typical. If we have closer to the "average" of other interpretation can be

In fact, the set of variables to form a multivariate distribution. Thus, as our example with we prefer values close to values of the already filled describe data showing just pairs. The extent to which depend on how probable the particular schema. One might language comprehension as (1975; Barclay, Bransford, Ford, & Solomon, 1973).

been abstracted and directly associated with the BREAK schema or it could be discovered by consulting our memories for various instances of things like windows and bubbles having been broken.

To summarize, schemata have variables with which are associated variable constraints. These serve two functions. First, they help in the assignment of values to variables by specifying the sorts of things that can fill the various roles in the schema. Second, when such an assignment cannot be made merely on the basis of the current input or from memory, the constraints can help to generate default assignments. Once an assignment of variables has been made, either from the environment, from memory, or by default, the schema is said to have been *instantiated*. It will transpire in our discussion of comprehension, that the instantiation of schemata is only the first step in comprehension. From there the process of activating subschemata or dominating schemata may continue, thereby perhaps modifying the original assignments of variables. This process of activating related schemata is akin to Craik and Lockhart's (1972) notion of depth of processing.

Before leaving our discussion of variables and variable constraints, it should be mentioned that variable constraints are seldom absolute. It is rarely the case that a variable *cannot ever* accept a value of a certain sort. Rather it is useful to think of variable constraints as representing *distributions* of possible values. A particular variable can take on any of a range of possible values, but some values are more typical than others. Empirical evidence for the view that distributional information has to be represented can be found in Walker (1975) who found that judgments about values of attributes of physical objects were quickest when these values were rated as being either extreme or typical. The primacy of prototypical values is argued for by, for example, Rosch (1973). Thus, "give" tends to require the giver to be a person, but of course governments and other institutions can give in much the same sense as a person; such values, however, are less typical. If we have a choice, the variable constraint will prefer values closer to the "average" of the distribution, but will *accept* deviant values if no other interpretation can be made.

In fact, the set of variable constraints for a given schema should be considered to form a multivariate distribution with correlations among the several variables. Thus, as our example with BREAK illustrates, when filling an unfilled variable we prefer values close to the "average" for that variable *conditional on* the values of the already filled variables. Half, Ortony, and Anderson (in press) describe data showing just this kind of context-sensitivity with adjective-noun pairs. The extent to which a particular schema fits a particular state will roughly depend on how probable that particular configuration of variables is for that particular schema. One might regard some of the experimental findings in language comprehension as mild support for this last claim (Anderson & Ortony, 1975; Barclay, Bransford, Franks, McCarrell, & Nitsch, 1974; Johnson, Bransford, & Solomon, 1973).

Schemata Embed

In much the same way as the entries for lexical items in a dictionary consist of other lexical items, so the structure of a schema is given in terms of relationships among other schemata. As we shall see, in some cases schemata can even be embedded within themselves, that is, some schemata are recursive structures. In this section, we shall restrict our discussion to simpler schemata such as those illustrated in Fig. 1 and 2 where the terms in circles are the names of embedded schemata, that we call the *subschemata*. These are represented within the schemata in which they appear, the *dominating schemata*, by names or labels, not by their entire structures. Clearly, representing the structures themselves would have the absurd consequence that every schema in memory would contain the knowledge to be found in at least most, if not all, of the other schemata in memory. This explosive multiplication of knowledge representations is arrested by incorporating only uniquely identifying references to the subschemata, for such names do not themselves incorporate other names.

Consider the FACE schema illustrated in Fig. 3a which is based on the model described by Palmer (1975). It contains within it references to schemata for eyes, ears, mouth, and so on. The substance of the FACE schema is the specification of such normal constituent parts, the subschemata, and the specification of the relationships that normally hold between them. Notice that the schema for EYE in Fig. 3b has as its subschemata those for pupil, iris, eyelid, etc., none of which appear in the FACE schema.

The overall organization which results is hierarchical, not just in the sense of a hierarchy of concepts related by class inclusion (as in Collins & Quillian, 1969), but in a more general way. This organization seems to lead to an infinite regress, in which each schema is characterized in terms of lower level constituents, or subschemata. Presumably, the dependence that schemata have on lower level subschemata must ultimately stop, that is to say, some schemata must be atomic in the sense that they are not characterized by reference to any other constituent schemata. These atomic schemata correspond to what Norman, Rumelhart, and LNR (1975) call *primitives*. Many of them probably represent basic sensory-motor procedures, while others may represent unanalyzable conceptual components of human knowledge such as that of "causal connection" which, as Hume pointed out two hundred years ago, cannot be extracted from experience alone. Thus, our entire knowledge system would appear to ultimately rest upon a set of atomic schemata.

The property of embedding which schemata have provides a number of important advantages. Foremost among these is that a situation or object can be comprehended in terms of its major constituents without necessary reference to the internal structure of the constituents themselves. Yet, at the same time, a "deeper" understanding can be achieved if reference is made to the internal structure of these constituents. Thus, for example, a face can be thought of as a certain configuration of eyes, nose, mouth, etc., rather than an enormously

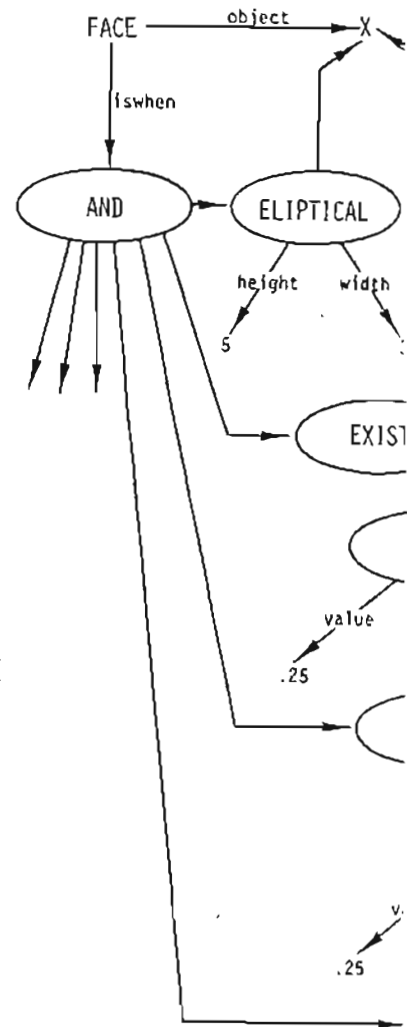


FIG. 3. Diagrammatic representation of the FACE schema. The indicated value is some standard metric. Rather than a metric concerning relative sizes and distances.

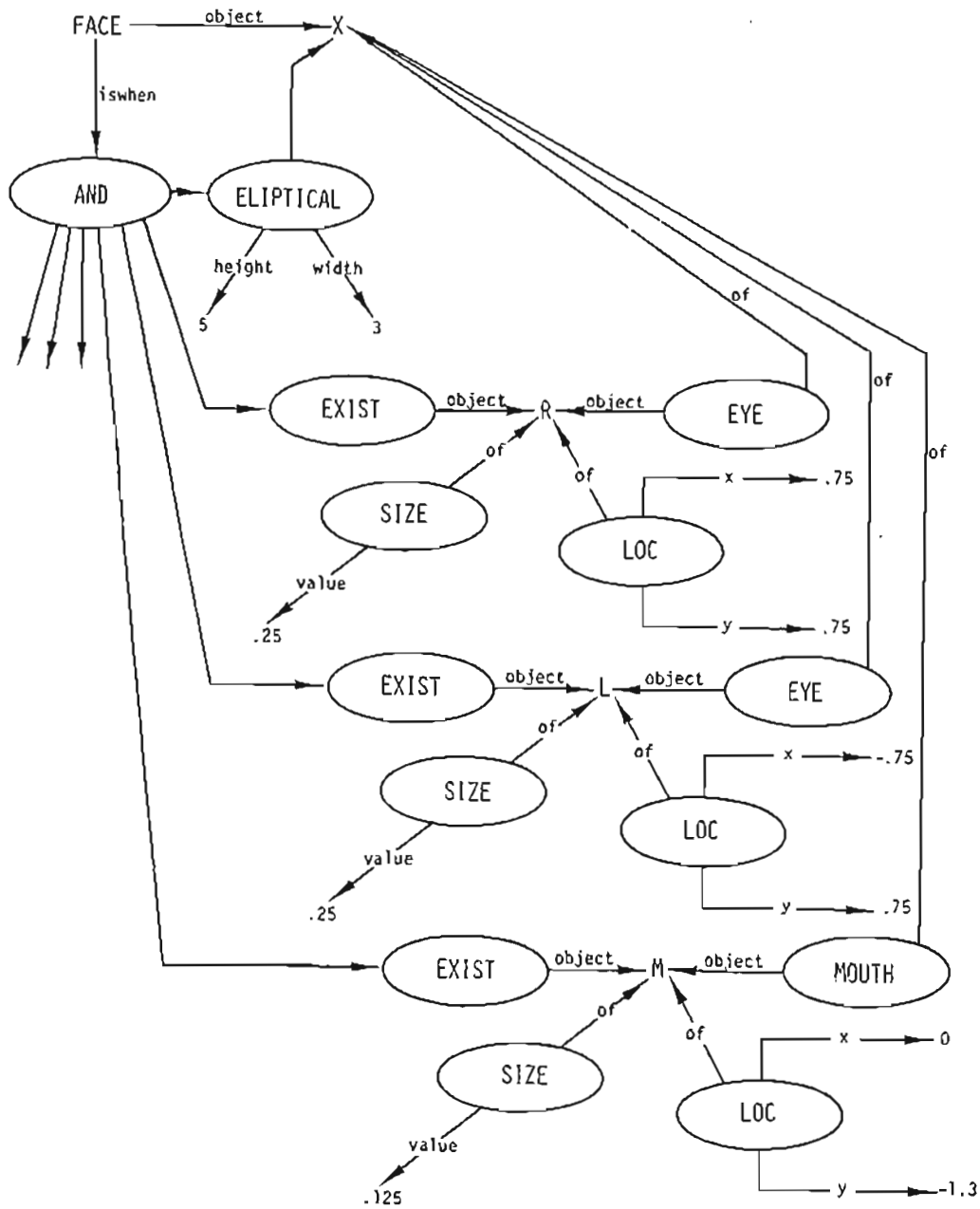


FIG. 3. Diagrammatic representation of (a) part of a FACE schema, and (b) part of an EYE schema. The indicated values should not be taken to represent distances in terms of some standard metric. Rather, they represent numerical approximations of knowledge concerning relative sizes and distances.

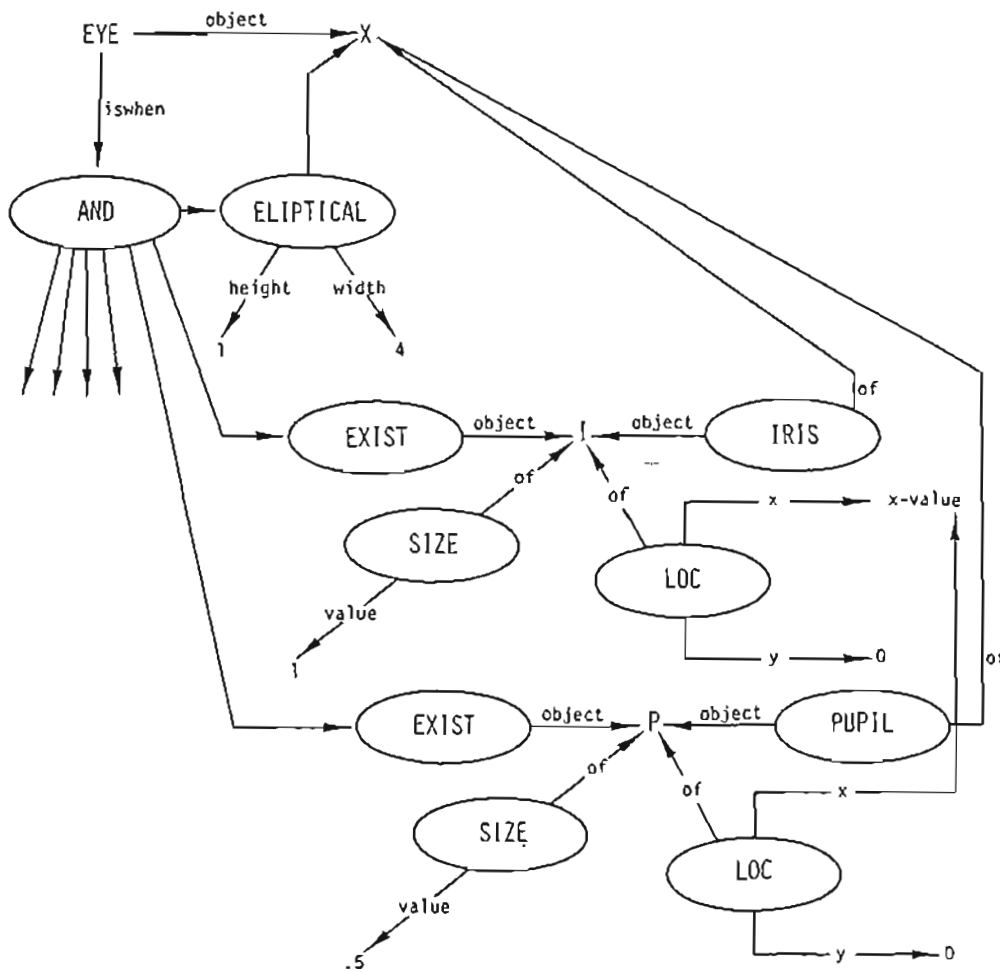


FIG 3B

complex configuration of those elementary perceptual attributes upon which the perception of a face must be presumed to ultimately depend. Similarly, the structure of a schema allows us to distinguish between the relationships that exist between the subschemata and those that exist between the constituents within any one of those subschemata. Thus, suppose that an event is observed in which two people are sitting next to one another talking. Because the PERSON schema is utilized, the feet of one person are seen as being more closely tied to the head of that person than to the feet of the second person. The event would be perceived in this way, even though their feet might in fact be physically closer than either person's feet were to that person's head. Put another way, we can say

that our knowledge about inhibits us from mindless or proximity. Objects or schemata which are being

A second advantage of optional economy of variational subschema which appear. The particular sense of "do and (thereby) cause" will include, but not be schema is much closer to or the BREAK schema. one way or another, it represent higher level action from primitives"). Thus, visible from the schema for the BREAK schema, it is some low level schema appropriate to answering instead of requiring variations and manner to be associated. kind of action, the binding schema. It may or may not that John broke a window do it?" at least as readily our ordinary language structures play. However, on conditions.

Schemata and Levels of

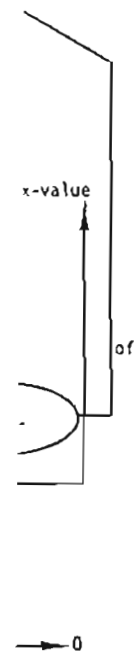
The third major characteristic section. There are several deserves special mention apart our position from general, previous works (1968; Rumelhart, Lindsay on representing the internal recently have attempted abstract levels such as: been done in this area: Rumelhart, 1975; Schacter representation sufficient

that our knowledge about people guides us in perceiving people as people and inhibits us from mindlessly grouping objects together on the basis of similarity or proximity. Objects *are* grouped together, but only on the basis of the schemata which are being employed in their interpretation.

A second advantage of embedding, related to the first, is that of representational economy of variables. In the last section we discussed a GIVE schema. A subschema which appears in it will be for one of the several senses of "cause." The particular sense of "cause" in question can be roughly translated as being "do and (thereby) cause." There will be a schema for this sense of "cause" that will include, but not be identical with, the atomic schema for "cause." The DO schema is much closer to some primitive action schema than is the GIVE schema or the BREAK schema in which we have it referenced directly. Consequently, one way or another, it will be accessible from almost all schemata which represent higher level action verbs (where "higher level" means "more remote from primitives"). Thus, sometimes, as in the GIVE schema, DO will be accessible from the schema for the appropriate sense of "cause" and other times, as in the BREAK schema, it may itself be a subschema. In any event, associated with some low level schema such as DO will be various subsidiary variables, those appropriate to answering such questions as When? Where? Why? and How? Thus, instead of requiring variables for such aspects of actions as time, place, reason, and manner to be associated with every single high level schema representing a kind of action, the binding of these variables can often take place in the DO schema. It may or may not be an interesting observation that when we are told that John broke a window, we tend to ask "When (where, why, or how) did he *do* it?" at least as readily as we ask such questions about his *breaking* it. Perhaps our ordinary language sometimes reflects the role that these underlying structures play. However, one should not exaggerate the significance of such observations.

Schemata and Levels of Abstraction

The third major characteristic of schemata emerged clearly at the end of the last section. There are schemata at all levels of abstraction. This characteristic deserves special mention because it is this aspect which most completely sets apart our position from earlier attempts to represent semantic memory. In general, previous works (see Anderson & Bower, 1973; Kintsch, 1972; Quillian, 1968; Rumelhart, Lindsay, & Norman, 1972; Schank, 1972) have concentrated on representing the internal structure of, at most, lexical items. Not until very recently have attempts been made to represent conceptualizations at more abstract levels such as action sequences or plots of stories. The work that has been done in this area is very much in its infancy (see Charniak, 1972; Rumelhart, 1975; Schank & Abelson, 1975). The need for theories of knowledge representation sufficiently powerful to be able to handle higher level conceptual-



ch the
y, the
s that
tuents
ved in
RSON
ied to
would
closer
an say

izations becomes more obvious when one considers how essential they seem to be to account for our ability to organize, summarize, and retrieve information about connected sequences of events. Particularly, if one includes the reading of connected discourse in this context, as of course one must, it would seem that such activities constitute the vast bulk of the information processing that people do, both in and out of formal educational settings. It is indeed rare that comprehension is required completely out of context, and on those occasions when it is, we seem to have an uncanny ability to construct a context within which, or perhaps, by means of which, to interpret the input.

Thus, we envision the human memory system as containing countless packets of information, each packet referring to other packets which normally form its constituents. Such packets represent knowledge at all levels of abstraction ranging from basic perceptual elements, such as the configuration of lines which form a square, to abstract conceptual levels which allow us to give cogent summaries of sequences of events occurring over substantial periods of time. We see no great discontinuity between perception and comprehension. Perception is comprehension of sensory input. Nor do we see any great discontinuity between plans and actions. Perhaps actions can be viewed merely as plans instantiated with motor values, which may themselves be either action schemata, like swimming, or primitive actions such as those required to swim. Indeed, in performance, the use of action schemata without reference to the internal structures of their constituents may be regarded as being those "automatic" performances which, Polanyi (1958, 1966) reminds us, become marred if the performer tries to attend to the constituent actions.

Schemata Represent Knowledge

In the discussion so far, reference has frequently been made to the fact that schemata represent the constituents and interrelations that are "normally" to be found. In the section discussing the embedding of schemata we noted, for example, that the FACE schema gave a specification of the "normal constituent parts" and of the relationships that "normally hold between them." We also suggested in the section on variables that the variable constraints are best regarded as distributions rather than inviolable limits. This notion seems to accord well with some recent psychological research as well as the linguistic analyses of Labov (1973) and Lakoff (1972).

If the three characteristics of schemata that we have so far discussed were the only ones, schemata would be much closer to dictionary entries than we would like to imply. To be sure, the fact that schemata can be at all levels of abstraction does exclude many of them from being candidates for dictionary entries. However, there are at least two additional reasons why it should be emphasized that schemata are not the same kinds of things as dictionary entries. First, schemata represent knowledge that is encyclopaedic rather than definitional in character, and even when "essential" characteristics are represented,

they are represented in a way that can be said to pertain. Second, while we say that schemata are "of words," schemata are frequently, they are not like dictionary entries, knowledge which we can use to look up for understanding language.

The characteristic of schemata is that they provide definitions and knowledge that is not like the knowledge of Odysseus and the Cyclops, nor do we deny that there is only one eye. Just how distorted before the scene account is largely an empirical matter rather than what is needed for typical occurrences, rather than the deviations, rather than the constraints and attempts to overcome such a way as to allow one-eyed faces can still be understood. Semantic feature theories seem to us inadequate (or at least they represent knowledge in a way that is not for vagueness, imprecision).

TH

The characterization of schemata as building blocks of the human mind we propose to elaborate on in this section. We discuss their primary role in the central mechanism. In addition, we discuss experience, and memory encoding and organizing actions.

Comprehension

Schemata are the key units in the framework presented here. They are selected and used to comprehend, and they are used to be comprehended, and they are used to be interpreted. We say that a schema is used to be interpreted as an instance of a class.

they are represented in most cases as characteristics which normally or typically pertain. Second, while dictionaries attempt to provide records of "the meanings of words," schemata represent knowledge associated with concepts. Consequently, they are not linguistic entities, but abstract symbolic representations of knowledge which we express and describe in language, and which may be used for understanding language, but which are nevertheless not themselves linguistic.

The characteristic of flexible variable constraints is a very important feature of schemata and it provides a way of seeing one of the differences between definitions and knowledge. We do not fail to understand the story about Odysseus and the Cyclops when we discover that the Cyclops is a one-eyed giant, nor do we deny that the Cyclops has a face, for a face is still a face, even if it has only one eye. Just how many, and to what extent, normal characteristics can be distorted before the schema in question no longer will provide an adequate account is largely an empirical question. But that such distortions and deviations from the typical occur is indisputable, and in representing what is normally true, rather than what is necessarily true, schemata have the capacity to tolerate such deviations, rather than to fail because of logical contradictions between variable constraints and attempted assigned values. Knowledge has to be structured in such a way as to allow that dead animals are nevertheless animals, and that one-eyed faces can still be faces. It is precisely for reasons like these that semantic feature theories, in so far as they require defining features for concepts, seem to us inadequate (see Rips, Shoben, & Smith, 1973). Schemata attempt to represent knowledge in the kind of flexible way which reflects human tolerance for vagueness, imprecision, and quasi-inconsistencies.

THE FUNCTIONS OF SCHEMATA

The characterization of schemata offered so far has portrayed them as the basic building blocks of the human information-processing system. In this section, we propose to elaborate on some of the ways in which schemata fulfill this role. We discuss their primary role of comprehension, for which we take them to be the central mechanism. In addition, we discuss their function in creating records of experience, and memories, as vehicles for inferential reasoning, and of representing and organizing action structures.

Comprehension

Schemata are the key units of the comprehension process. Within the general framework presented here, comprehension can be considered to consist of selecting schemata and variable bindings that will "account for" the material to be comprehended, and then verifying that those schemata do indeed account for it. We say that a schema "accounts for" a situation whenever that situation can be interpreted as an instance of the concept the schema represents. Thus, the

bulk of the processing in a schema-based system is directed toward finding those schemata which best account for the totality of the incoming information. On having found a set of schemata which appears to give a sufficient account of the information, the person is said to have "comprehended" the situation. In this view, when a person uses a schema to comprehend some of the aspects of the situation, the schema constitutes a kind of theory about those aspects. Thus, in general, the process of comprehension can be regarded rather like the process a scientist goes through in testing a theory; evidence is sought which either tends to confirm it, or which leads to its rejection. Upon finding a theory which, to our satisfaction, accounts for the observations we have made, we feel that we understand the phenomenon in question.

One of the most important aspects of theories that is shared by schemata is the role of prediction. We need not have performed every experiment to be able to predict with some confidence the outcome of many proposed experiments. Thus, for example, astronomers, confident in their theories, were able to predict the existence and location of Pluto before they were able to observe it. Similarly, a schema allows us to predict aspects of the input which have not been (and perhaps never will be) observed. For instance, once having determined satisfactorily that a certain object is an electric lamp, one tends to assume that it has an on/off switch even though it has not been observed. Similarly, on being told that someone went to a movie, it is normally assumed that the person, or a companion, went to the ticket window and bought a ticket prior to watching the movie. We make these assumptions because the schemata we employ in comprehending the scenes in question, or linguistic descriptions of them, predict that those aspects very probably exist by providing variables for them. It is often possible, in effect, to run the experiment by looking for the switch or inquiring about the purchasing of the movie ticket. In fact, we rarely do this, because first we have enough confidence in the schemata correctly predicting the outcomes of such "experiments," and second, the aspects of a situation which are left to be assumed in this way are usually not very important. Certain environments, such as courts of law, are interesting because very often neither of these reasons pertain.

As a means of introducing a more detailed discussion we now deal with one of the objections that could be raised against theories of the kind we are proposing. If comprehension is achieved by utilizing a schema or set of schemata to account for the input, how is the absurd conclusion that there exists a schema for every conceivable input to be avoided? The first response to this objection is to accept, as we do and must, that there are not specific schemata available for every situation we might encounter. Yet, it is equally true that we could not *understand* every situation we might encounter; indeed, in reality we do not even understand all those situations we *do* encounter. Of course, we do usually manage to achieve at least partial comprehension; that is to say, while we may not be able to find a single schema which fully accounts for some particular

situation, we may well understand some part of it. The problem, then, is to find a schema which satisfactorily accounts for the situation.

This first response, however, does not deny that we often use a specific schema for understanding a situation. A response is required, and we try to illustrate with the following examples:

- (4a) Mary heard the door open.
- (4b) She remembered the key.
- (4c) She rushed into the house.

Sentences (4a)–(4c) together with the situation of us can rather easily give rise to a schema which is along the lines that Mary had some ice cream. Buying the ice cream was the source of funds. She remembered the key, which, presumably, was used to get the money by the time she got home.

An interpretation of this schema in ordinary language interpretation we do not have a detailed account for cases of "heuristic" rather abstract "problem-solving" schema has roughly the following form:

Problem-solving schema

1. *E* causes *P* to want *G*.
2. *P* tries to get *G* until *A*.

That is, a problem solving schema for someone (*P*) which initiates an action which continues to make attempts until *G* is achieved. Suppose, further, that the goal *G* is:

Try (Person *P*, Goal *G*)

1. *P* decides on an action *A*.
2. While any condition *C* holds, *P* does *A*.
3. *P* does *A*.

³ A detailed discussion of this schema is given in Rumelhart (1975) and in Rumelhart (1975).

situation, we may well be able to find schemata to account for particular aspects of it. The problem, then, is only that we cannot find a single schema which satisfactorily accounts for the entire situation.

This first response, however, appears to leave the objection standing, for we do not deny that we often can understand situations for which the postulation of a specific schema for understanding it would be gratuitous. Thus, a second response is required, and this is that novel situations can generally be handled by using not only specific schemata, but also high level abstract schemata, as we will try to illustrate with the following example:

- (4a) Mary heard the ice cream man coming.
- (4b) She remembered her pocket money.
- (4c) She rushed into the house.

Sentences (4a)–(4c) together constitute a kind of story snippet for which most of us can rather easily get a good interpretation. Presumably, this interpretation is along the lines that Mary heard the ice-cream man coming and wanted to buy some ice cream. Buying ice cream costs money, so she had to think of a quick source of funds. She remembered some pocket money she had not yet spent which, presumably, was in the house. So, Mary hurried into the house trying to get the money by the time the ice-cream man arrived.

An interpretation of Sentences (4a)–(4c) of the kind we have given is an ordinary language interpretation, but how is it achieved using schemata? Clearly, we do not have a detailed schema like the GIVE schema or FACE schema to account for cases of "hearing ice-cream men." However, we probably do have a rather abstract "problem solving" schema which, in conjunction with a number of less abstract schemata, will account for the inputs.³ Suppose our problem-solving schema has roughly the following structure:

Problem-solving schema (Person *P*, Event *E*, Goal *G*)

1. *E* causes *P* to want *G*.
2. *P* tries to get *G* until *P* gets *G* or until *P* gives up.

That is, a problem solving episode is one in which something (*E*) happens to someone (*P*) which initiates in him a desire for something (*G*). The person then continues to make attempts to get the goal until he finally attains it or he gives up. Suppose, further, that the TRY schema has the following internal structure:

Try (Person *P*, Goal *G*)

1. *P* decides on an action *A* which could lead to *G*.
2. While any condition *A'* for *A* is not satisfied, *P* tries to get *A'*.
3. *P* does *A*.

³A detailed discussion of these problem solving schemata is presented in Rumelhart (1975) and in Rumelhart (1977a).

Thus, the TRY schema consists of three parts: deciding on an appropriate plan of action (A), fulfilling any preconditions (A') on that plan of action, and finally carrying out the plan itself.

Let us now consider how these schemata might help in the comprehension of the sentences. On encountering the first sentence, (4a), the comprehension system activates a number of schemata based on the surface clues in the sentence. (In the general case, schemata may also be activated by contextual clues, but more of that later.) In this particular case, therefore, "Mary" and something like "the coming of the ice-cream man" will be bound to variables in the HEAR schema. Similarly, a schema for ICE-CREAM MAN will be activated, which, in turn will activate its subschemata which together constitute our knowledge about ice-cream men. These will include the SELL schema with "ice-cream man" already bound to the *vendor* variable and "ice cream" bound to a variable representing "goods." However, the variables for *buyer* and *money* have yet to be bound. Meanwhile, since one of the things we know about ice-cream men is that they sell ice cream, the ICE-CREAM schema is activated, thereby making available our knowledge about ice cream. Part of this knowledge is the fact that many people like ice cream. This gives rise to the expectation of finding such a person, and the assignment of any such candidate to the *liker* variable in the now activated LIKE schema. At this point some of the unbound variables in the LIKE and SELL schemata can be identified with one another, since one ordinarily likes what one buys, the still unbound *buyer* of the SELL schema is a good candidate for the first variable of the LIKE schema, the *liker*. Similarly, "ice cream," which binds the *goods* variable in the SELL schema, is a good candidate for the second variable of the LIKE schema, the liked *object*. Thus far, the only candidate for the buyer is Mary and, "Mary" being the name of a person, the requirements for the *buyer* in the SELL schema are apparently satisfied and a tentative assignment of "Mary" to the *buyer* variable is made. At this point we have activated at least the schemata for ICE-CREAM MAN, SELL, ICE CREAM, and LIKE. Now once the LIKE schema is available, the WANT schema is activated, since we often want what we like. Thus, we are able to interpret (4a) by concluding that since ice-cream men sell ice cream to buyers, and since buyers usually like what they buy and want what they like, and since Mary is a good candidate for a buyer, Mary probably wants ice cream.

Before proceeding with the analysis, a couple of observations should be made. Even though we have taken a page or so to account for the comprehension of one simple sentence, our account is still not as detailed as we would like. We have doubtless still made some questionable assumptions and taken some dubious turns. Nevertheless, we hope that the general mechanism whereby schemata function to produce comprehension has been adequately described. It should at least be clear that the process is a complex one. The immediate result is a kind of inference, but it neither is, nor is intended to be, a deductively valid one. Rather, it relies very heavily on stereotypical, default values for variables, that is why we say things like "one ordinarily likes what one buys." Information

of this kind will be ir
reimage," the BUY sel
determined from the
earlier. Such assignn
tion. They combine :
certain sequel senten
followed by Sentence

(5) She drew her

we suspect that, in
reinterpretation woul
in Sentence (4a) as a r

Returning now to o
desire for ice cream.
solving schema, beca
Consequently, among
activated. This schema
fore. In this case, the
sufficient cause. Thus,
become bound by the
by "Mary" and the G
solving schema tells us
evidence for Mary try
processing is complet
pocket money, and we
least related to an att
need to make use of ou
on a plan of action whi
already have an approp
of the ice cream from t
choice is supported by
needs money; Mary rem
to buy the ice cream. V
some ice cream, or abou
Thus, when in Sentence
that she was intending t
these associations and up
with the interpretation o
snippet by settling upon
which appear to account
schema for dealing with t

The abstract schemata
Sentences (4a)-(4c) are r

of this kind will be incorporated into the SELL schema (and also into its "mirror image," the BUY schema), but its incorporation will be as a default assignment determined from the variable-constraining distributional information alluded to earlier. Such assignments are only made in the absence of conflicting information. They combine to give rise to probable interpretations which would render certain sequel sentences, at least, unexpected. Thus, were Sentence (4a) to be followed by Sentence (5)

(5) She drew her revolver and shot him.

we suspect that, in the absence of any contextual cues, a certain degree of reinterpretation would be required to modify some of the variable findings made in Sentence (4a) as a result of the sequel, Sentence (5).

Returning now to our example, we have concluded that Mary probably has a desire for ice cream. It is at this point that we invoke our abstract problem-solving schema, because desires are often parts of problem-solving episodes. Consequently, among others, the problem-solving schema outlined above will be activated. This schema requires an event which causes the desire to come to the fore. In this case, the event of becoming aware of the ice-cream man's arrival is a sufficient cause. Thus, the event variable *E* of the problem solving schema would become bound by the event of Mary hearing the ice-cream man, the *P* variable by "Mary" and the *G* variable by "ice cream." The second line of the problem-solving schema tells us that *P* tried to get *G*. We can therefore expect to find evidence for Mary trying to get the ice cream. Probably even before this processing is completed we have taken in the next sentence, (4b), about Mary's pocket money, and we now have a good reason for trying to interpret this as at least related to an attempt by Mary to get the ice cream. In order to do this, we need to make use of our TRY schema which reveals that the first step is deciding on a plan of action which could lead to the accomplishment of the goal. Now we already have an appropriate schema available, namely, that of Mary as the buyer of the ice cream from the ice-cream man (the instantiated SELL schema). This choice is supported by the evidence provided by Sentence (4b), for the buyer needs money; Mary remembers her pocket money, so she has probably decided to buy the ice cream. We now expect to find input either about Mary buying some ice cream, or about her trying to satisfy some precondition for so doing. Thus, when in Sentence (4c) we see Mary going into the house we may conclude that she was intending to get her pocket money from the house. Upon making these associations and upon finding no more input, we conclude our processing with the interpretation outlined above. That is, we have comprehended the story snippet by settling upon a configuration of schemata and their variable bindings which appear to account for all aspects of it, even though we had no single schema for dealing with the particular event described.

The abstract schemata that were used to illustrate the comprehension of Sentences (4a)–(4c) are really very general. They can be used to cover very many

instances of trying and problem solving. Consider the following somewhat more complex example from Schank and Abelson (1975):

- (6a) John knew his wife's operation would be expensive.
- (6b) There was always Uncle Harry. . . .
- (6c) John reached for the suburban telephone book.

Again, these sentences can easily be interpreted using the same two abstract schemata. John's awareness that his wife's operation would be expensive binds the *event* variable in the abstract problem-solving schema. John is thereby caused to want something, namely money, and he therefore tries to get it. He decides to borrow it from Uncle Harry. A condition on borrowing is asking, and asking requires contacting. One way to contact is by phone. A phone book is used as a part of telephoning. Thus, John is trying to phone Uncle Harry to ask for the money.

If comprehension does indeed proceed in the way we have suggested, one might suppose that it would be possible to produce a computer simulation of the process. In fact we know of cases in which the details of the operation of such a system have been adequately specified to determine whether or not comprehension can be so achieved. Schank *et al.* (1975) have developed a computer system called SAM which can apply schemata at an intermediate level of abstraction and one of us (DER) has developed a computer system called STORYWORLD which is able to apply very simple versions of the problem-solving schema just mentioned. Still, no one has yet developed a processing system of sufficient sophistication nor a knowledge base so rich that we can say with certainty that these proposed mechanisms will work at the level we are suggesting.

Schemata and Memories

A complete discussion of the representation of knowledge cannot restrict itself to *generic* knowledge alone. It must deal not only with what has traditionally been called *semantic* memory but also with what Tulving (1972) called *episodic* memory. The content of episodic memory, episodic knowledge (see Ortony, 1975a) is more specific, being those memories for particular events which we have directly or indirectly experienced. By contrast, generic knowledge is the knowledge that we have of concepts, abstracted from such memories. In this section we discuss the relationship between episodic knowledge, memories, and schema theory.

In a sense, our memories are natural side effects of the comprehension process. In comprehension various aspects of the input are associated with a configuration of schemata, and these instantiated schemata constitute our *interpretation* of the input. What gets stored in memory is, in effect, a copy or partial copy of these instantiated schemata, that is, what gets stored is not the input itself but the interpretation that was given to that input as a result of the comprehension

process. In fact, these originally instantiated schemata become inaccessible to the copies of the original fragments to try to recover" the input situation utilizes schemata to access information from a continuum between "imposition of constraints," and in the latter on "memorial fragments emphasized that although memory as the modalities not merely fragments of our interpretation.

Notice that this view "distortion" in our reconstruction involves a "fill in" which invariably allows some hand, our reconstruction imposition of yet a second experiment reported by processes.

Having already suggested standing, we will now discuss them. In our discussion process required only cannot be right, for the demands not only on example, when we hear require reference to old schemata if it is to be may well require reference of the *given-new* distinction certain variables are bound to aspects of our through such bindings: our memories, thus processes become directly related.

process. In fact, these memory traces are probably not complete copies of the originally instantiated schemata, but a more or less complete set of fragments of them. Perhaps such partial information storage results from some incompleteness in the original copying process due to time pressures or some inherent difficulties in the process. Perhaps various aspects of the memory trace decay, or become inaccessible over time. In any event, after some time only fragments of the copies of the originally instantiated schemata remain and we must use these fragments to try to reconstruct the original interpretation and thereby "remember" the input situation. This reconstruction is not, however, unguided, but utilizes schemata to assist in interpreting the fragments, just as comprehension utilizes schemata to assist in interpreting the sensory inputs. There is thus a kind of continuum between understanding and remembering, where in the former we have the imposition of an interpretation primarily on incoming "sensory fragments," and in the latter, we have the imposition of an interpretation primarily on "memorial fragments." In both cases schemata are employed. It should be emphasized that although remembering can be thought of as perceiving with memory as the modality, the episodic memories on which it is usually based are not merely fragments of the initial sensory input, but a fragmentary representation of our interpretation of that input.

Notice that this view of memory yields two sources of "importation" and "distortion" in our recollections. On the one hand, the initial comprehension process involves a "filling out" of the original sensory skeleton — this filling out invariably allows some latitude on the part of the understander. On the other hand, our reconstruction of the original interpretation may very well lead to the imposition of yet a slightly different interpretation. It seems to us that the experiment reported by Spiro (1975) provides support for both of these sorts of processes.

Having already suggested a close relationship between remembering and understanding, we will now introduce another very important connection between them. In our discussion of comprehension, it may have seemed that the entire process required only the generic knowledge captured by schemata. But this cannot be right, for the interpretation of the outside world very often makes demands not only on generic knowledge but also on specific memories. For example, when we hear a sentence, although some of it is *new* and does not require reference to old memories (although it does, of course, require existing schemata if it is to be understood), some of it is presupposed to be *given* and may well require reference to stored memories (see Clark, 1973, for a discussion of the *given-new* distinction in linguistic inputs). Thus, when binding variables, certain variables are bound to aspects of the current situation, other variables are bound to aspects of our memories of a related situation. It is presumably through such bindings that new information becomes interrelated with old in our memories, thus providing a way for memories about particular events to become directly related to one another.

Finally, the fact that our memories are representations of interpreted inputs rather than of inputs themselves has some important consequences for retrieval. Since the particular schemata which will be activated at the time of comprehension depend not only on the input but also on the context, different contexts may give rise to different patterns of schemata available for comprehension even though the input be the same. A second presentation of the input (or part of it) will tend to be helpful as a retrieval cue to the extent that it can be interpreted in the same way as the original. Consequently, changes in the contextual conditions prevailing at retrieval time, compared with those at the time of presentation, may result in a failure to recognize the second presentation as being the same item as the original. For the same reason, fragments of the original presented as retrieval cues may also be relatively ineffective. In this way schema theory accounts for the *encoding specificity* results of Tulving and Thomson (1973) and others.

Making Inferences with Schemata

We have discussed the use of schemata for comprehension, storage, and retrieval of input information. In addition to these, schemata serve an important function as powerful devices for making inferences.

Perhaps the most obvious way schemata serve to make inferences is as predictors of as yet unobserved input. Upon finding a schema which gives a good account for an input situation, we can infer likely aspects of the situation which we have not observed. Thus, if someone tells us that he went to a restaurant for dinner, we can *infer* that he was probably given a menu, gave his order to the waitress, and paid for the meal after eating. We can make such inferences because the RESTAURANT schema has things like *dining* (see Schank & Abelson, 1975, for the specification of a RESTAURANT DINING schema) as subschemata. The activation of such subschemata and their constituents serves as a vehicle for such inferences. Related to this is another inferential process that we have already discussed, namely, inferring the existence of a whole from the observation of a part. Thus, for example, if we see an eye, we can often infer the existence of a face. This inference is involved in the natural course of comprehending an input, and comes about by the activation of dominating schemata by their subschemata. A third kind of inference already discussed involves filling unspecified variables. The variable constraints along with our knowledge of particular cases allow us to make rather good guesses about unspecified variables by assigning typical default values. Thus, in our *break* example, we make a guess about the sort of action that was involved in breaking the bubble without actually having been told.

In addition to these sorts of inferences which naturally occur during comprehension, schemata are useful for what Collins, Warnock, Aiello, and Miller

(1975) call *functional* question (7a) cited by

(7a) Is the Chaco there.

(7b) I think it's sense I guess

To account for Senter goods" (a not unreasonably many claim which determine what Question (7a) can be given the schema are fixed then be initiated for question have the same reveals a candidate, on *productis* variable. In the determined that since that cattle were raised Collins and his colleagues same reasoning process fill certain variables of variables and assume instance we found. The unspecified variables be can be applied as a general

Functional reasoning an important part in a ogy problem:

(8) Neil Armstrong

A schema-based system memory in which the Armstrong," the other plier Columbus" for "and search memory (most) other variables the value of the variable schema relating "Neil repeated. Presumably exploratory expedition

(1975) call *functional reasoning*. Consider Sentence (7b), which answers the question (7a) cited by Collins *et al.* as an example of functional reasoning:

(7a) Is the Chaco the cattle country? I know the cattle country is down there.

(7b) I think it's more sheep country. It's like western Texas, so in some sense I guess it's cattle country.

To account for Sentence (7b) we can postulate a schema for "producing farm goods" (a not unreasonable assumption for students of geography). There are presumably many climatic variables such as temperature, rainfall, and vegetation which determine what agricultural products can be produced. The answer to Question (7a) can be generated if one assumes that first the climatic variables of the schema are fixed by being bound by their values for Chaco. A search can then be initiated for a related schema or memory in which the variables in question have the same or comparable values. If such a search successfully reveals a candidate, one can check to see if "cattle" is a value of its *agricultural products* variable. In the case of Sentence (7b) the person answering presumably determined that since Western Texas matched on values of climatic variables and that cattle were raised there, indeed they might well raise cattle in the Chaco. Collins and his colleagues give a number of similar examples which illustrate the same reasoning process. It appears, then, that a typical reasoning strategy is to fill certain variables of a schema and then search for cases which match those variables and assume that the unspecified variable has the same value as the instance we found. This is exactly the same process that we discussed for filling unspecified variables by consulting episodic memory. Apparently, this process can be applied as a general reasoning strategy.

Functional reasoning is a kind of analogical inference. Schemata seem to play an important part in explicit analogical reasoning. Consider the following analogy problem:

(8) Neil Armstrong is to the moon as Christopher Columbus was to what?

A schema-based system would have to first find an instantiated schema, or memory in which there were at least two variables, one bound by "Neil Armstrong," the other bound by "the moon." It would then substitute "Christopher Columbus" for "Neil Armstrong," replace "the moon" by a free variable, x , and search memory for another instantiated schema which matched on all (or most) other variables (not values). At that point, it would permit as a response the value of the variable, x . In case this procedure failed to find a match, a new schema relating "Neil Armstrong" to "the moon" can be found and the process repeated. Presumably, we would eventually find that Neil Armstrong led an exploratory expedition to the moon and that Christopher Columbus led an

exploratory expedition to America. Thus, we could produce the answer "America."

In addition to the particular inferential procedures such as those just described, there are also much more abstract *reasoning schemata* which will allow conclusions to be drawn from premises. In many cases, such schemata will be general statements of the kinds of rules frequently found in logic textbooks. Like other schemata, they may contain variables which get bound when the schemata are utilized; and like still other schemata, they can vary in their level of abstractness. Thus, while there is probably a general-purpose TRANSITIVITY schema with complicated variable constraints, there may also be more specific ones in which the *relation* variable is fixed. Thus, for example, if the *relation* variable in the transitivity schema is fixed for "physical cause" we would get the special-purpose CAUSAL TRANSITIVITY schema below:

Causal Transitivity (event E_1 , event E_2 , event E_3).

1. If E_1 causes E_2 and E_2 causes E_3 , then certainly E_1 causes E_3 .

The use of such a schema would be identical to that of schemata in general. It would be activated at appropriate times (see the section on processing principles for a more detailed discussion of activation of schemata) and once activated, available candidates for binding the variables would be sought. Thus, the reasoning strategies which people normally employ can readily be incorporated into a schematic representation of knowledge such as the one we propose. These principles may not only include, but may also transcend, the "laws of thought" which are described in textbooks. Finally, at a more general level, it is interesting to note that one can regard the entire comprehension process in schema theory as itself being a case of analogical reasoning. When we determine that a situation fits a certain schema we are in a sense determining that the current situation is analogous to those situations from which the schema was originally derived. Moreover, when we make inferences about unobserved aspects of the situations we are, in effect, assuming their existence *by analogy* from the situations from which the schemata were derived.

Schemata and the Structure of Actions

At the end of the section on the characteristics of schemata we indicated that schemata can also constitute the underlying knowledge used to perform actions. We turn now to a more detailed examination of this idea.⁴

⁴The application of schemata to actions is not new. Bartlett (1932) suggested that we have motor schemata for such activities as playing tennis. Much more recently, Schmidt (1975) has proposed a schema theory of motor skill learning which further develops Bartlett's notions. Many of the particular ideas discussed in this section derive from those

Most people know how to solve a problem that concerns actions. What characterizes a TRANSFER action? The following schemata, it has variable

TRANSFER (object O)

1. TOSS O from H_i to H_f
2. CATCH O with H_f

H_i is taken to be the initial hand position (e.g., the hand holding the object) and H_f is the hand which catches the object. In this case, then, TRANSFER consists of two subschemata – TOSS and CATCH – and a complex schemata with a specific configuration of subschemata and the following internal structure:

CATCH (object O , with hand H_f)

1. POSITION H at IN
2. When O contacts H

The invocation of the CATCH schema (the thing we might call the "point schema"). The TRANSFER schema variables to be fed back to the CATCH schema. POSITION SCHEMA is a subschema. In sum, the fine tuning of the CATCH schema is under the control of a TRANSFER schema. (Petrie, 1974; Powers, 1975)

Of course, the TRANSFER schema is an abstract action schema. It involves an object in each hand and a hand on the other hand:

EXCHANGE (object O , hand H , hand T).

1. TRANSFER O_1 from H to T
2. TRANSFER O_2 from T to H

discussed by D. A. Norman, Department of Psychology, the University of California, San Diego, 1975.

Most people know how to toss an object from one hand to another. The problem that concerns us is how to represent that knowledge. How do we characterize a TRANSFER schema that organizes and coordinates the set of actions involved? The schema below could serve this purpose. Like all other schemata, it has variables and subschemata:

TRANSFER (object O , from hand H_i , to hand H_f , at time T).

1. TOSS O from H_i to location (H_f) at T .
2. CATCH O with H_f at $T + \delta T$.

H_i is taken to be the initial hand which holds the object (O) at the outset (time T) and H_f is the hand which finally holds the object as a result of the transfer. In this case, then, TRANSFER is assumed to have four variables and two embedded subschemata – TOSS and CATCH. TOSS and CATCH themselves are of course complex schemata with variables, and they are, in turn, represented by a configuration of subschemata. Thus, for example, CATCH may have roughly the following internal structure:

CATCH (object O , with hand H).

1. POSITION H at INTERCEPTION-POINT (of O with H).
2. When O contacts H , GRASP O with H .

The invocation of the CATCH schema will result in the activation of something we might call the TRAJECTORY schema. (by way of the interception-point schema). The TRAJECTORY schema will enable the values of some of its variables to be fed back to the CATCH schema. This, in turn, would allow the POSITION SCHEMA to move the hand closer to the interception point. Ultimately, the fine tuning which takes place is probably best regarded as being under the control of a negative-feedback system with the perceived disparity between the hand position and the object's position being successively reduced (Petrie, 1974; Powers, 1973).

Of course, the TRANSFER schema can itself serve as a constituent of more abstract action schemata. Consider, for example, the case in which we have an object in each hand and want to EXCHANGE the objects by tossing each to the other hand:

EXCHANGE (object O_1 , with object O_2 , from hand H_i , to hand H_f , at time T).

1. TRANSFER O_1 from H_i to H_f at T .
2. TRANSFER O_2 from H_f to H_i at APEX (O_1).

discussed by D. A. Norman, Ross Bolt, and other members of the LNR Research Group at the University of California, San Diego, during a number of research meetings in the fall of 1975.

The APEX schema enables the determination of a region within which the object is at its highest point. Clearly, some kind of distribution of positions will have to be represented, although the coordination of transfers does not depend upon an object being exactly at the highest point. The initiation of step 2 in the EXCHANGE schema could thus vary within limits. An interesting feature of the APEX schema is its subtle blend of cognitive and motor aspects; it involves understanding or interpreting perceptual inputs and perceptual tracking. We doubt that any sharp separation of action schemata from those we have discussed already as means for interpreting inputs can usefully be made.

Directly or indirectly, the action schemata we have just described all find their place as subschemata within a yet more complex schema, namely, juggling by the cascade method. Assuming that of the three objects, two (O_1 and O_2) start off in the right hand (H_r) while the third (O_3) starts in the left (H_L) we have a JUGGLE schema as follows:

JUGGLE (object O_1 , object O_2 , object O_3 , at time T).

1. EXCHANGE O_1 with O_3 from H_r to H_L at T .
2. JUGGLE O_2 , O_3 , O_1 at APEX (O_3).

Juggling is thus represented recursively as invoking first an EXCHANGE of two objects and then, as EXCHANGE is being completed, reinvoking the JUGGLE schema, which in turn initiates a new EXCHANGE, and so on.

With these examples we have tried to show that knowledge underlying the performance of actions can be represented in the same way as knowledge underlying comprehension. For clarity of exposition, we can distinguish between these two kinds of knowledge as being based on action schemata and comprehension schemata. At the same time it should be emphasized that these schemata are almost always highly interdependent. The coordination of many actions requires an interpretation of perceptual cues which are often selected because activated comprehension schemata are purposefully "looking for" variables relevant to the action in question. The interdependence of action and comprehension schemata in the other direction was discussed when we dealt with the TRY schema which has an action schema as a constituent.

Schemata appear to handle the representation of actions and action sequences rather naturally because the basic characteristics of schemata map conveniently onto some of the crucial characteristics of actions. In the first place, the existence of *variables* in action schemata permit the flexibility required for the performance of actions. Thus, when we shoot a basketball, we are probably doing so from a position on the floor from which we have never before shot. Nevertheless, we make some estimate of the distance and angle (both variables in a schema for shooting basketballs) and thus determine how forcefully and in what direction the shooting should be initiated. Secondly, the *embedding* of action schemata within one another also captures some important intuitions

about actions, in particular about juggling. Although juggling is a difficult task, it is learned faster if the subaction schemata are mastered. The constituents of juggling are mastered by jugglers. Finally, the embedding principle allows us to understand action schemata and their finger twitches.

SCHEMA

For the most part, our cognitive tools that explain whereby new schemata are learned, the nature of schema specialization, the nature of new schemata can be predicted by mechanisms, specialization kinds of learning.

Schema Specialization

Schema specialization occurs to form a less abstract schema, as illustrated in Fig. 2 will be an instance, to fix the *object* variable to "bubble" would be other variables, such as time. The WINDOW schema, and that was used in the initial cream, Sentences (4a) - CREAM FROM AN ICE CREAM SCHEMATA. The concepts for which there are simple The fact that schema specialization are frequent

about actions, in particular, the fact that they have constituent structures. Thus, although juggling is a single action, it does have complex subactions as constituents. Papert (personal communication) reports that people learn to juggle faster if the subactions corresponding to our TRANSFER and EXCHANGE schemata are mastered first. This suggests that these subschemata are *real* constituents of juggling in spite of the apparent unity of the actions for skilled jugglers. Finally, the existence of action schemata at *all levels of abstraction* in principle allows us to account for the relations between *plans* (very abstract action schemata) and the execution of those plans, even down to the smallest finger twitches.

SCHEMA ACQUISITION AND MODIFICATION

For the most part, our discussion to this point has taken schemata as givens, cognitive tools that exist from the start. We have postulated no mechanisms whereby new schemata can grow and old ones evolve. Indeed, this is a central problem for schema theories and very little work has been done on it. Nevertheless, the nature of schemata suggests a number of plausible mechanisms whereby new schemata can be produced. In this section we will concentrate on two such mechanisms, specialization and generalization, both of which can be regarded as kinds of learning.

Schema Specialization

Schema specialization occurs when one or more variables in a schema are fixed to form a less abstract schema. The BREAK schema discussed earlier and illustrated in Fig. 2 will serve as an example. It would be quite possible, for instance, to fix the *object* variable, *Y*, to "bubble." Since, as was mentioned in the discussion on variables, the variable constraints interact, fixing the *object* variable to "bubble" would have repercussions for the constraints associated with other variables, such as the *method* variable. Thus, the original BREAK schema could be specialized to produce a new BREAK BUBBLE schema, or a BREAK WINDOW schema, and so on. Similarly, the abstract problem-solving schema that was used in the interpretation of the sentences about Mary and the ice cream, Sentences (4a)–(4c), could be specialized to produce a BUY ICE CREAM FROM AN ICE-CREAM MAN schema. Notice that there are no constraints on the complexity of our ordinary language descriptions of schemata. The concepts which schemata represent are not restricted to concepts for which there are simple lexical items in the language.

The fact that schema specialization can occur tells us nothing about the circumstances under which it does occur. Presumably the criteria for schema specialization are frequency and utility. If a schema is frequently used with the

region within which the object
of positions will have to
ofers does not depend upon an
initiation of step 2 in the
s. An interesting feature of the
nd motor aspects; it involves
and perceptual tracking. We
ata from those we have dis-
usefully be made.

ive just described all find their
schema, namely, juggling by
bjects, two (O_1 and O_2) start
arts in the left (H_L) we have a

ic T).

first an EXCHANGE of two
ted, reinvoking the JUGGLE
and so on.

at knowledge underlying the
he same way as knowledge
i, we can distinguish between
ction schemata and compre-
phasized that these schemata
ordination of many actions
h are often selected because
lly "looking for" variables
ence of action and compre-
sed when we dealt with the
uent.

ctions and action sequences
schemata map conveniently
ns. In the first place, the
flexibility required for the
basketball, we are probably
we have never before shot.
and angle (both variables in
nine how forcefully and in
secondly, the *embedding* of
some important intuitions

same values assigned to some of its variables then the generation of a more specialized schema with those values fixed may occur. At the same time, some schemata may be so general that their utilization involves a great deal of work and leaves a great deal of uncertainty as to the probability of default assignments fitting. Since schema specialization constrains the default assignments and reduces the amount of work to be done, its use may be more effective. Thus, in the ice-cream example, a good deal of processing was required to determine that the problem-solving schema should be invoked and exactly how the variables within the subschemata should be bound together. Were we to construct a more specific schema much of this processing could be bypassed.

Consider another more extreme example of the usefulness of this specialization process. Suppose we have a schema for a *thing* or physical object. Suppose that among its variables were, its name, and a list of its properties. If it turned out that a significant number of those properties correlated highly with the "name" of the concept it may well be useful to build a specialized schema for that subset of *things* which went by that particular name. It might well be that the prior existence of a number of such abstract schemata coupled with the machinery for specializing these schemata might be enough to account for all of our schemata. A final example of the potential role of schema specialization in learning comes from the learning of motor skills and the operation of action schemata. Consider what happens when we learn, say, how to throw a dart at a bullseye. At first we invoke a rather general THROW schema and attempt to determine the proper variable values for throwing particular kinds of darts particular distances. The THROW schema is very general, but there are many variables to set and our ability to set these properly may not be great. However, once we have thrown a particular dart several times we become increasingly better at determining the proper angle, amount of force, etc., for throwing it. It would thus make sense that we might well build these values into a THROW DART schema which could later be called upon when we again want to throw darts. The tremendous savings on relearning motor skills would appear to support such a view.

A final point should be made about specialization and that concerns the question of the storage/processing tradeoff. One of the virtues of relatively general schemata is that they are able to assist in the comprehension of a diverse array of inputs. Specific schemata, on the other hand, provide a faster, more detailed interpretation of a smaller range of inputs. If we allow the generation of too many specialized schemata, the differences between them may not be sufficient to enable the correct ones to be isolated. Consequently the processing saved in comprehension may be taken up in selection, and little would be gained. In general, the more structures there are in memory, the greater are the storage demands and processing time for selection. For this reason, the production of specialized schemata has to be limited to cases where a reasonable payoff can be expected.

Schema Generalization

Schema generalization is, of course, exactly the converse of specialization. That is, some fixed portion of an old schema is replaced with a variable to construct a new and more abstract schema. This mode of learning would be especially useful when we have a case in which no schema fits a particular situation exactly, but one schema gives a very close fit to the situation except for one aspect. If we repeatedly encounter such a situation we might very well construct a new schema similar to the old one, but in which the troublesome constant has been replaced by a variable. The variable constraints for this new schema would presumably be determined by the distribution of values we actually observed which forced us to create this new variable.

Again the BREAK schema from Fig. 2 will be useful for illustration. One might decide that an object's being rigid or brittle only represents a subset of particular cases and that the object should have more general characteristics. Suppose that the object were constrained so that it either possessed a natural structure, or that it was normally used to perform some function. Now the required action would be different. It would have to be some action which was capable of destroying the natural structure, or fouling up the normal performance of the object's function. If this were the case, one could view the BREAK schema in Fig. 2 as a particular case of a more general BREAK schema. The virtue of the more abstract schema would be that it could handle other cases of breaking as in Sentences (7) and (8):

- (7) John broke the sewing machine.
- (8) John broke his promise.

In these cases our original BREAK schema would be inadequate, more so perhaps for Sentence (8) than for Sentence (7). In the case of Sentence (7), we clearly have some notion of destroying the normal function involved in our understanding of it, for a broken sewing machine normally is a machine that no longer functions properly, rather than a physically mutilated, twisted, unrecognizable lump of metal. In the case of Sentence (8), a promise can be viewed as serving a social function of providing a certain kind of commitment or guarantee. The original examples (1), (2), and (3) can also be handled by the more general schema. In the case of Sentences (1) and (2), both concerned with the breaking of a window, the normal function of the window can be regarded as having been fouled, indeed, one might well maintain that the difference between a broken window and a cracked one lies in just this fact. In the case of Sentence (3), the natural, generally convex structure of the bubble, is destroyed. All that would be required for the more general schema to be applied would be knowledge of the natural structure or normal functions in the schemata representing the objects in question.

If the more general BREAK schema we have just sketched can be used to account for a greater range of "breakings" than the more specific one of Fig. 2, the question again arises as to whether and why we need both. Again, the answer is that it depends on whether or not the more specialized BREAK schema is sufficiently useful sufficiently often for it to be stored as a separate schema. The particular distinctions within schemata that an individual has, will depend on their utility for that individual. It would be entirely reasonable to expect a football player to have a special schema for the concept of "breaking a tackle," separated from his general purpose BREAK schema. Such a specialized schema would incorporate a great deal of more specific knowledge and would presumably be used rather frequently.

The GIVE schema shown in Fig. 1 can also be used to illustrate the point. Whereas the most common sense of the word "give" concerns causing something to change possession, when someone "gives you trouble" there does not seem to be any object changing possession at all. Yet, one does end up "having the gift" when someone gives it to you. It would thus appear that the sense of "give" in "giving trouble" is a kind of generalization from a more specific, and probably prior, GIVE schema. Thus, generalized schemata may also constitute a means for interpreting what Gentner (1975) calls "metaphorical extensions."

The importance of generalization of schemata for learning is obvious. Schemata need to be generalized to the extent that they permit the interpretation of the inputs to the system. Thus, a great deal of learning may be dealt with by supposing that when a radically new input is encountered, a schema without variables is constructed. Then, when comparable inputs are encountered, which are sufficiently close to the original schema, a new one is created in which the differences become variables and the consistencies get built into the structure. In the other direction, more general schemata may be acquired as a result of learning, for example, general principles, and such schemata may become more specialized as the range of their application becomes more apparent.

In addition to these two modes of schema formation, there appears to be one other related learning mechanism natural for schema theories. This is related to, but not identical to, the generalization mechanism. Suppose we encounter a situation in which we cannot find a schema which will account for the entire configuration of subschemata we have discovered. In this case, we can store our partial interpretation of the situation, a number of unrelated aspects. If, subsequently, we encounter very similar configurations of schemata for which we again can find no overall schema, we might well build a new schema whose internal structure matches the similar aspects of the configurations and whose variables match the variable portion of these situations. In this way, we can find repeatedly co-occurring configurations of schemata and thereby gain a specification of a new, more abstract schema.

Before leaving our discussion of schema change, it should be noted that it seems reasonable to suppose that not only can new schemata be grown (by the

just sketched can be used to the more specific one of Fig. 2, we need both. Again, the answer a specialized BREAK schema is stored as a separate schema. The individual has, will depend on entirely reasonable to expect a concept of "breaking a tackle," schema. Such a specialized schema specific knowledge and would pre-

be used to illustrate the point. "ive" concerns causing something trouble" there does not seem to he does end up "having the gift" appear that the sense of "give" in a more specific, and probably may also constitute a means for ical extensions."

ata for learning is obvious. hat they permit the interpreta- al of learning may be dealt with encountered, a schema without inputs are encountered, which ew one is created in which the s get built into the structure. In y be acquired as a result of h schemata may become more es more apparent.

nation, there appears to be one ma theories. This is related to, ism. Suppose we encounter a ich will account for the entire . In this case, we can store our of unrelated aspects. If, subse- ns of schemata for which we ll build a new schema whose the configurations and whose tions. In this way, we can find a and thereby gain a specifica-

age, it should be noted that it w schemata be grown (by the

mechanisms outlined above) but old schemata can evolve or be "tuned." Within schema theory as we have developed it here, there are three ways in which this can come about. First, we can get more precise information on the nature of the "distributions" underlying the variable constraints. Everytime we determine that a particular schema gives a good account of a situation we can use the value of its variables to modify the variable constraints and the correlations among the various variable values. Second, we can drop out apparently irrelevant aspects of a schema. If a certain variable is rarely filled from the input situation, it is probably not a very important aspect of the schema and perhaps could be dropped from the specification of the schema. A similar argument would apply for presumed, "fixed properties" of a schema. If such properties simply are rarely or never observed, they cannot be very important aspects of the schema. Finally, old schemata can be tuned by adding new variables or fixed properties that appear to be relevant. If a given schema always differs from the situations for which it is intended to account by a small constant difference, that constant element should be added to the specification of the schema.

A real case in point is that of a five-year-old boy we know who currently believes that a sauna "is a wooden room where a lot of men sit around." Presumably, in the not too distant future, this child, with or without being influenced by the Women's Movement, will relax the constraint on sauna users to include women, and, hopefully will introduce variables for purpose, and for temperature, with a default value like "hot"!

Our discussion of learning has been necessarily vague. Nevertheless, it does appear that there are a number of mechanisms which can operate naturally within schema theory allowing for the natural growth and evolution of a schema system which can carry out the tasks required of it.

PROCESSING PRINCIPLES

It used to be fashionable for a rather sharp distinction to be made between cognitive structures and cognitive processes. More recently, however, emphasis on this distinction has waned for two related reasons. In the first place, models have been developed in which knowledge has been represented procedurally (see Hewitt, 1975; Winograd, 1972)). Second, models have been developed which represent many cognitive processes as structures identical in their characteristics to those used to represent more static knowledge. Our own suggestions for general problem-solving and inference schemata are good examples of this latter development. Yet, it is still the case that some more global processing principles are required in order to account for the availability of the right concepts at the right time. A theory of knowledge representation ought not to ignore this issue. Throughout the course of this paper we have repeatedly used such terms as "activation" and "invocation" which have cropped up at key points in our

discussions of almost all aspects of the theory. It is therefore now necessary for us to take a somewhat closer look at them.

The entire memory system contains an enormous number of schemata and memories. At any one time only a few of them are required and no procedure of random search could possibly lead to their efficient discovery. The search for likely candidate schemata must, therefore, be somehow guided, and it must be sensitive to the context, for the "correct" choices often depend on the context in which the processing is occurring. The same input is differentially interpreted by an observer depending on the conditions under which he observes it, what he has just observed, and what he expects to observe. In addition, although expectations are obviously important, unexpected events can be interpreted without going through all possible interpretations first. Thus, what seems to be required is a process which allows for the convergence of information so that information derived directly from the input can be combined with expectations to lead (more or less directly) to plausible candidate schemata.

We believe that schemata have characteristics which readily enable these requirements to be satisfied. The convergence is achieved by the combination of *bottom up* and *top down* processing. Bottom up processing occurs when aspects of the input directly suggest or activate schemata which correspond to them and when these schemata themselves activate or suggest dominating schemata of which they are constituents. In our example of Mary and the ice-cream man, Sentence (4a)–(4c), the occurrence of a word like "hear" (or a cognate) in the input would directly activate the HEAR schema. The HEAR schema may itself activate a dominating schema like one for "becoming aware" and this being an event would suggest the problem-solving schema. In general, we want to say that schemata activated by their own constituents are activated from the bottom up, so that bottom up processing is the activation of dominating schemata. Top down processing, on the other hand, arises from schemata activating their constituent subschemata. In our ice-cream example, the activation of the SELL schema by the ICE-CREAM MAN schema is a case in point, as is the activation of the TRY schema by the general problem-solving schema. These processes are called "top down" because they lead from conceptual expectations towards the data in the input where the satisfaction of these expectations might be found. In fact, such processes need not go all the way back to the input since they can meet with the bottom up processing. Bobrow and Norman (1975) call this latter type of processing *conceptually driven*, since it is ultimately concepts which generate a search for particular constituents which they suggest. They contrast this with *data driven* processing in which it is ultimately the input data which generates suggestions for particular concepts.

It may be helpful to think of these processing issues in terms of a computer-programming metaphor, for one can think of a schema as being a kind of *procedure*. Procedures have *subroutines* and one can think of the activation of a schema as being like the invocation of a procedure. The variables of a schema are

is therefore now necessary for
 ous number of schemata and
 required and no procedure of
 ient discovery. The search for
 how guided, and it must be
 often depend on the context
 ut is differentially interpreted
 which he observes it, what he
 rserve. In addition, although
 d events can be interpreted
 first. Thus, what seems to be
 gence of information so that
 e combined with expectations
 e schemata.

which readily enable these
 hieved by the combination of
 rocessing occurs when aspects
 hich correspond to them and
 ggest dominating schemata of
 Mary and the ice-cream man,
 "hear" (or a cognate) in the
 The HEAR schema may itself
 ing aware" and this being an
 n general, we want to say that
 ctivated from the bottom up,
 f dominating schemata. Top
 m schemata activating their
 s, the activation of the SELL
 in point, as is the activation
 g schema. These processes are
 ual expectations towards the
 ections might be found. In
 to the input since they can
 orman (1975) call this latter
 s ultimately concepts which
 they suggest. They contrast
 mately the input data which

ues in terms of a computer-
 schema as being a kind of
 think of the activation of a
 The variables of a schema are

thus analogous to the variables of a procedure while the subschemata are analogous to the subroutines which may be invoked from within it. The activation of subschemata within a schema is like the calling-up or invocation of the subroutines within a procedure. This is the paradigm case of conceptually-driven processing. However, unlike ordinary procedure calls, in which the flow of control is only from procedure to subroutine, the flow of control in a schema system operates both ways. It is as though a given procedure not only could invoke its own subroutines (conceptually driven processing), but also could invoke those procedures in which it was itself a subroutine (data driven processing). Finally, one must imagine these procedures as operating simultaneously.

If the combination of data driven and conceptually driven processing exhausted the processing mechanism, we would have a serious problem on our hands. For, were this to be the case, there would appear to be no way of preventing every schema in memory from becoming activated as soon as one was activated. The solution to this problem lies in the notion of "accounting for the input" which was discussed in the section on comprehension. In the normal course of processing, some schemata will find more evidence for themselves than will others, and in general, these will be schemata which are suggested from a number of different sources. It is upon these that processing will focus. "Finding good evidence" happens in a number of ways. First, a schema needs to find good bindings for its variables. Thus, if the GIVE schema is invoked there should be candidates for the giver, the recipient, and the gift. Second, a schema should find some evidence for its subschemata, so that for the GIVE schema there should be some evidence to suggest that a recipient did indeed GET the gift and perhaps that it did not happen by chance. Third, it should be possible to find a dominating schema which to some extent offers a good fit. Those schemata which fail to find such evidence cease processing and are deactivated so that the dominating schemata in which they occur can use the failure-to-fit information toward the assessment for their own goodness of fit. The details of the evaluation mechanism are beyond the scope of this paper, but a detailed mathematical formulation of it can be found in Rumelhart (1977b).

We mentioned earlier that contextual and situational factors influence the way in which inputs are interpreted. Since schemata are structures which provide interpretations for inputs in all modalities, the simplest way to understand the mechanism whereby such factors affect comprehension is to regard the input as including those factors. For example, consider hearing an utterance of (9) in the rather different situations of a bar and a children's birthday party.

(9) I would like something to drink.

In a bar one is continually seeing and hearing things connected with bars and one thus anticipates that many bar-related schemata will be activated, unless, that is, one is totally oblivious to one's surroundings. When (9) is encountered in such a situation and it activates a DRINK schema, drinks which can be found within

the BAR schema are going to be suggested from more sources than they would at a children's party where they might well not be currently active at all. So, in the context of a bar one would expect to find beer and liquor suggested more strongly and one would expect to find more evidence for them than one might for lemonade and milk which would presumably be more prevalent at the party. Thus, the utterance of Sentence (9) will give rise to different expectations in different situations. More local context effects, such as the influence of what immediately went before, are handled in exactly the same way. To summarize, information (including both the "stimulus" and the context) enters the system and directly suggests certain plausible candidate schemata to account for it. At the same time as this data driven processing is going on, such postulated schemata activate their dominating schemata, which in turn look for other as yet unsuspected aspects of the situation. This conceptually driven processing allows internal contextual constraints to be effective.⁵ A schema is said to provide a good account of (aspects of) the input situation when it can find good evidence for itself.

CONCLUSIONS

We claim that the schema theory provides both the concepts and the vocabulary for theorizing about the organization of knowledge. Indeed, the prevalence of schema-related notions in this volume attests to this fact. At the same time, it cannot be denied that the terms we utilize need to be constrained so as to prevent them from being absorbed into all manner of incompatible accounts, and to this end, we have tried to characterize the concepts to which they correspond in greater detail. A schema theory cannot be expected to completely describe the makeup and machinery of the mind, but as part of it we think it is promising.

The notion of schemata as varying in their abstractness relates rather directly to the findings of Meyer as reported not only in this volume but also in Meyer (1975) and Meyer and McConkie (1973). Meyer's research indicates that an independent objective characterization of the logical structure of a prose passage enables quite specific predictions to be made about the relative memorability of different ideas occurring within it. In particular, higher level ideas, ones which are more dominant in the logical structure, are better remembered than particular details, and, they are better remembered if their order of appearance in the passage is congruent with their priority in the structure. As Meyer herself points

⁵It should be noted that variable binding comes about as a result of either or both types of processing. A schema may be actively looking for an (aspect of the) input to bind to one of its variables (conceptually driven) and/or such an input may demand some variable to which it can itself be bound (data driven).

out, these findings tend to support Ausubel's (1963) claims about the importance of having a higher-order structure to which to attach the details. This can be translated very readily into schema theory. In fact, Rumelhart (1975) describes in some detail the structure one might anticipate certain kinds of STORY schemata to have, in particular, children's stories with "morals." Since a STORY schema can be regarded as a partially ordered set of rewrite rules, and since these rules (the STORY grammar) embody the (presumed) logical structure of a class of stories, it would follow with minimal assumptions that less processing would be required to fit a story to a schema when the story corresponded more closely to the schema structure than when it did not and, if understanding is considered to be finding such a fit, one can conclude that it is easier to understand a story whose structure closely matches that of the STORY schema. Arguing from better understanding to better "memory for gist" is not difficult. It can be done on both empirical and theoretical grounds. A story is just a special case of a prose passage and there is no reason to believe that Meyer's findings could not be replicated in the domain of children's stories.

Expressed in terms of schema we could thus come to two conclusions. First, if the binding of variables within a schema is normally assumed to proceed most smoothly from the top down, then providing information in a structured form most closely resembling the structure of the schema which will be required for its interpretation maximizes the likelihood that the interpretation will be appropriate and minimizes the processing required. In such a case, each successive piece of information, as it is assimilated, provides additional support that the interpretation will be appropriate and minimizes the processing required. In such a case, each successive piece of information, as it is assimilated, provides additional support that the interpretation so far achieved is indeed appropriate or "satisfactory." Second, and by parallel argument, one might expect to "unpack" some used schema in recall, and this unpacking will be most efficient if it is done from the top down. In this case, the major structural aspects would appear before the details, which are themselves less predictable as they become more specific.

The implications that schema theories, or indeed any other theories of knowledge organization, have for education must still be regarded as only potential. Awaiting more detailed models, we can nevertheless point to a few general considerations. It is certainly the case that one of the purposes of instruction is to provide the kind of knowledge that will prove useful to a person in processing new information and dealing with novel situations. This goal can be regarded as equivalent to that of producing knowledge structures in which new information can be processed and understood. The provision, therefore, of new knowledge structures which do not have this characteristic is as pointless as is the provision of new information for which no interpretive structure can be found. The purpose of a schema is that of a cognitive template against which new inputs can be matched and in terms of which they can be comprehended. Thus, the role of

more sources than they would be currently active at all. So, in order and liquor suggested more evidence for them than one might expect to be more prevalent at the party. due to different expectations in such as the influence of what (the same way. To summarize, the context) enters the system schemata to account for it. At this is going on, such postulated that in turn look for other as yet actually driven processing allows. A schema is said to provide a when it can find good evidence

the concepts and the vocabulary. Indeed, the prevalence of this fact. At the same time, it is said to be constrained so as to avoid the presence of incompatible accounts. the concepts to which they are not expected to completely fit but as part of it we think it is

effectiveness relates rather directly to this volume but also in Meyer's research indicates that an overall structure of a prose passage affects the relative memorability of higher level ideas, ones which are better remembered than particular order of appearance in the structure. As Meyer herself points

is a result of either or both types (aspect of the) input to bind to one that may demand some variable to

examples in instruction can be regarded as providing individual cases in which a schema can have its variables bound; well-chosen examples will fully exploit such a schema by showing the nature and bounds of values that its variables can take. The generation of new knowledge structures and demonstrations of the way in which they can be used can thus be regarded as one of the principle goals of instruction. Ortony (1975b) has argued that metaphor is a powerful instructional device. In the current context, one might regard metaphors as aids to selecting an old schema, which with relatively little modification, can be used to produce a new one. One might thus use a "flowing water" schema as the basis for the generation of an "electric current" schema. The former might incorporate knowledge concerning unidirectionality of flow, branching, capacity of the conduit, and so on, all of which would have their analogs in an "electric flow" schema. Good instruction would clarify the metaphor of electricity in wires as water in pipes by specifying which variables stay and which go. What makes it a metaphor, after all, is that some of the new information will not fit into the old schema.

The generation, modification and instantiation of schemata seems to us to characterize both informal learning and formal schooling. There are many ways in which they can occur, ranging from discovery through play, to insight through instruction. In all cases, existing knowledge is utilized in and required for the acquisition of new knowledge. We leave the implications of this for the new born child for the consideration of biologists and philosophers.

ACKNOWLEDGMENTS

The work reported herein was supported in part by Grant NS 07454 from the National Institutes of Health, by Grant NIE-G-74-0007 from the National Institute of Education, and a grant to the second author from the Sloan Foundation.

REFERENCES

- Anderson, J. R., & Bower, G. *Human associative memory*. Washington, D.C.: Winston, 1973.
- Anderson, R. C., & Ortony, A. On putting apples into bottles - A problem of polysemy. *Cognitive Psychology*, 1975, 7, 167-180.
- Ausubel, D. P. *The psychology of meaningful verbal learning: An introduction to school learning*. New York: Grune & Stratton, 1963.
- Barclay, J. R., Bransford, J. D., Franks, J. J., McCarrell, N. S., & Nitsch, K. Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 471-481.
- Bartlett, F. C. *Remembering*. Cambridge, England: Cambridge University Press, 1932.
- Bobrow, D. G., & Norman, D. A. Some principles of memory schemata. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press, 1975.

providing individual cases in which a few examples will fully exploit such values that its variables can take. and demonstrations of the way in which it is used as one of the principle goals of the metaphor is a powerful instruction. It might regard metaphors as aids to thought which, with little modification, can be used to describe "flowing water" schema as the basis for a "flow" schema. The former might incorporate flow, branching, capacity of the pipe, and their analogs in an "electric flow" metaphor of electricity in wires as to how it flows and which go. What makes it a new formation will not fit into the old

formation of schemata seems to us to be a result of schooling. There are many ways to gain insight through play, to insight through metaphors utilized in and required for the applications of this for the new born philosophers.

REFERENCES

Grant NS 07454 from the National Science Foundation, the National Institute of Education, and the National Science Foundation.

memory. Washington, D.C.: Winston, 1975.

to bottles - A problem of polysemy.

learning: An introduction to school

U. N. S., & Nitsch, K. Comprehension and Verbal Behavior, 1974, 13,

Cambridge University Press, 1932.

memory schemata. In D. G. Bobrow & R. S. Lockhart (Eds.), *Learning: Studies in cognitive science*. New

- Charniak, E. *Toward a model of children's story comprehension*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, 1972. (Also, MIT Artificial Intelligence Laboratory Tech. Rep. AI-TR, 266, 1972.)
- Charniak, E. Organization and inference in a frame-like system of common sense knowledge. In proceedings of *Theoretical issues in natural language processing: An interdisciplinary workshop*. Cambridge, Mass.: Bolt, Beranek, & Newman, Inc., 1975. Pp. 42-51.
- Clark, H. H. *Comprehension and the given-new contract*. Paper presented to a conference on "The role of grammar in interdisciplinary linguistic research," University of Bielefeld, Bielefeld, Germany, December 1973.
- Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, 240-247.
- Collins, A. M., Warnock, E. H., Aiello, N., & Miller, M. L. Reasoning from incomplete knowledge. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press, 1975.
- Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.
- Fillmore, C. The case for case. In E. Bach & R. I. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, Rinehart, & Winston, 1968.
- Gentner, D. Evidence for the psychological reality of semantic components: The verbs of possession. In D. A. Norman, D. E. Rumelhart, & the LNR Research Group, *Explorations in cognition*. San Francisco: Freeman, 1975.
- Halford, H. M., Ortony, A., & Anderson, R. C. A context-sensitive representation of word meanings. *Memory and Cognition*, in press.
- Hewitt, C. Stereotypes as an ACTOR approach towards solving the problem of procedural attachment in FRAME theories. In proceedings of *Theoretical issues in natural language processing: An interdisciplinary workshop*. Cambridge, Mass.: Bolt, Beranek, & Newman, Inc., 1975.
- Johnson, M. K., Bransford, J. D., & Solomon, S. K. Memory for tacit implications of sentences. *Journal of Experimental Psychology*, 1973, 98, 203-205.
- Kant, E. *Critique of pure reason* (1st ed. 1781, 2nd ed. 1787, translated by N. Kemp Smith). London: Macmillan, 1963.
- Kintsch, W. Notes on the structure of semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- Labov, W. The boundaries of words and their meanings. In C. J. Bailey & R. Shuy (Eds.), *New ways of analyzing variation in English*. Washington, D.C.: Georgetown University Press, 1973.
- Lakoff, G. Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Papers from the Eighth Regional Meeting, Chicago Linguistics Society*. Chicago: University of Chicago Linguistics Department, 1972.
- Meyer, B. J. F. *The organization of prose and its effects on memory*. Amsterdam: North-Holland Publ., 1975.
- Meyer, B. J. F., & McConkie, G. W. What is recalled after hearing a passage? *Journal of Educational Psychology*, 1973, 65, 109-117.
- Minsky, M. A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision*. New York: McGraw-Hill, 1975.
- Norman, D. A. *Resources and schemas replace stages of processing*. Paper presented at the Sixteenth Annual Meeting, The Psychonomic Society, Denver, Colorado, November 1975.
- Norman, D. A., Rumelhart, D. E., & the LNR Research Group. *Explorations in cognition*. San Francisco: Freeman, 1975.
- Ortony, A. How episodic is semantic memory? In proceedings of *Theoretical issues in*

- natural language processing: An interdisciplinary workshop*. Cambridge, Mass.: Bolt Beranek & Newman, Inc., 1975. (a)
- Ortony, A. Why metaphors are necessary and not just nice. *Educational Theory*, 1975, 25, 45-53. (b)
- Palmer, S. E. Visual perception and world knowledge: Notes on a model of sensory-cognitive interaction. In D. A. Norman, D. E. Rumelhart, & the LNR Research Group, *Explorations in cognition*. San Francisco: Freeman, 1975.
- Petrie, H. G. Action, perception and education. *Educational Theory*, 1974, 24, 33-45.
- Polanyi, M. *Personal knowledge*. Chicago: University of Chicago Press, 1958.
- Polanyi, M. *The tacit dimension*. New York: Doubleday, 1966.
- Powers, W. T. *Behavior: The control of perception*. Chicago: Aldine, 1973.
- Quillian, M. R. Semantic memory. In M. Minsky (Ed.), *Semantic information processing*. Cambridge, Mass.: M.I.T. Press, 1968.
- Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 1-20.
- Rosch, E. H. On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press, 1973.
- Rumelhart, D. E. Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press, 1975.
- Rumelhart, D. E. Understanding and summarizing brief stories. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977. (a)
- Rumelhart, D. E. Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance, VI*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977. (b)
- Rumelhart, D. E., & Levin, J. A. A language comprehension system. In D. A. Norman, D. E. Rumelhart, & the LNR Research Group, *Explorations in cognition*. San Francisco: Freeman, 1975.
- Rumelhart, D. E., Lindsay, P. H., & Norman, D. A. A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- Schank, R. C. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 1972, 3, 552-631.
- Schank, R. C., & the Yale A. I. Project. *SAM - A story understander* (Res. Rep. 43). New Haven, Connecticut: Yale University, Department of Computer Science, August 1975.
- Schank, R. C., & Abelson, R. P. Scripts, plans and knowledge. In *Advance Papers of the Fourth International Joint Conference on Artificial Intelligence*. Tbilisi, Georgia, USSR, 1975. Pp. 151-157.
- Schmidt, R. A. A schema theory of discrete motor skill learning. *Psychological Review*, 1975, 82, 225-260.
- Spiro, R. J. *Inferential reconstruction in memory for connected discourse* (Tech. Rep. No. 2). Urbana, Ill.: Laboratory for Cognitive Studies in Education, University of Illinois, 1975.
- Tulving, E. Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.
- Tulving, E., & Thomson, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, 80, 352-373.

ary workshop. Cambridge, Mass.: Bolt

l just nice. *Educational Theory*, 1975, 25,

nowledge: Notes on a model of sensory-
Rumelhart, & the LNR Research Group,
an, 1975.

Educational Theory, 1974, 24, 33-45.

ity of Chicago Press, 1958.

eday, 1966.

. Chicago: Aldine, 1973.

: (Ed.), *Semantic information processing*.

: distance and the verification of semantic
Behavior, 1973, 12, 1-20.

d and semantic categories. In T. E. Moore
of language. New York: Academic Press,

In D. G. Bobrow & A. M. Collins (Eds.),
cognitive science. New York: Academic

ng brief stories. In D. LaBerge & S. J.
tion and comprehension. Hillsdale, N.J.:

reading. In S. Dornic (Ed.), *Attention and*
m Associates, 1977. (b)

chension system. In D. A. Norman, D. E.
plorations in cognition. San Francisco:

A process model for long-term memory.
tion of memory. New York: Academic

of natural language understanding. *Cogni-*

story understander (Res. Rep. 43). New
of Computer Science, August 1975.

d knowledge. In *Advance Papers of the*
tial Intelligence. Tbilisi, Georgia, USSR,

or skill learning. *Psychological Review*,

or connected discourse (Tech. Rep. No.
ss in Education. University of Illinois,

lving & W. Donaldson (Eds.), *Organiza-*

ity and retrieval processes in episodic
3.

Walker, J. H. Real-world variability, reasonableness judgments, and memory representations for concepts. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 241-252.

Winograd, T. *Understanding natural language*. Edinburgh: Edinburgh University Press, 1972.

Winograd, T. Frame representations and the declarative-procedural controversy. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press, 1975.