Psychological Review
1975, Vol. 82, No. 6, 407–428

A Spreading-Activation Theory of Semantic Processing

Allan M. Collins
Bolt Beranek and Newman Inc., Cambridge, Massachusetts

Elizabeth F. Loftus
University of Washington

This paper presents a spreading-activation theory of human semantic processing, which can be applied to a wide range of recent experimental results. The theory is based on Quillian's theory of semantic memory search and semantic preparation, or priming. In conjunction with this, several of the misconceptions concerning Qullian's theory are discussed. A number of additional assumptions are proposed for his theory in order to apply it to recent experiments. The present paper shows how the extended theory can account for results of several production experiments by Loftus, Juola and Atkinson's multiple-category experiment, Conrad's sentence-verification experiments, and several categorization experiments on the effect of semantic relatedness and typicality by Holyoak and Glass, Rips, Shoben, and Smith, and Rosch. The paper also provides a critique of the Smith, Shoben, and Rips model for categorization judgments.

Some years ago, Quillian[1] (1962, 1967) proposed a spreading-activation theory of human semantic processing that he tried to implement in computer simulations of memory search (Quillian, 1966) and comprehension (Quillian, 1969). The theory viewed memory search as activation spreading from two or more concept nodes in a semantic network until an intersection was found. The effects of preparation (or priming) in semantic memory were also explained in terms of spreading activation from the node of the primed concept. Rather than a theory to explain data, it was a theory designed to show how to build human semantic structure and processing into a computer.

This research was supported by the National Institute of Education, U.S. Department of Health, Education, and Welfare, Project 1-0420, under Contract OEC-1-71-0100(508) with Bolt Beranek and Newman Inc. Revision of the paper was supported by a grant from the John Simon Guggenheim Memorial Foundation to the first author. We would like to thank Stephen Woods, Colin MacLeod, Mark L. Miller, and the reviewers for their comments on previous drafts of this paper.

Requests for reprints should be sent to Allan M. Collins at Bolt Beranek and Newman Inc., 50 Moulton Street, Cambridge, Massachusetts 02138.

[1] Quillian's theory of priming appeared in the unpublished version of the 1967 paper (i.e., CIP Paper No. 79, Carnegie Institute of Technology, 1965).

Since the theory was proposed, there have been a number of experiments investigating retrieval and priming in semantic memory. In the present paper, we attempt to show how an elaboration of Quillian's basic theory can account for many of the results. In the first section, we briefly review the original theory while trying to correct a number of the common misunderstandings concerning it. In the second section, we extend the theory in several respects, and in the third section show how the extended theory deals with some recent experimental findings. In the fourth section we compare the theory to the model of Smith, Shoben, and Rips (1974).

Quillian's Theory of Semantic Memory

The fact that Quillian's theory was developed as a program for a digital computer imposed certain constraints on the theory, which Quillian felt were psychologically unrealistic. We will recount the theory as he proposed it, and then elaborate the theory in psychological terms. The theory made a number of assumptions about structure and processing in human semantic memory. A brief discussion of these assumptions follows.

People's concepts contain indefinitely large amounts of information. Quillian used the example of a machine. If one asks people to tell everything they know about machines, they will start off giving obvious properties, for example, that machines are man-made and have moving parts. But soon people run out of obvious facts and begin giving facts that are less and less relevant, for example, that a typewriter is a machine or even that the keys on the IBM electric typewriters select the position of a ball that strikes the ribbon against the paper. The amount of information a person can generate about any concept in this way seems unlimited.

In these terms concepts correspond to particular senses of words or phrases. For example, not only is the noun "machine" a concept, but the verb "to machine" is a concept; the "particular old car I own" is a concept; the notion of "driving a car" is a concept; even the notion of "what to do if you see a red light" has to be a concept. Thus, people must have a very large number of concepts, and concepts must have very complicated structures.

A concept can be represented as a node in a network, with properties of the concept represented as labeled relational links from the node to other concept nodes. These links are pointers, and usually go in both directions between two concepts. Links can have different *criterialities,* which are numbers indicating how essential each link is to the meaning of the concept. The criterialities on any pair of links between two concepts can be different; for example, it might be highly criterial for the concept of a typewriter that it is a machine, and not very criterial for the concept of machine that one kind is a typewriter. From each of the nodes linked to a given node, there will be links to other concept nodes and from each of these in turn to still others. In Quillian's theory, the full meaning of any concept is the whole network as entered from the concept node.

The links are not simply undifferentiated links, but must be complicated enough to represent any relation between two concepts. In the original theory, Quillian proposed five different kinds of links: (a) superordinate ("isa") and subordinate links, (b) modifier links, (c) disjunctive sets of links, (d) conjunctive sets of links, and (e) a residual class of links, which allowed the specification of any relationship where the relationship (usually a verb relationship) itself was a concept. These different kinds of links could be nested or embedded to any degree of depth, so that the format was designed to be flexible enough to express anything, however vague or specific, that can be expressed in natural language.

The search in memory between concepts involves tracing out in parallel (simulated in the computer by a breadth-first search) along the links from the node of each concept specified by the input words. The words might be part of a sentence or stimuli in an experimental task. The spread of activation constantly expands, first to all the nodes linked to the first node, then to all the nodes linked to each of these nodes, and so on. At each node reached in this process, an activation tag is left that specifies the starting node and the immediate predecessor. When a tag from another starting node is encountered, an *intersection* between the two nodes has been found.[2] By following the tags back to both starting nodes, the *path* that led to the intersection can be reconstructed.

When an intersection has been found, it is necessary to *evaluate* the path to decide if it satisfies the constraints imposed by syntax and context. The complicated kinds of decision rules that are invoked for comprehension of sentences in this evaluation phase are described in Quillian (1969). For categorization tasks these rules are described by Collins and Quillian (1972b) and in the next section of this paper. As an example, in a phrase such as "the fall leaves," a path found between the concept "to fall" and the

[2] There can be intersections with more than two starting nodes, but we have limited our discussion in this paper to the case of two nodes (as did Quillian, 1966, initially). The basic assumptions in Quillian's theory and our elaboration can apply to intersections with more than two starting nodes, but this leads to complications in the evaluation of intersections. Some of these were discussed by Quillian (1969) in regard to comprehension, but for the experiments considered in this paper, only the case of two starting nodes needs to be considered.

concept "tree leaf" would be rejected as a wrong interpretation, because syntax requires a participial form of "fall" to fit that interpretation. If a path found is rejected, other paths are considered in the order in which they are found.

Priming (or preparation) involves the same tracing process that was described for memory search. When a concept is primed, activation tags are spread by tracing an expanding set of links in the network out to some unspecified depth. When another concept is subsequently presented, it has to make contact with one of the tags left earlier to find an intersection. One of the nonobvious implications of this view of priming is that links as well as nodes will be primed. This is because Quillian treated links themselves as concepts (see above). Thus priming a node such as "red" will prime the links involving the relation "color" throughout the network. This provides a very powerful context mechanism.

Common Misinterpretations of Quillian's Theory

There is a rich variety of misinterpretations of Quillian's theory, many of them deriving from Collins's (Collins & Quillian, 1969, 1970a, 1970b) simplifications of the theory. The problem arose because Collins and Quillian were investigating specific aspects of the theory and only described enough of the theory to motivate a particular experiment. In turn experimenters made interpretations of these simplified versions, which did not fit with the theory as described elsewhere (Bell & Quillian, 1971; Quillian, 1966, 1969).

Perhaps the most prevalent misinterpretation of Quillian's theory concerns the idea of cognitive economy (Anderson & Bower, 1973; Conrad, 1972). In this regard, it is important to distinguish the strong theory of cognitive economy, which Conrad takes issue with in her attack on Collins and Quillian (1969), and the weak theory of cognitive economy, which Collins and Quillian were testing (though they did not spell it out clearly enough). As Conrad (1972) states, she rejects the "hypothesis that all properties are stored only once in memory and must be retrieved through a series of inferences for all words except those that they most directly define" (p. 153). This is a statement of the strong theory of cognitive economy. Undoubtedly the Collins and Quillian (1969) paper gave rise to this notion, but the authors cautioned against making that interpretation of the theory. As they said, "people surely store certain properties at more than one level in the hierarchy" (p. 242), and they cited the maple leaf as an example of this general rule.

The strong theory requires erasing information whenever it applies at a more general level. If a person learns a robin can fly and then later that birds fly, the strong theory implies that "flying" must be erased from "robin." The weak theory of cognitive economy merely assumes that every time one learns that X is a bird, one does not at that time store all the properties of birds with X in memory. Thus, an inference will be necessary to decide that X can fly, unless one encounters this fact directly. Hence, Collins and Quillian, in testing the weak theory of cognitive economy, picked instances where people were not likely to have encountered the general property with the specific instance (e.g., "A wren can fly"). The point of the experiment was to test whether it was possible to measure inference time, when the weak theory of cognitive economy implies that an inference is likely to be necessary for most subjects.

Another assumption sometimes made about Quillian's theory is that all links are equal (Anderson & McGaw, 1973; Rips, Shoben, & Smith, 1973; Wilkins, Note 1). In Quillian's original theory, there were criteriality tags on links, as we described earlier. In Collins and Quillian (1969, 1972b) links were assumed to have differential accessibility (i.e., strength or travel time). The accessibility of a property depends on how often a person thinks about or uses a property of a concept. Whether criteriality and accessibility are treated as the same or different is a complex issue, but network models allow them to be treated either way. Thus for example, even though "lungs," "hands," and "warts" are all linked directly to the concept "human," these links need not be in any sense equal. The same is true for the links between "bird" and its exemplars, such as "robin," "chicken," or "penguin." Rips et al. (1973) suggest that intermediate nodes are necessary for a network model to explain the reaction time differences they find in categorizing different birds. This makes the mistaken assumption that all links are equally criterial or accessible in any network model. It turns out, however, that differences in links are crucial to many different aspects of human semantic processing as Carbonell and Collins (1973) point out in their discussion of importance (or criteriality) tags.

A related implication of the Rips et al. (1973) paper and also a more recent paper of Smith et al. (1974) is that feature models can account for data that network models cannot. A feature model posits that a concept consists of a set of values on a large number of semantic dimensions (e.g., animateness, color, etc.). What is strange about this argument is that network models were developed as a method of representing features in a computer. Any process that can be represented in a feature model is representable in a network model; in particular, the Smith et al. model itself could be implemented in a semantic network (Hollan, 1975). In fact, network models are probably *more* powerful than feature models, because it is not obvious how to handle inferential processing or embedding in feature models.

Smith et al. (1974) argued in favor of feature models because their data for comparison of concepts seemed to fit a feature comparison process. What should be emphasized about Quillian's theory is that the parallel search would inevitably lead to just such a feature comparison process, though the process would take place over a period of time as different connections are found. One way that Quillian's theory is different from the Smith et al. models is that superordinate connections, if they exist, would also be found and evaluated. The distinction between these two theories is so crucial that we will discuss it at length in conjunction with the spreading activation theory's explanation of the Rips et al. (1973) results.

Another common misconception of Quillian's theory shows up in Juola and Atkinson's (1971) work on categorization judgments. In a categorization task, response time is measured for a subject to decide whether or not a particular instance (e.g., "car") is a member of one or more categories (e.g., "flower" or "vehicle"). Juola and Atkinson assume that in Quillian's theory the memory search to make a categorization judgment proceeds from the instance to the category. In fact, the wording in Collins and Quillian (1969) mistakenly gives that impression. But Quillian's theory (1966, 1969) assumes the search proceeds from both the instance and category in parallel. However, if one or the other is presented first, this gives the search from that node a head start, which is the notion of priming. Juola and Atkinson's experiment involves priming in a complicated way, which we will discuss below.

Anderson and Bower (1973) reject a Quillian-like model of a parallel search, while acknowledging that their data are compatible with "a parallel model whose search rate is slower in proportion to the number of paths that must be searched" (p. 371). Anderson and Bower's argument implies wrongly that Quillian has made the independence assumption for his parallel search. An independent parallel search is like a race where the speed of each runner is independent of the other runners. This is a common assumption in psychology, because it makes it possible to assign an upper bound to reaction time (see Sternberg, 1966). But there is no difficulty for Quillian's theory if the parallel search rate depends on the number of paths searched. Hence, Anderson and Bower's data are perfectly compatible with Quillian's parallel search.

The above discussion, then, shows what Quillian's theory is *not*, or at least some of what it is not. Several other misconceptions are discussed in Collins and Quillian (1972b), in particular the notion that Quillian's theory of memory is rigidly hierarchical, which Anderson and Bower (1973, p. 379) still believe, and Schaeffer and Wallace's (1970) argument that Quillian's theory predicts it will always take less time to compare concepts that are close together in the semantic network. We will return to some of these same papers below, in order to describe how the extended version of Quillian's theory accounts for some of the results these experimenters have used to reject Quillian's theory.

THE EXTENDED THEORY

In order to deal with the specific experimental results that have appeared in recent years, several more processing and structural assumptions must be added to the basic Quillian theory. These do not bend the theory, but merely elaborate it in such a way that it can be applied to the kinds of experiments on semantic memory that have been performed recently. The elaboration may itself be wrong, so our mistakes should not be held against Quillian's theory.

Local Processing Assumptions

There are four local processing assumptions in the extended theory. These four assumptions transform the theory from computer terms to quasi-neurological terms, a la Pavlov. But all the assumptions of the original theory should be preserved despite the transformation, except that activation tags are to be considered as source-specific activation (i.e., activation that is traceable to its node of origin).

1. When a concept is processed (or stimulated), activation spreads out along the paths of the network in a decreasing gradient. The decrease is inversely proportional to the accessibility or strength of the links in the path. Thus, activation is like a signal from a source that is attenuated as it travels outward.

2. The longer a concept is continuously processed (either by reading, hearing, or rehearsing it), the longer activation is released from the node of the concept at a fixed rate. Only one concept can be actively processed at a time, which is a limitation imposed by the serial nature of the human central process (Collins & Quillian, 1972b). This means that activation can only start out at one node at a time. But it continues in parallel from other nodes that are encountered as it spreads out from the node of origin.

3. Activation decreases over time and/or intervening activity. This is a noncommittal assumption that activation goes away gradually by some mechanism. Assumptions 2 and 3 impose a limitation on the amount of activation that can be allocated in priming more than one concept, because the more concepts that are primed, the less each will be primed.

4. With the assumption that activation is a variable quantity, the notion of intersection requires a threshold for firing. The assumption is that activation from different sources summates and that when the summation at the point of intersection reaches threshold, the path in the network producing the intersection will be evaluated.

Global Assumptions About Memory Structure and Processing

There are three assumptions in the extended theory concerned with the global structure of memory and its processing. These are generalizations of Loftus's (Note 2) arguments that semantic memory is organized primarily into noun categories and that there is a "dictionary" (or lexical memory) separate from the conceptual network.

5. The conceptual (semantic) network is organized along the lines of semantic similarity. The more properties two concepts have in common, the more links there are between the two nodes via these properties and the more closely related are the concepts. This means that different vehicles or different colors will all be highly interlinked through their common properties. This also implies that red things (e.g., fire engines, cherries, sunsets, and roses) are *not* closely interlinked, despite the one property they have in common. In these terms semantic relatedness is based on an aggre-

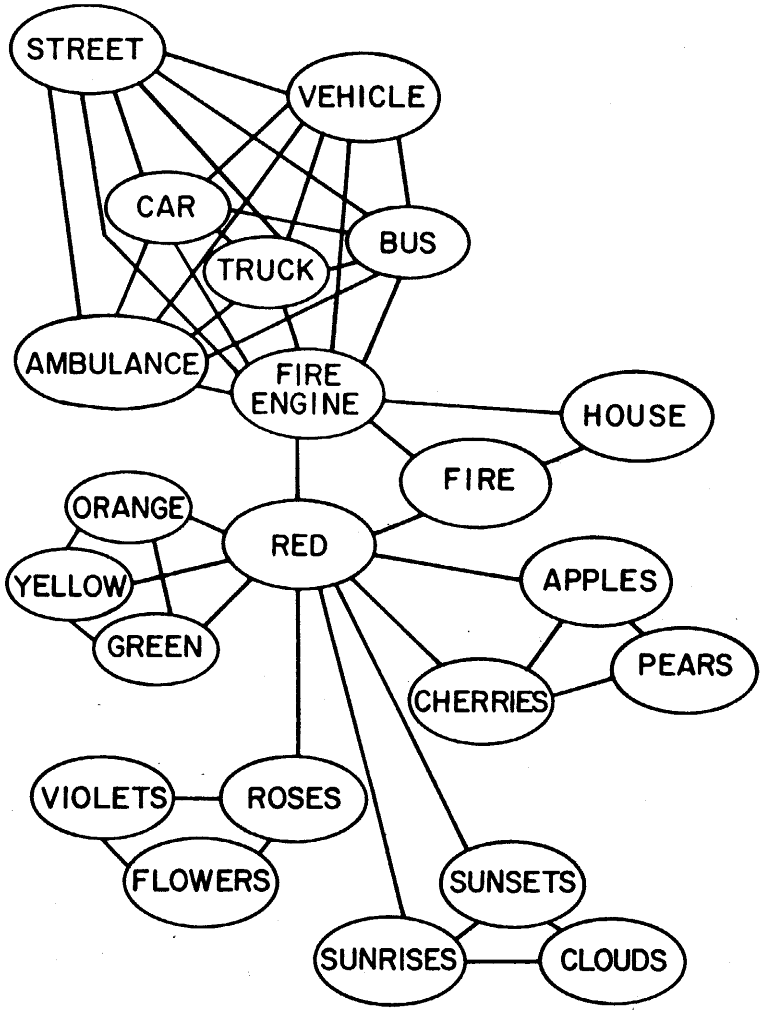

FIGURE 1. A schematic representation of concept relatedness in a stereotypical fragment of human memory (where a shorter line represents greater relatedness).

gate of the interconnections between two concepts.[3]

Figure 1 illustrates this aggregate notion of concept relatedness for a hypothetical human memory. (It is the kind of diagram that the scaling techniques of Rips et al., 1973, would produce.) In the figure the various vehicles are shown as closely related, because of the numerous individual connections that are assumed to exist between them. Conversely, the concepts associated with "red" are shown as less related, because of the presumed paucity of interconnections between them.

From the assumption that memory is organized according to semantic similarity, to-

[3] Semantic relatedness is a slightly different notion from semantic distance, though the two terms are sometimes used interchangeably. Semantic distance is the distance along the shortest path, and semantic relatedness (or similarity) is an aggregate of all the paths. Two concepts may be close in distance, say by a path through "red," and still not be closely related because that is the only path. Our use of *close* to refer to both relationships is admittedly confusing. In this paper we shall use close to refer to relatedness or similarity, though in some tasks (Quillian, 1966) it is only distance that matters.

gether with earlier assumptions, it follows that if "vehicle" is primed, activation at any type of vehicle will accumulate from many neighboring nodes. That is to say, to the degree that "fire engine" is primed by "vehicle," it will in turn prime "ambulance," "truck," "bus," etc., and each of these in turn will prime the others. On the other hand, if "red" is primed, the activation that spreads to "fire engine" will not prime "cherries," "roses," or "sunsets" to any great extent, because there are so few connections between these concepts. Instead, "fire engine" will tend to prime other vehicles, and "cherries" to prime other fruits. Hence, the same amount of activation will be diffused among a greater number of concepts.

6. The names of concepts are stored in a lexical network (or dictionary) that is organized along lines of phonemic (and to some degree orthographic) similarity. The links from each node in the lexical network are the phonemic properties of the name, specified with respect to their position in the word. The properties stored about names are assumed to be the properties that Brown and McNeill (1966) found people could identify about words on the "tip of their tongue." Each name node in the lexical network is connected to one or more concept nodes in the semantic network.

7. Loftus's (Note 2) data lead to the further assumption that a person can control whether he primes the lexical network, the semantic network, or both. For example, a person can control whether to prime (a) words in the lexical network that sound like "bird," (b) concepts in the semantic network related to "bird," or (c) words in the lexical network corresponding to the concepts in (b). This control over priming can be thought of in terms of summation of diffuse activation for an entire network (perhaps in a particular part of the brain) and source-specific activation released from a particular node. Thus, (a) would derive from activation of the lexical network together with the word "bird," (b) would derive from activation of the semantic network together with the concept "bird," and (c) would derive from activation of both networks together with the concept "bird."

Assumptions About Semantic Matching Process

There are a number of assumptions about the decision process for evaluating whether or not two concepts match semantically. This is a fundamental process that occurs in many aspects of language processing, such as matching referents, assigning cases, and answering questions (Collins & Quillian, 1972b; Collins, Warnock, Aiello, & Miller, 1975). Categorization tasks, which ask "Is X a Y?" (where X and Y are concepts), directly investigate this process. The decision process described here is a more explicit and somewhat revised version of the process postulated by Collins and Quillian (1972b), with additions to encompass the results of Holyoak and Glass (1975).

8. In order to decide whether or not a concept matches another concept, enough evidence must be collected to exceed either a positive or a negative criterion. The evidence consists of various kinds of intersections that are found during the memory search. Evidence from different paths in memory sum together. Positive and negative evidence act to cancel each other out, as shown by dialogue excerpts in Carbonell and Collins (1973). Failure to reach either criterion before running out of relevant evidence leads to a "don't know" response (Collins et al., 1975). This process is essentially the Bayesian decision model that is common in the reaction time literature (see, for example, Fitts, 1966; Stone, 1960).

There are a number of different kinds of paths between the two concepts that constitute positive or negative evidence. Any of these types of evidence might contribute to a particular decision. The different types are listed in Table 1 and described below in Assumptions 9–13.

9. If the memory search finds that there is a superordinate (or a negative superordinate) connection from X to Y, that fact alone can push the decision over the positive (or negative) criterion. Superordinate links act like highly criterial property links

TABLE 1

TYPES OF PATHS FOUND IN MEMORY THAT CONSTITUTE POSITIVE OR NEGATIVE EVIDENCE

Positive evidence	Negative evidence
Superordinate connection	Negative superordinate connection
Property comparison, matching property	Property comparison, distinguishing property
Wittgenstein strategy, matching property	Wittgenstein strategy, distinguishing property
	Mutually exclusive subordinates
	Counterexamples

(see below). For example, it is conclusive positive evidence that a mallard is a bird, if superordinate links are found between "mallard" and "duck" and between "duck" and "bird." Similarly, if a negative superordinate link is found between "bat" and "bird," it is conclusive evidence that a bat is not a bird.

10. If the memory search finds properties on which X and Y match (i.e., common properties), this is positive evidence proportional to the criteriality of the property for Y. If the memory search finds properties on which X and Y mismatch (i.e., distinguishing properties), this is negative evidence proportional to the criteriality of the property for Y. There is an asymmetry in the weighing of positive and negative evidence in a property comparison, because a mismatch on just one fairly criterial property can lead to a negative decision, whereas most of the highly criterial properties must match in order to reach a positive decision (Collins & Quillian, 1972b).

It is important to note that property comparisons and superordinate connections sum together in reaching either criterion as the memory search finds them. Thus, distinguishing properties make it harder to reach the positive criterion when there is a superordinate connection and therefore slow down the process.

As an example of property comparison, suppose there is no superordinate connection in a particular person's memory between "mink" and "farm animal" and between "cat" and "farm animal." Then the decision as to whether minks or cats are farm animals might be based on a comparison of the properties of minks or cats on one hand, and farm animals on the other. The most criterial properties of farm animals are presumably being animate and being kept on farms, but other less criterial properties include being domesticated, being raised for some purpose, or being kept in barns or outside. How a particular person would weigh the various properties of minks and cats to decide whether they are farm animals would vary from person to person (our intuition is that minks are farm animals and cats are not, even though both have the two most criterial properties of farm animals—what Smith et al., 1974, called defining properties). This decision strategy is similar to that proposed by Smith et al., as we will discuss later.

11. The Wittgenstein strategy is a variant of the property comparison strategy. It is postulated on the basis of Wittgenstein's (1953) observation that to decide whether something is a game (for example, frisbee), a person compares it to similar instances that are known to be games. Our assumption is that if any properties of X are found that match properties of another instance whose superordinate is Y, these constitute positive evidence. Similarly, any distinguishing properties constitute negative evidence. In the Wittgenstein strategy, unlike the property comparison strategy, matching properties count just as much toward a positive decision as distinguishing properties count toward a negative decision.

To illustrate the Wittgenstein strategy, Collins and Quillian (1972b) pointed out that in deciding whether a stagecoach is a vehicle, it might be compared to a car. The many properties that a stagecoach has in common with a car constitute strong positive evidence that a stagecoach is a vehicle. But notice that a stagecoach does not have a motor, which is highly criterial for being a car. Though this is strong evidence that a stagecoach is not a car, it is only weak evidence that it is not a vehicle. This illustrates how the same evidence is weighed differently in the property comparison strategy and the Wittgenstein strategy. The final decision that a stagecoach is a vehicle might depend both on matching properties between a stagecoach and vehicles in general (conveyance, motion, etc.) and matching properties between a stagecoach and particular vehicles like a car (seats, doors, etc.). Thus the property comparison strategy and the Wittgenstein strategy might combine to determine a person's response.

12. The mutually exclusive subordinates strategy was necessary for programming a computer to answer questions (Collins et al., 1975). Holyoak and Glass (1975) argue that this strategy accounts for some of their reaction time data (they call it a contradiction). The assumption is that if two concepts have a common superordinate with mutually exclusive links into the common superordinate, then this constitutes strong negative evidence, almost comparable to a negative superordinate link.

For example, if the question is whether a mallard is an eagle, the fact that a mallard is a duck and ducks and eagles are mutually exclusive kinds of birds is rather conclusive evidence that a mallard is not an eagle. Though Holyoak and Glass (1975) do not mention it, the mutually exclusive restriction is necessary. For example, the fact that Mike Mansfield is a politician does not exclude him from being a lawyer. Although "politician" and "lawyer" are both occupational roles, they are not mutually exclusive and in fact most politicians are lawyers. But lacking specific information to the contrary, people may make a default assumption of mutual exclusivity when two concepts have a common superordinate.

13. Counterexamples also can be used as negative evidence. This strategy derives from Holyoak and Glass (1975), who argue that statements of the kind "All birds are canaries" are disconfirmed by finding a counterexample, such as "robin." If the question is of the form "Is X a Y?" and there is a superordinate link from Y to X, then finding a counterexample involves finding a Z that also has X as superordinate and is mutually exclusive from Y. This is conclusive evidence that X is not always a Y.

Holyoak and Glass (1975) discusss counterexamples in the context of the universal quantifier "all," but the same process would occur for a question of the kind "Is a marsupial a kangaroo?" In such a case, retrieving a counterexample (such as a wallaby) can be used to determine that kangaroos are a subset of marsupials and not equivalent (e.g., "automobiles" and "cars" are equivalent concepts).

Though these five kinds of evidence (Assumptions 9–13) are the only ones we have postulated for the semantic matching process seen in categorization tasks, there may be other kinds of evidence of this sort. We should stress that there are many other kinds of evidence people use for answering more complicated questions (Collins et al., 1975). It is beyond the scope of this paper, however, to consider all the different ways people use evidence to make semantic decisions.

Recent Experiments

In this section, we discuss how the theory deals with some different kinds of recent experiments. The four types of studies to which we apply the theory are (a) several production experiments by Loftus (Freedman & Loftus, 1971; Loftus, 1973a, 1973b, Note 2); (b) Juola and Atkinson's (1971) multiple-category experiment; (c) the Conrad (1972) sentence-verification experiment; and (d) several categorization experiments on the effects of semantic relatedness and typicality (Holyoak & Glass, 1975; Rips et al., 1973; Rosch, 1973; Smith et al., 1974). We intend to deal with the major kinds of available findings to which the Quillian theory has not yet been applied. Our objective is to show how a spreading-activation theory can handle these results, not to consider all the possible alternative explanations of the experiments.

Production Experiments of Loftus

There are several Loftus experiments we want to discuss in terms of the spreading-activation theory. The first of these is an experiment by Freedman and Loftus (1971), in which subjects had to produce an instance of a category that began with a given letter or was characterized by a given adjective. For example, subjects might be asked to name a fruit that begins with the letter *A* or a fruit that is red. On some trials the category was shown first and on some trials second. Hence, this was a priming experiment in that one concept was activated before the other. Reaction time was measured from the onset of the second stimulus.

Our concern is with the finding that subjects were faster when the category (e.g., "fruit") was given first than when either the letter or the adjective was given first. This basic result was later replicated even for cases in which the instance named was a more frequent associate to the adjective than to the category noun (e.g., "lemon" is a closer associate of "sour" than of "fruit").

The explanation in terms of the theory is as follows: When a noun, such as "fruit," is presented first, the activation spreads to nodes connected to "fruit," among which are instances such as "apple," "pear," "peach," "orange," and "lemon." But these concepts are all highly interlinked with each other (though some, such as "orange" and "lemon," are more closely interlinked than others). Thus, the total amount of activation is spread among a relatively small number of closely interlinked concepts (see Assumption 5). However, when an adjective or letter is presented first, say "red" or "A," the activation spreads to a much wider set of concepts, which are not particularly interlinked with each other. Thus, the large variety of different things that are red or that start with the letter *A* will receive relatively little priming when the adjective or letter are given first. Because priming the noun leads to a greater accumulation of activation on the instances, these are closer to their threshold for firing, so that it takes less stimulation, and hence less time, to trigger an intersection when the second stimulus is presented.

Freedman and Loftus (1971) explained their finding in terms of entering the category when a noun is presented and entering a cluster within the category when the adjective or letter is presented. Thus if the noun is presented first, the subject can enter the category immediately and need only choose the correct cluster when the adjective or letter is presented. But if the adjective or letter is presented first, the subject must wait until the category is presented, because the cluster is specific to that category. (However, Loftus, Note 2, has revised this explanation for the letter stimulus in her dictionary-network model.)

The Freedman and Loftus explanation is not altogether different from the explanation offered here, though our theory is less rigidly hierarchical. The rigid hierarchy gets into trouble with errors such as one we encountered where a subject produced "Ben Franklin," given the stimulus pair "president" and "F," although he later recognized his mistake. In an activation theory, "Franklin" is a very likely intersection starting at "president" and "F," because he is so closely linked with the concept, "president," and some of its foremost instances, such as "Washington." Such a wrong intersection was likely in this case because the correct answer (prior to Ford) was "Fillmore," who is rather inaccessible and unlikely to be found quickly enough to preclude finding "Franklin." Once such an intersection is found, it is only by evaluating the connection between "president" and "Franklin," that it can be rejected (see Assumptions 8–13). It is a general problem of category-search models that they cannot deal with such errors.

Perhaps the major advantage of the spreading-activation theory over the Freedman and Loftus (1971) explanation is in tying their result to a parallel result in a quite different experiment by Loftus (1973b). In a categorization experiment, Loftus found that the direction of the association between the category and the instance determined whether subjects were faster when given the category first or the instance first. In the experiment she used four kinds of category–instance pairs: (a) pairs where both the category and instance evoked the other with high frequency (e.g., "tree–oak"); (b) pairs where the category evoked the instance with high frequency, but the instance evoked the category with low frequency (e.g., "seafood–shrimp"); (c) pairs where the category evoked the instance with low frequency, but the instance evoked the category with high frequency (e.g., "insect–butterfly"); and (d) pairs where both the category and instance evoked the other with low frequency (e.g., "cloth–orlon"). When the category was presented before the instance, reaction time for Conditions (a) and (b) was approximately equal and significantly faster than for Conditions (c) and (d). However, when the instance was presented first, reaction time for Conditions (a) and (c) was approximately equal and significantly faster than for Conditions (b) and (d). That is to say, subjects are fast when the category is presented first, if the category evokes the instance with high frequency, and subjects are fast when the instance is presented first, if the instance evokes the category with high frequency. The spreading-activation theory explains the pattern of reaction times in the following way, assuming that production frequency is a measure of the strength or accessibility of the path from one concept to another. When the first concept (i.e., the one presented first) evokes the second with a relatively high frequency, this means that more activation spreads to the second, and it takes less time to reach the threshold for an intersection. Thus, the amount the first concept primes the second concept determines the reaction time.

By comparing this experiment with the Freedman and Loftus (1971) study, it can be seen that the two results are exactly parallel. Based on our structural assumptions, "fruit" primes "apple" more than "red" or the letter "A" primes "apple" in the Freedman and Loftus study. Hence, the shorter reaction time occurs when "fruit" or "A" is presented first. Similarly in the Loftus (1973b) study, when the category primes the instance most highly, the shortest reaction times occur when the category is presented first. But when the instance primes the category most highly, the shortest reaction times occur when the instance is presented first. A spreading-activation explanation is quite compelling to account for the Loftus (1973b) results, and the theory offered here encompasses the order effect in both the Loftus study and the earlier Freedman and Loftus experiment within a single framework.

Recently Loftus (Note 2) has found two different ways in which presenting a letter acts differently from presenting an adjective, in variations of the Freedman and Loftus paradigm. This has led her to the development of a dictionary-network model, which we will translate into spreading-activation terms. The first difference between presenting a letter and an adjective appeared when Grober and Loftus (1974) compared reaction time in two conditions: one where noun–adjective (e.g., "fruit–red") and noun–letter (e.g., "fruit–A") trials were randomly intermixed, and one where noun–adjective and noun–letter trials were separated into blocks. In all cases the noun preceded the adjective or letter. The results of this experiment are shown in Figure 2. It is clear that when the subject knows a letter is coming, he can prepare for it. But in the mixed condition, the subject apparently prepares for either kind of trial the same way he prepares for an adjective trial, since adjective trials take the same amount of time in either case. The theory's description of semantic processing on the adjective trials is the same as that given earlier for the Freedman and Loftus (1971) experiment, with the amendment that only the semantic network and not the lexical network would be diffusely primed before the adjective is presented. When an intersection is found in the semantic network, then the subject must retrieve the name from the lexical network.

Loftus (Note 2) described what must happen on noun–letter trials in the blocked condition as follows:

> The first step of the process is entering the category. The next step is a quasi-parallel simultaneous search towards the Dictionary. That is to say, the subject traces some number of pathways leading from category instances to the Dictionary representations of those instances. This step can be started during the interval between the

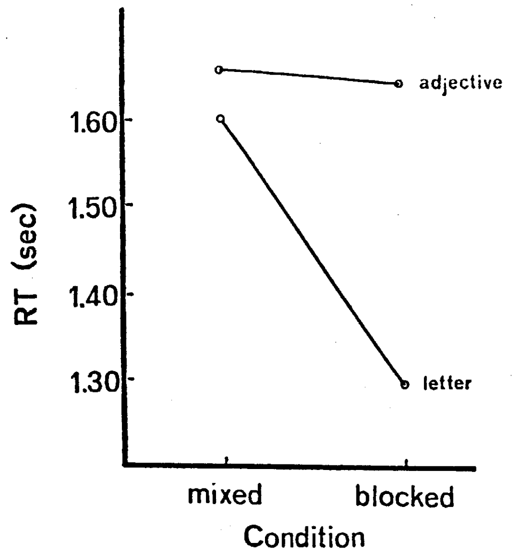

FIGURE 2. Reaction time for noun–adjective and noun–letter stimuli in mixed and blocked conditions (from Loftus, Note 2).

presentation of the category name and the restricting letter *if* the subject knows a letter is coming. (p. 13)

This is essentially the spreading-activation explanation, if the dictionary is taken to be a lexical network. Rather than saying that "the subject traces some number of pathways," which suggests a conscious tracing process, the present theory would say that activation spreads along some number of pathways, because the subject has activated the lexical network in addition to the semantic network (see Assumption 7). Hence, in the present explanation, the subject's control is reduced to diffusely activating whole networks rather than specific pathways (in addition to the specific nodes activated by the stimuli in the experiment). The difference in the results for the noun–letter trials in the two conditions then depends on whether the subject primes both networks (as in the blocked condition) or only the semantic network (as in the mixed condition). The reason he only activates the semantic network in the mixed condition may be either because of a principle of least effort (hence he could speed up his reaction time if he tried) or because there is less activation available to the semantic network if both are primed (hence he will be slower on noun–adjective trials if he primes both).

As can be seen in Figure 2, the subject is much slower on noun–adjective trials than on noun–letter trials in the blocked condition. This is accounted for by the fact that an intersection on a noun–adjective trial occurs in the semantic network and requires the further step of retrieving the corresponding name in the lexical network. On the other hand, the intersection on a noun–letter trial occurs at the name in the lexical network. Therefore, the name does not then need to be retrieved.

The second result that shows up the difference between adjectives and letters was predicted by Loftus from the dictionary-network model. In this experiment (Loftus & Cole, 1974) subjects saw three stimuli, ordered either noun, adjective, letter or noun, letter, adjective. For example, the three stimuli might be "animal," "small," and "M," for which an appropriate response is "mouse." The prediction was that the subject should be faster when the adjective is presented before the letter, and this was the result found. The reasoning is as follows: When the adjective appears before the letter, activation will spread from a small set of instances in the semantic network to the lexical network where the intersection occurs, since the letter can be expected just as in the blocked condition. When the letter is presented before the adjective, activation will spread from a small set of instances in the lexical network back to the semantic network where an intersection with the adjective will occur. Then the subject must return again to the lexical network to retrieve the name, so there is an extra transit necessary in this condition.

Loftus has also run a series of experiments in which subjects were asked to produce a member of a category and a short time later asked to produce a different member of that category (Loftus, 1973a; Loftus, Senders, & Turkletaub, 1974; Loftus & Loftus, 1974). This was accomplished by showing a category–letter pair (e.g., "fruit–P"), which asked the subject for an appropriate instance, then, following 0, 1, or 2 intervening items, showing the same category paired with a different letter (e.g., "fruit–A"), which asked for a different instance. The general finding is that reaction time for the second instance is shorter than reaction time for the first instance and increases monotonically with the number of intervening items. For example, in Loftus (1973a) a subject's baseline time to name a fruit beginning with the letter "P" was 1.52 sec. However, it took him 1.22 sec to produce the same response if he had named a different fruit on the previous trial and 1.29 sec to produce the response if he had named a different fruit two trials back.

The spreading-activation theory predicts these results by assuming that when an item is processed, other items are activated to the extent that they are closely related to that item. That is, retrieving one category member produces a spread of activation to other category members, facilitating their later retrieval. The assumption (Assumption 3) that activation decreases over time or trials predicts the lag effect.

Meyer and Schvaneveldt (Meyer, 1973; Meyer & Schvaneveldt, 1971; Schvaneveldt & Meyer, 1973; Meyer, Schvaneveldt, & Ruddy, Note 3) have also shown that the time to retrieve information from memory is faster if related information has been accessed a short time previously. Their paradigm is somewhat different. Subjects were required to classify letter strings as words or nonwords. The general finding was that the response time to classify a letter string as a word is faster if the subject has just classified a semantically similar word as opposed to a semantically dissimilar word. Thus, for example, the time it takes to classify "butter" as a word is faster if "butter" is preceded by "bread" than if it is preceded by "nurse." Their results have led Meyer and Schvaneveldt to an explanation in terms of spreading activation and illustrate the widely different paradigms that such a theory can encompass.

Juola and Atkinson's Study with Multiple Categories

An increase in reaction time with multiple categories has been found by Juola and Atkinson (1971) in a task where subjects had to decide whether a stimulus word belonged to one of a variable number (1–4) of prespecified (target) categories. They compared this task with one where subjects decided if the stimulus word was the same as one of a variable number (1–4) of target words. Their experiment was designed to distinguish between two kinds of models, one they attribute to Landauer and Freedman (1968) and one they attribute to Collins and Quillian (1970a). In most respects, their results fit the model they derived from Landauer and Freedman, but since the spreading-activation theory provides an alternative explanation for their results, we want to compare their two models with our theory.

The model Juola and Atkinson (1971) derived from Landauer and Freedman (1968) is very similar to what Landauer and Meyer (1972) call the "category-search model." It assumes that the subject searches through instances of the categories in memory seeking a match for the stimulus word. Such a model predicts that as the number of categories or words in the memory set increases, reaction time for the category-matching task should increase at a greater slope than reaction time for the word-matching task. This is because each additional target category adds more instances that must be searched, whereas each additional target word only adds one, the word itself. This result was essentially what Juola and Atkinson found.

The model they ascribed to Collins and Quillian (1970a) assumed that subjects perform the category-matching task by retrieving their stored category for the stimulus word and comparing this to the given categories to see if it matches one of them. This model would predict that the slope for the two tasks should be about the same, and the intercept for the category-matching task should be greater than for the word-matching task. Their results clearly reject this model. Although attributed to Collins and Quillian, this model is quite different from Quillian's theory, because the semantic search in Quillian's (1966) theory is assumed to spread in parallel from both categories and instances. When the categories are given first, as in Juola and Atkinson's experiment, then activation would spread out from the categories before the instance even appeared.

In order to explain our interpretation of Juola and Atkinson's results, it is necessary to describe their procedure in more detail. They chose 10 large categories and 12 common instances from each category as stimuli. This makes a total of 120 instances in all. In the word-matching task, they presented from 1 to 4 of the 120 instances as targets on each trial. In the category-matching task, they presented from 1 to 4 of the 10 categories as targets on each trial. In both cases the negative stimuli were chosen from the same set of 120 instances. In the word-matching task, then, the discrimination necessary to categorize the stimulus was between one of the target words or a word that had not occurred as a target for a large number of trials (on the average about 24 trials earlier). In the category-matching task, however, the discrimination was between a word in one of the target categories and a word in one of the categories from a recent trial (on the average about 2 trials earlier). The discrimination, therefore, was rather easy in the word-matching task and quite difficult in the category-matching task.

What we think must be happening in the task is that the discrimination between positive and negative responses is made (at least partly) on the basis of activation level. In the category-matching task, as the number of categories increases, the amount of activation allocated to each category decreases (see Assumptions 2 and 3). Furthermore, activation will be left over from previous trials on the categories corresponding to negative instances, though it will have partly decayed (Assumption 3). For example, suppose "tree" was a target category on a particular trial and "body part" was not, but "body part" was a target on a previous trial. Then "tree" will have a higher activation level than "body part," but the difference will not be very large because "body part" was presented so recently. If a positive instance such as "oak" is presented, it will intersect with "tree"; if a negative instance such as "arm" is presented, it will intersect with "body part." The less activation on "tree," which depends on how many other targets there are, the harder it is to discriminate that it is in fact a target or that "body part" is not a target. The more difficult the discrimination is, the longer it takes to make, and thus there will be a fairly large effect of the size of the target set on reaction time.

In the word-matching task, however, the difference in activation level as the number of targets is varied will not be so critical a factor. This is because the absolute difference between the activation level of targets and nontargets is so much greater in the word-matching task, given Juola and Atkinson's experimental procedure. That is to say, each nontarget was presented as a target so many trials previously (approximately 24) that the activation level for a nontarget would have decayed (Assumption 3) to a very low activation level as compared to any target. Hence, the large absolute difference in activation levels between targets and nontargets makes the differences due to target-set size relatively unimportant.

In conclusion, the spreading-activation theory's explanation of Juola and Atkinson's (1971) results is that the effect of differences in activation level due to target-set size matter more when the discrimination is difficult and matter less when the discrimination is easier. Furthermore, as Juola and Atkinson point out, there are two aspects of their data (namely, the fact that the data for positive responses are not linear, and the marked recency effects in the serial position curves) that fit much better with a parallel model, such as spreading-activation theory, than they do with a serial model, such as the one they derive from Landauer and Freedman (1968).

There are two implications of this view that could be tested fairly easily. One is that reaction time to decide that an instance such as "arm" is a negative instance will depend on the recency with which "body part" was presented as a target category. The more recent its presentation, the longer it will take to say "no." A more global implication is simply that the slope of the curve with respect to target size depends on those factors that affect the difficulty of discrimination, such as the recency with which negative instances occurred as targets (and probably as nontargets as well). This has important implications for the memory-search literature as a whole.

Conrad's Study

Using a true–false reaction time technique for sentences (e.g., the task is to decide whether "A salmon can eat" is true or false), Conrad (1972) found results which she interpreted as contradictory to Quillian's (1966, 1969) theory of semantic processing. In fact, the results of her study are quite close to what Quillian's theory would predict given Conrad's methodology.

In her first experiment, which was like the Collins and Quillian (1969) study, Conrad selected 2-level and 3-level hierarchies from the common culture (e.g., salmon → fish → animal) and properties associated with the objects at different levels. Then she constructed sentences with instances, such as "salmon," from the lowest level and properties from all three levels. The results Collins and Quillian found were that reaction time increased as the property was farther removed from the instance in the hierarchy. The reason for the increases in reaction time according to spreading-activation theory is that as the instance and property are farther apart in the hierarchy, it takes activation longer to spread between them and to trigger an intersection (and perhaps to evaluate the path found as well).

Unlike Collins and Quillian (1969), Conrad (1972) broke down the properties in her sentences into three groups on the basis of the frequency (high, medium, low) with which people generated each property, given the different objects in the hierarchies. Another difference from Collins and Quillian's study is that she collected data over 5 days by repeating all the sentences each day.

The results of her first experiment were generally in the same direction as the increases in reaction time that Collins and Quillian found when the property was farther removed from the instance in the hierarchy. There was one reversal in her data out of nine comparisons, and this occurred for the high-frequency properties, where it was not unexpected given the weak theory of cognitive economy (as we will argue below). However, the increases she found were much smaller on the average than those of Collins and Quillian. The weak theory of cognitive economy predicts that people store a property with whatever instance it is linked to in a sentence, so Conrad's repetition of sentences over 5 days should lead to the smaller reaction time increases she found. This is because an inference necessary on the first day would be less likely on the second day, and so on. Conrad, in fact, reports a large Level × Day interaction.

In general, Conrad found that the higher the frequency of the property, the smaller the increases betwen levels. Given the weak theory of cognitive economy, we would expect that high-frequency properties are more likely to be stored at several levels in the hierarchy, because they are more likely to be encountered in contexts involving specific instances. For example, "leaves" are more likely to be stored as a property with particular types of trees (such as "maple" and "oak") than is "bark," because leaves are a higher frequency property. Thus, the effect of property frequency found by Conrad is consistent with the weak theory.

Conrad (1972) argued that Collins and Quillian's (1969) results could be explained by a confounding of property level and property frequency. Her argument was that the sentences Collins and Quillian used may have been based on high-frequency properties for Level 1 sentences (e.g., "A salmon is pink"), moderate-frequency properties for Level 2 sentences (e.g., "A salmon has fins"), and low-frequency properties for

Level 3 sentences (e.g., "A salmon can eat"). To support her argument, she showed that if one plots her reaction time data in the above way, one obtains approximately the same slope as Collins and Quillian did, whereas if one plots the slope for low-, medium-, or high-frequency properties separately, one obtains much smaller slopes.

However, there are two weaknesses in Conrad's argument. First, she did not use her frequency data to evaluate systematically the frequency of the properties in Collins and Quillian's sentences, so her conjecture about such a confounding had no empirical basis. Collins and Quillian (1969) did obtain subject ratings of importance of the property for the relevant level concept, which should correlate with Conrad's frequency measure. These ratings averaged 1.90 for Level 1 sentences, 1.92 for Level 2 sentences, and 2.16 for Level 3 sentences (based on a 5-point scale, where 1 = very important and 5 = not important). These are small differences and certainly do not support the notion that the slope between Level 1 and Level 2 sentences was due to the confounding Conrad hypothesized. The difference between Level 3 sentences and the others may have contributed to the greater slope that Collins and Quillian found, but even that is doubtful. For those subjects who had sentences with all 3 levels, the slope was actually larger (approximately 100 msec rather than 75 msec), but this was offset by a group of subjects who were slower overall and saw only Level 1 and 2 sentences. So the latter group acted to cancel out any exaggeration of the slope due to the lower importance of Level 3 properties.

Second, the comparison Conrad (1972) made in plotting her data against Collins and Quillian's data compared data based on five responses to the same sentence with data based on one response to a sentence. As indicated above, the weak theory of cognitive economy predicts that repetition of a sentence makes an inference less likely and should reduce the slopes in the way Conrad found. A fairer comparison would be between her data on the first day and Collins and Quillian's data. But even that comparison has the problem that she may well have included sentences of the kind, "A maple has leaves," where the property is a general property of trees, but where most people would store it as a property of maples as well. This suggests that the fairest comparison is between Collins and Quillian's data and her data on the first day for low-frequency properties, where the properties were least likely to be stored at more than one level. But we cannot make this comparison because she did not break down her data by days. In conclusion, the differences between the two experiments and the fact that the only relevant data do not particularly support the conjecture about a confounding of property level and property frequency make Conrad's argument rather tenuous.

It was Conrad's second experiment that appears more damaging to Quillian's theory, but here she made a crucial methodological change. She presented the object 1 sec before the property, and this turned the experiment into a priming study. In the study she presented properties true of the highest level nodes, together with objects (e.g., "salmon," "fish," or "animal") at different levels in the hierarchy. Therefore, she predicted from Quillian's theory that the lower level objects, such as "salmon," would take longer to confirm, since it would take activation longer to spread between lower level objects and higher level properties. But by presenting the object 1 sec before the property and by using only high-level properties, she made it possible for her subjects to prepare during the interval by priming the object's superordinates. For example, if a subject saw "salmon," his best strategy was to retrieve the superordinates, "fish" and "animal," because the property to appear would be a high-level property, such as "eating." In these circumstances, there is little reason to expect systematic differences between objects such as "salmon," "fish," and "animal." Thus, this particular experiment had real methodological problems as a test of Quillian's theory, and it is weaker evidence *against* spreading-activation theory than her first experiment is evidence for the theory.

Effects of Typicality and Semantic Relatedness in Categorization Tasks

In recent experiments, Rips et al. (1973), Rosch (1973), and Smith et al. (1974) have shown that reaction time in a categorization task corresponds very closely to ratings of how typical the instance is of the category. For example, robins and sparrows are considered typical birds whereas chickens and geese are not. The effect of typicality on reaction time is quite large even when frequency of the particular instances in the language is controlled. Like Smith et al., we would argue that the typicality effect is one more manifestation of the fact that semantic similarity speeds up positive decisions and slows down negative decisions. Such an effect has been found repeatedly (Collins & Quillian, 1969, 1970a, 1972b; Schaeffer & Wallace, 1969, 1970; Wilkins, 1971). While Landauer and Meyer (1972) argued that the evidence for similarity effects at that time was either questionable or artifactual, the evidence now seems so overwhelming that any viable theory must account for them. They are very damaging to the category-search model.

There are two reasons why spreading-activation theory predicts that atypical instances will take longer to categorize than typical instances. The most important reason derives from the way evidence is aggregated (see Assumptions 8–13). Because different connections that are found are combined as evidence, distinguishing properties can slow down a positive decision based on a superordinate connection or on matching properties. For example, the decision that a chicken is a bird (i.e., an atypical instance) might be made on the basis of a superordinate connection from "chicken" to "bird," which people learn because chickens are frequently referred to as birds. But the fact that people eat chickens, that they are raised on farms, and that they are rather large are all properties that distinguish chickens from most birds. If these distinguishing properties are found during the memory search, as some are likely to be, they act to slow down the positive decision, because they are negative evidence. Similarly, matching properties can slow down a negative decision. For example, the decision that a goose is not a duck might be made on the basis of the difference in their necks (a distinguishing property) or simply because they are stored as mutually exclusive kinds of birds (see Assumption 12), but the matching properties that are found (e.g., their affinity to ponds, their webbed feet, their large size) will slow down the decision that they are different. The argument here is similar to that of Smith et al. (1974), which we will discuss in comparing the two theories.

The second reason for the typicality effects relates to those cases where a superordinate connection is found. As we indicated earlier, superordinate links differ in accessibility (or strength), and accessibility depends on use. If a person frequently uses the link that a robin is a bird, and less frequently uses the link that a chicken is a bird (assuming approximately equal frequency for chickens and robins), then the accessibility of "bird" from "robin" will be greater than from "chicken." Because of this, accessibility will be highly correlated with typicality ratings. All the factors acting to make a chicken or a goose an atypical bird in the real world will also act to make the use of the superordinate link in a person's mind from "chicken" or "goose" to "bird" infrequent. It is because they are atypical that the superordinate link is weak, and this will also act to slow down reaction time in making categorization judgments about atypical instances.

The way evidence is aggregated in the theory also explains the common finding (Collins & Quillian, 1970a, 1972b; Holyoak & Glass, 1975; Rips et al., 1974) that people are fast to decide that semantically unrelated concepts are different (e.g., that a book is not a dog). In comparing such concepts, there are not likely to be any superordinate connections, and almost all property connections will involve distinguishing rather than matching properties. Therefore, almost all the connections found will constitute negative evidence, and subjects will be quite fast to reach the negative criterion in such cases. This too is similar to the explanation in the Smith et al (1974) model.

Recently, Holyoak and Glass (1975) have isolated two different cases where semantic relatedness or typicality does not produce the usual effect on reaction time for negative judgments. One case arises when the decision depends on what they call a contradiction and what we have called mutually exclusive subordinates. The other case arises when the decision depends on a counterexample.

In the first case, Holyoak and Glass found that people are faster to reject sentences such as "All fruits are vegetables" or "Some chairs are tables" than sentences such as "All fruits are flowers" or "Some chairs are beds." In these four sentences the two nouns are mutually exclusive subordinates. The difference between the sentences is that "vegetables" and "tables" are generated with high frequency, while "flowers" and "beds" are generated with low frequency, when subjects are given the frame "All fruits are . . ." or "Some chairs are . . ." and asked to produce a false sentence. This difference is in the opposite direction of the usual finding that negative judgments are slower when the two concepts are more closely related semantically. The explanation for this reversal according to the theory (and to Holyoak and Glass) is that people make these decisions not on the basis of distinguishing properties (though some might be considered), but because they are stored as mutually exclusive subordinates (Assumption 12). Generation frequency in this case is a measure of the strength of the connection between the two concepts and therefore of how long it will take to find the contradiction between the two mutually exclusive concepts.

The second finding of Holyoak and Glass (1975) involves sentences where people reject the sentence by finding a counterexample (Assumption 13). For example, "All animals are birds," can be rejected by finding another kind of animal, such as a mammal. In this case Holyoak and Glass varied the production frequency of the predicate noun (e.g., "birds") independently of the production frequency of the counterexample (e.g., "mammals"). Their finding was that reaction time depended not on the production frequency of the predicate noun (which is a measure of the semantic relatedness of the concepts in the sentence) but on the frequency of producing a counterexample. Here again where a decision strategy that is not based on distinguishing properties is appropriate, the reaction time data do not depend on the semantic relatedness of the two concepts.

The importance of these two findings by Holyoak and Glass (1975), in our view, is that they demonstrate that different kinds of evidence can be involved in making categorization judgments. This suggests that approaches such as that of Meyer (1970) and Smith et al. (1974), which try to formulate a single strategy for making such judgments, will inevitably fail.

RELATION OF THE THEORY TO THE MODEL OF SMITH, SHOBEN, AND RIPS

Quillian's (1966, 1969) theory was a forerunner of a number of global theories of semantic processing based on network representations, in particular those of Anderson and Bower (1973), Norman and Rumelhart (1975), and Schank (1972). These theorists have made important advances on the Quillian theory (especially in the representation of acts and causes) which in no way contradict the basic thrust of Quillian's theory. There are some differences between these theories and Quillian's, but the basic intent of this paper is to deal with those aspects of semantic processing where the model of Smith et al. (1974) is the major competitor to Quillian's theory.

Unlike the various network models, the model of Smith et al. represents concepts as bundles of semantic features. Their model has the virtues of being quite clear and explicit, and it agrees quite well with the reaction time data for categorization judgments, except for the Holyoak and Glass (1975) results. Because it is such an initially compelling model, we want to emphasize how it differs from spreading-activation theory and point out what we think are its inherent difficulties.

In the model of Smith et al. (1974), the meaning of a concept is assumed to be represented by semantic features of two kinds: defining features and characteristic features. Defining features are those that an instance must have to be a member of the concept, and the model assumes that features can be more or less defining. Characteristic features are those that are commonly associated with the concept, but are not necessary for concept membership. For example, "wings" might be a defining feature of "birds" and "flying" a characteristic feature, since all birds have wings but not all fly. In a categorization task, the model assumes that the two concepts are first compared in Stage 1 with respect to all their features, both characteristic and defining. If the match is above a positive criterion, the subject answers "yes"; if it is below a negative criterion, the subject answers "no"; and if it is in-between, the subject makes a second comparison in Stage 2 based on just the defining features. If the instance has all the defining features of the category, the subject says "yes" and otherwise says "no." If the subject can decide in Stage 1, his reaction time will be faster than if he decides in Stage 2.

There are several minor differences between the model of Smith et al. (1974) and the spreading-activation theory that could be minimized by slightly changing their model. The difference in wording between comparing features in their model and finding links between properties in our theory is really a nondifference. But the distinction between defining and characteristic features has the inherent difficulty, pointed out through the ages, that there is no feature that is absolutely necessary for any category.[4] For example, if one removes the wings from a bird, it does not stop being a bird. Furthermore, we doubt if people can make consistent decisions as to whether a feature is defining or characteristic, either from time to time or from one person to another. Smith et al. recognized that features are more or less defining (or criterial), but they were forced into making the artificial distinction between defining and characteristic in order to have a two-stage model. Still, the model could be revised to work without the two stages and make essentially the same reaction time predictions.

The revision is as follows: If features are compared over time, as in Quillian's (1966) theory, then as the process goes on longer, more features will be compared (assuming features have different accessibilities). The comparison process can have a positive criterion and a negative criterion just as before, and features can be weighted by their criteriality. If the match at any point in time is above the positive criterion, the subject says "yes"; if the match falls below the negative criterion, the subject says "no"; and otherwise he goes on comparing features. Finally, if he is running out of relevant information, he says "I don't know." This is simply the Bayesian decision model described in Assumption 9 of the extended theory, where the evidence consists of matching and mismatching features as in the property comparison of Assumption 11.

Thus, we agree that a decision process similar to the one that Smith et al. (1974) postulate does occur for *some* categorization decisions. But there is a fundamental disagreement, because they argue that *all* categorizations judgments are made by comparing features of the instance and category, whereas we argue that people use whatever evidence they find, including superordinate links.

Because they exclude the use of superordinate links, the model of Smith et al. has several inherent difficulties. The most obvious is the assumption that even when people have superordinate information stored, they do not use it. While most people may not have learned some superordinate relations (e.g., that a beaver is a mammal, or a sled is a vehicle), there are many they have learned (e.g., that a wren is a bird, and a beaver is an animal). Why would they not

[4] There is for living things a biologists' taxonomy, which categorizes objects using properties that are not always those most apparent to the layman. Thus, there are arbitrary, technical definitions that are different from the layman's ill-defined concepts, but this is not true in most domains. There is no technical definition of a game, a vehicle, or a country that is generally accepted.

use such information if it is stored? How in fact can they avoid using it? It is an unlikely model which postulates that people use information that is less relevant to make a decision, instead of information that is more relevant.

Another obvious difficulty with the Smith et al. (1974) model is that people seldom know the defining properties of concepts. For example, consider whether a whale is a mammal, a sponge is an animal, a bat is a bird, or a wren is a sparrow. In the Smith et al. model, these difficult (and slow) decisions would be made in Stage 2 on the basis of defining properties. But people generally have no idea what the defining properties of a mammal, an animal, a bird, or a sparrow are. Even if they know that one of the most criterial properties for being a mammal is that it bears its young alive, it seems highly unlikely that they know whether whales (or beavers for that matter) bear their young alive. Neither of the authors has any idea what properties of a sponge make it an animal, but if asked in an experiment whether a sponge was an animal, we would answer "yes," and we would be comparatively slow about it. The reason we would answer "yes" is simply that we were told at one time that a sponge is an animal. We were also told that a bat is not a bird, and if we had not been told, we fear we might have responded "yes" if asked whether a bat is a bird in a categorization experiment. The decision that a wren is not a sparrow would be made because they are mutually exclusive kinds of birds (See assumption 12). They are both small songbirds, and it is hard to believe that many people know what the defining features of a sparrow are that a wren does not have. The fact that there are cases where people must use superordinate information to make correct categorization judgments makes it unlikely that they do not use such information in other cases where they could make the decision simply by matching features or properties. This is one of the strongest arguments for a hybrid theory.

We would like to close this section by raising the question of why one should adopt such a complicated theory when the Smith et al. (1974) model is simpler and predicts the reaction time data quite well. We have tried to stress the inherent difficulties that their model has in ignoring superordinate information and in relying on defining properties. Experimental tests can probably be devised that will show up those difficulties. We will suggest one such test, but first we might point out that the results of the Loftus (1973b) categorization experiment described earlier do not fit the Smith et al. model very well. If a person is merely comparing features between the instance and the category, then it should not matter whether the instance or category is presented first. It is the asymmetry in the superordinate connections that predicts the asymmetry Loftus found in reaction time, and it is hard to imagine how one could have an asymmetry of that kind in comparing features of two concepts.

One experiment that might show difficulties with the Smith et al. (1974) model is a categorization task. The categories and instances used are based on their multidimensional scaling of birds and animals on the one hand, and mammals and animals on the other. As both Collins and Quillian (Note 4) and Rips et al. (1973) report, subjects are faster at deciding that bird names are in the category "bird" than in the category "animal," whereas they are slower at deciding that mammal names are in the category "mammal" than in the category "animal." Collins and Quillian argue that this is the way people learn the superordinates: that pigeons are birds and lions are animals. Smith et al. argue that it is based on shared features, and they show by their scaling solution that most birds are closer to "bird" than to "animal," and most mammals are closer to "animal" than to "mammal." But there are several bird names that are closer to "animal" than to "bird" (in particular, "goose," "chicken," and "duck"; "pigeon" is equidistant), and there are several mammal names that are closer to "mammal" than "animal" (in particular, "deer," "bear," and "lion"; "horse" is equidistant). We would predict that even for those instances the above pattern would hold, whereas a pure feature-matching theory, such as the Smith et al. model, makes the opposite prediction. So this is a possible test of the two theories. There are undoubtedly many other tests.

Finally, we want to explain why we have been led to such a complicated theory. In trying to write computer programs that answer different types of questions, it becomes apparent that any decision procedure that gives correct answers must be flexible enough to deal with many different configurations of knowledge in memory. This is because people have incomplete knowledge about the world (see Collins et al., 1975), and they often do not have stored particular superordinate links or criterial properties. Any realistic data base for a computer system will have this same kind of incomplete knowledge. Therefore, perhaps our strongest criticism of the Smith et al. (1974) model is that it breaks down when people lack knowledge about defining features.

While at one level this is a complicated theory, at another level it is a simpler theory than the Smith et al. model. By viewing superordinate links as highly criterial properties, the theory becomes a simple Bayesian model. It is only in specifying the particular configurations of knowledge that constitute positive or negative evidence for the Bayesian process that the theory becomes complicated. The difference between the two theories is that the Smith et al. model allows only one kind of evidence (matching or mismatching features), whereas the theory presented here allows other kinds of evidence as well. Thus the theory encompasses a revised version of the Smith et al. model as a special case of a more general procedure.

CONCLUSION

We have extended Quillian's spreading-activation theory of semantic processing in order to deal with a number of experiments that have been performed on semantic memory in recent years. The result is a fairly complicated theory with enough generality to apply to results from many different experimental paradigms. The theory can also be considered as a prescription for building human semantic processing in a computer, though at that level many details are omitted about decision strategies for different judgments that arise in language processing (see Carbonell & Collins, 1973; Collins et al., 1975; Quillian, 1969). We would argue that the adequacy of a psychological theory should no longer be measured solely by its ability to predict experimental data. It is also important that a theory be sufficiently powerful to produce the behavior that it purports to explain.

REFERENCE NOTES

1. Wilkins, A. J. *Categories and the internal lexicon.* Paper presented at the meeting of the Experimental Psychology Society, Oxford, England, 1972.
2. Loftus, E. F. *How to catch a zebra in semantic memory.* Paper presented at the Minnesota Conference on Cognition, Knowledge, and Adaptation, Minneapolis, 1973.
3. Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. *Activation of lexical memory.* Paper presented at the meeting of the Psychonomic Society, St. Louis, 1972.
4. Collins, A. M., & Quillian, M. R. *Categories and subcategories in semantic memory.* Paper presented at the meeting of the Psychonomic Society, St. Louis, 1971.

REFERENCES

Anderson, J. R., & Bower, G. H. *Human associative memory.* Washington, D.C.: V. H. Winston, 1973.

Anderson, R. C., & McGaw, B. On the representation of meanings of general terms. *Journal of Experimental Psychology,* 1973, *101,* 301–306.

Bell, A., & Quillian, M. R. Capturing concepts in a semantic net. In E. L. Jacks (Ed.), *Associative information techniques.* New York: American Elsevier, 1971.

Brown, R. W., & McNeill, D. The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior,* 1966, *6,* 325–337.

Carbonell, J. R., & Collins, A. M. Natural semantics in artificial intelligence. *Proceedings of the Third International Joint Conference on Artificial Intelligence,* 1973, 344–351.

Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior,* 1969, *8,* 240–248.

Collins, A. M., & Quillian, M. R. Does category size affect categorization time? *Journal of Verbal Learning and Verbal Behavior,* 1970, *9,* 432–438. (a)

Collins, A. M., & Quillian, M. R. Facilitating retrieval from semantic memory: The effect of repeating part of an inference. *Acta Psychologica,* 1970, *33,* 304–314. (b)

Collins, A. M., & Quillian, M. R. Experiments on semantic memory and language comprehension.

In L. W. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972. (a)

Collins, A. M., & Quillian, M. R. How to make a language user. In E. Tulving & W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972. (b)

Collins, A., Warnock, E. H., Aiello, N., & Miller, M. L. Reasoning from incomplete knowledge. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press, 1975.

Conrad, C. Cognitive economy in semantic memory. *Journal of Experimental Psychology*, 1972, *92*, 149–154.

Fitts, P. M. Cognitive aspects of information processing: III. Set for speed versus accuracy. *Journal of Experimental Psychology*, 1966, *71*, 849–857.

Freedman, J. L., & Loftus, E. F. Retrieval of words from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 107–115.

Grober, E., & Loftus, E. F. Semantic memory: Searching for attributes versus searching for names. *Memory & Cognition*, 1974, *2*, 413–416.

Hollan, J. D. Features and semantic memory: Set-theoretic or network model? *Psychological Review*, 1975, *82*, 154–155.

Holyoak, K. J., & Glass, A. L. The role of contradictions and counterexamples in the rejection of false sentences. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 215–239.

Juola, J. F., & Atkinson, R. C. Memory scanning for words versus categories. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 522–527.

Landauer, T. K., & Freedman, J. L. Information retrieval from long-term memory: Category size and recognition time. *Journal of Verbal Learning and Verbal Behavior*, 1968, *7*, 291–295.

Landauer, T. K., & Meyer, D. E. Category size and semantic memory retrieval. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 539–549.

Loftus, E. F. Activation of semantic memory. *American Journal of Psychology*, 1973, *86*, 331–337. (a)

Loftus, E. F. Category dominance, instance dominance, and categorization time. *Journal of Experimental Psychology*, 1973, *97*, 70–74. (b)

Loftus, E. F., & Cole, W. Retrieving attribute and name information from semantic memory. *Journal of Experimental Psychology*, 1974, *102*, 1116–1122.

Loftus, E. F., Senders, J. W., & Turkletaub, S. Retrieval of phonetically similar and dissimilar category members. *American Journal of Psychology*, 1974, *87*, 57–64.

Loftus, G. R., & Loftus, E. F. The influence of one memory retrieval on a subsequent memory retrieval. *Memory & Cognition*, 1974, *2*, 467–471.

Meyer, D. E. On the representation and retrieval of stored semantic information. *Cognitive Psychology*, 1970, *1*, 242–300.

Meyer, D. E. Correlated operations in searching stored semantic categories. *Journal of Experimental Psychology*, 1973, *99*, 124–133.

Meyer, D. E., & Schvaneveldt, R. W. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 1971, *90*, 227–234.

Norman, D. A., & Rumelhart, D. E. *Explorations in cognition*. San Francisco: W. H. Freeman, 1975.

Quillian, M. R. A revised design for an understanding machine. *Mechanical Translation*, 1962, *7*, 17–29.

Quillian, M. R. *Semantic memory*. Unpublished doctoral dissertation, Carnegie Institute of Technology, 1966. (Reprinted in part in M. Minsky [Ed.], *Semantic information processing*. Cambridge, Mass.: M.I.T. Press, 1968.)

Quillian, M. R. Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioral Science*, 1967, *12*, 410–430.

Quillian, M. R. The Teachable Language Comprehender: A simulation program and theory of language. *Communications of the ACM*, 1969, *12*, 459–476.

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. On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and acquisition of language*. New York: Academic Press, 1973.

Schaeffer, B., & Wallace, R. Semantic similarity and the comparison of word meanings. *Journal of Experimental Psychology*, 1969, *82*, 343–346.

Schaeffer, B., & Wallace, R. The comparison of word meanings. *Journal of Experimental Psychology*, 1970, *86*, 144–152.

Schank, R. C. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 1972, *3*, 552–631.

Schvaneveldt, R. W., & Meyer, D. E. Retrieval and comparison processes in semantic memory. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press, 1973.

Smith, E. E., Shoben, E. J., & Rips, L. J. Comparison processes in semantic memory. *Psychological Review*, 1974, *81*, 214–241.

Sternberg, S. High-speed scanning in human memory. *Science*, 1966, *153*, 652–654.

Stone, M. Models of choice-reaction time. *Psychometrika*, 1960, *25*, 251–260.

Wilkins, A. J. Conjoint frequency, category size, and categorization time. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 382–385.

Wittgenstein, L. *Philosophical investigations* (G. E. M. Anscombe, trans.). Oxford: Blackwell, 1953.

(Received November 4, 1974)