Retrieval of Information from Long-Term Memory

John R. Anderson

Summary. Information is represented in long-term memory as a network of associations among concepts. Information is retrieved by spreading activation from concepts in working memory through the network structure. The time required to retrieve information is a function of the level of activation that it achieves. Fanning of multiple paths from a node dissipates the activation the node sends down any path and increases retrieval time. Fan effects are reduced as subjects overlearn the material or when they can change their task from a recognition judgment to a consistency judgment.

Human memory is often considered to have two parts: a long-term memory, which is the repository of all our factual knowledge, and a short-term or working memory, which holds the information we are currently processing (*1*). Working memory is assumed to have such a limited capacity that at any time it can hold only a small subset of the information in long-term memory. When information is no longer needed, it is dropped from working memory or rapidly decays. As new information is needed, it must be retrieved from long-term memory, a process that takes a long time relative to the other human information processing steps. This article is concerned with the mechanisms that determine the temporal properties of the retrieval of information.

According to the general research procedure, subjects are presented with a statement and asked to indicate whether it is a fact they know. A subject typically sits in front of a computer terminal screen with his or her right index finger on a key that indicates *yes* and left finger on a key that indicates *no*. (This assignment facilitates response compatibility.) Sentences of the following sort appear on the screen:

1) *Ronald Reagan won the election.*
2) *Jimmy Connors plays tennis.*
3) *Jimmy Carter is the senator from New Jersey.*
4) *Joe Namath sees a psychiatrist.*
5) *Count Dracula would avoid the Arctic Circle in the summer.*

Most subjects would make relatively rapid decisions about the first three sentences, but would find it difficult to say anything about how they came to a decision. The apparent automaticity of these retrieval operations is one of their distinctive features. Sentence 4 is one about which the subject would conclude he does not know in contrast to 3, which he would judge false. Most of my research has been concerned with contrasting positive decisions like 1 and 2 with "don't know" decisions like 4. Although a subject may come to a positive judgment about sentence 5, he would probably come to a series of inferences and could probably report these inferences after the judgment. In contrast, sentences 1 and 2 require less response time (often under 1 second), and there is no experience that can be analyzed into substeps. In this article, I am concerned with the simpler kind of memory judgment. I believe that such simple memory retrievals are subcomponents of more complex judgment like that required for sentence 5.

The Fan Paradigm

Most of my research has not been with real-world facts such as those given above. Such facts have many uncontrolled properties that make theoretical inferences perilous. Rather, subjects typically learn a set of manufactured facts, often simple and homogeneous. After the subjects have learned this material they are tested to determine how the information is retrieved from long-term memory.

In the first experiments in this series (*2*), subjects committed to memory 26 facts of the form *A person is in the location*. Subjects studied between one and three facts about a person and between one and three facts about a location. Examples are:

A doctor is in the bank. (1-1)
A fireman is in the park. (1-2)
A lawyer is in the church. (2-1)
A lawyer is in the park. (2-2)

A lawyer is in the church is a (2-1) fact in that the list contains two facts about lawyers and one about the church. After studying each of the 26 sentences, subjects were asked to recall all of the locations in response to questions like "Where are the lawyers?" and all of the people in response to questions like "Who is in the park?" They continued this drill until they achieved perfect recall.

This knowledge can be represented in terms of propositional networks developed in cognitive psychology (Fig. 1) (*3, 4*). Propositional networks consist of nodes connected by links. There are two types of nodes—propositions and concepts. Each oval node signifies a proposition that encodes a fact. Links connect these proposition nodes with concept nodes. In terms of the network structure we are creating, various numbers of links emanate from the concepts. We refer to the number of links connected to a concept as the *fan* of the concept. *Park* and *lawyer* both occur with two facts—a fan of two is associated with each. A fan of one is associated with *doctor*. The experiment measured the effect of fan manipulation on speed of recognizing these facts.

Subjects presented with probes simply decided whether they had studied these probes or not. Studied probes are referred to as *targets* and those not studied are referred to as *foils*. The foils were created by presenting people and locations from the original material in new combinations. For instance, if the subjects had studied the material given above, *The doctor is in the park* would be a foil. We designed the study material to enable the orthogonal manipulation (for both targets and foils) of the fan of one to three associated with people and of the fan of one to three associated with locations.

For both targets and foils, reaction time increases systematically as the person or location fan increases (Table 1). Subjects also take about 100 millisec-

The author is a professor of psychology at Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

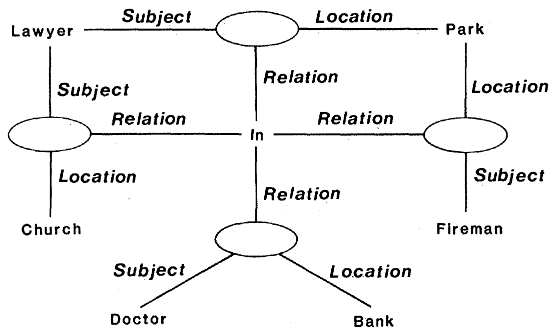


Fig. 1. A propositional network indicating the associative linkages among concepts that are created by studying a set of facts.

onds longer to reject foils than to accept targets. The error rate is low, and those errors tend to be a result of the speed stress and not true failures of memory. Subjects typically correct themselves immediately after pressing the wrong key. A number of experiments performed according to this procedure (*5–8*) show these same basic effects. They have used a large variety of materials and have found over a wide range of circumstances that judgment time for a probe—target or foil—increases as more facts are associated with any concept in the probe. This fan effect has guided the development of a theory of memory retrieval.

Spreading Activation

When a subject sees a probe, the memory representations of the concepts in the probe are activated. So when the subject sees *The lawyer is in the park*, the concepts in Fig. 1 corresponding to *lawyer, in*, and *park* will become active. Activation spreads down all the paths leading from the concepts. (Activation can flow in either direction along a link.) A subject is able to recognize a particular fact when activation from the various concepts has intersected and the level of activation has reached some threshold value; thus, the spread of activation determines the latency of the recognition judgment. Each concept is proposed to have a limited capacity for spreading activation. The more paths leading out of that concept, the less activation will be spread down any one path and the slower the recognition judgment. This slowing of recognition produces the fan effect.

This limited-capacity spreading-activation model (*5*) predicts that recognition time will vary with the fan of all the concepts in the probe. Activation spreads simultaneously from each of the concepts in the sentence. A contrasting model proposes that the subject gains access to memory from only one concept in the probe (for example, the first) and tries to retrieve information from that particular concept. In such a model there should be effects of the fan out of only the one concept rather than the approximately equal effects shown in Table 1 and in other experiments (*5–8*).

The Waiting Process

The spreading-activation model accounts only for the recognition of targets. We have also developed a waiting model for the rejection of foils (*5, 7*). After being presented with a probe, a subject waits for a set period for an intersection of activation. If none occurs, the subject concludes that the probe combination has not been studied and responds *no*. This waiting will lead to correct behavior if the subject waits longer to say *no* than it would take for an intersection to occur. The model predicts that it should take longer to reject foils than to accept targets, as is confirmed in Table 1. More time must be allowed for the waiting process when probes are constructed from higher fan concepts. Otherwise, the process would erroneously reject high fan targets, which take longer to reach threshold activation. Thus, the waiting process will take longer to reject foils of higher fan. The fan effect for foils is produced as this waiting process reproduces the temporal properties of spreading activation.

Table 1. Mean verification times for targets and proportions of errors for foils in the person-location experiment (*A doctor is in the bank*). [Adapted from tables 2, 3, and 4 in (*2*)]

Propositions per location	Reaction time (seconds)				Error rate (proportion)			
	Propositions per person			Mean	Propositions per person			Mean
	1	2	3		1	2	3	
				Targets				
1	1.111	1.174	1.222	1.169	0.051	0.042	0.046	0.046
2	1.167	1.198	1.222	1.196	0.065	0.056	0.060	0.060
3	1.153	1.233	1.357	1.248	0.063	0.044	0.054	0.054
Mean	1.144	1.202	1.267	1.204	0.059	0.048	0.054	0.053
				Foils				
1	1.197	1.221	1.264	1.227	0.019	0.042	0.030	0.030
2	1.250	1.356	1.291	1.299	0.014	0.037	0.044	0.032
3	1.262	1.471	1.465	1.399	0.042	0.079	0.051	0.057
Mean	1.236	1.349	1.340	1.308	0.025	0.053	0.042	0.040

King and Anderson (*7*) reported evidence for both the parallel spread of activation and the claim that foils are typically rejected by waiting for a fixed period without an intersection of activation. We had subjects commit to memory connected pairs of sentences like *The doctor touched the debutante* and *The doctor amused the lawyer* and unconnected pairs like *The fireman hated the mailman* and *The soldier bandaged the sailor*. (Connected pairs shared the same sentence subject, whereas unconnected pairs did not.) After learning this material, the subjects' task was to recognize whether a presented verb and noun came from the same sentence. In all cases the verbs and nouns used were the verbs and objects of studied sentences. In the case of foils, they were taken from different members of a connected or unconnected pair. The probes could be classified as targets and foils and as coming from connected pairs or not. For example, *touched debutante* is a connected target, *hated mailman* an unconnected target, *touched lawyer* a connected foil, and *hated sailor* an unconnected foil.

For targets, connected pairs did not differ from unconnected ones on either reaction time or error rate. For foils, however, subjects were much slower and made more errors with connected than with unconnected pairs. Figure 2 illustrates the relevant network structure for connected pairs. To recognize a verb-object connection, we assumed that activation spreads from the two concepts and the memory system monitors for an intersection. The time for activation to intersect between verb and object on a target probe should not be affected by the subject connection. With foils, however, the subject connection creates a spurious path for activation to intersect between verb and object. We assume

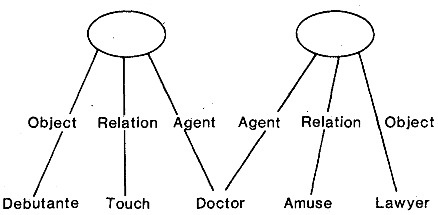


Fig. 2. The propositional network structure for connected pairs of sentences [as used in (*7*)].

that the subject will either respond positively to this spurious intersection (make an error) or first consider and then reject the intersection (increase latency beyond that required for unconnected foils).

Thus, we propose that a subject decides by default that he or she does not know something about a particular probe by waiting for a certain length of time during which intersection of activation occurs. Any intersection will have to be checked and will slow the decision process. Glucksberg and McCloskey (*9*) recently compared subjects' time to decide they did not know a fact like *John has the rifle* under two conditions. In the implicit condition, they had learned nothing relevant to this fact. In the explicit condition, they had committed to memory *It is not known whether John has the rifle*. Subjects more rapidly decided they did not know in the implicit condition. It is easier to simply wait a fixed period with no intersection than to have to evaluate the intersection of activation in the explicit condition.

Strength of Network Paths

The discussion to this point has assumed that each path from a concept to a proposition is equally activated. However, more frequently studied and tested facts about a concept can be recognized more rapidly; the relevant factor is not the total number of facts studied about a concept but rather the relative frequencies of these facts (*5*). I contrasted speed of recognition of facts about two-fan versus four-fan concepts, comparing facts that were tested with equal frequency in the two conditions. Suppose the subject studied the following facts in the two-fan condition:

The lawyer is in the church. (f = 3)
The lawyer is in the bank. (f = 3)

and the following facts in the four-fan condition:

The farmer is in the house. (f = 3)
The farmer is in the store. (f = 1)
The farmer is in the forest. (f = 1)
The farmer is in the boat. (f = 1)

where f is the frequency with which that fact was studied and tested in each block of the experiment. The first fact in each condition was studied half the time (three of six times) and thus has the same frequency but different fan.

In comparisons that equated for frequency and contrasted fan (for example, a two-fan fact studied three times and a four-fan fact studied three times), fan had little or no effect. In comparisons that equated fan and contrasted frequency (for example, a four-fan fact studied once and a four-fan fact studied three times), frequency had a large effect. Unpublished research from my laboratory has established that the relevant factor is relative, not absolute, frequency. For instance, a subject responds more slowly to a two-fan fact studied four times when the other fact has been studied eight times than when the other has been studied two times. This finding has led to a model that proposes that activation is divided among competing links according to their relative strength. In a normal fan experiment, relative strength and fan are confounded, because when each fact is studied equally often, the relative strength of a fact from a concept with N facts will be $1/N$.

Nonexperimental Facts

Lewis and Anderson (*10*) investigated whether these same basic retrieval phenomena could be obtained with actually true material that the subject knew from the real world. We had subjects learn sentences such as *George Washington wrote* Tom Sawyer, *Napoleon Bonaparte was from India*, and *Napoleon Bonaparte was a singer*. Subjects were told they were learning facts about a strange fantasy world and that they should regard these facts as one regards facts about famous figures in a historical novel, that is, as true in the context of the work of fiction. After having learned the material, subjects made judgments about whether particular facts were true of the person (recognition phase). They were to regard a fact about a person as true if it was true of the person in the real world or true of the person in the fantasy world, and false only if the fact was true of the person neither in the real world nor in the fantasy world. Examples of actually true facts were *Fidel Castro is a Cuban* and *George Washington crossed the Delaware*. The actual facts, the experimental facts, and the false assertions (not studied or false) could all be classified according to the number of facts studied about the famous person in the experiment. Recognition

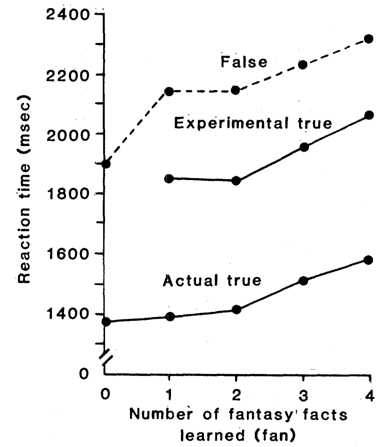


Fig. 3. Reaction times to nonexperimental facts, experimental facts, and foils as a function of fan. [Adapted from figure 2 in (*10*)]

times increased with fan for all three categories of responses (Fig. 3). In particular, learning new experimental facts about these people slowed subjects' retrieval times for old real-world facts. The actual trues were responded to much more rapidly than the experimental trues because the former should be much more strongly encoded given their high frequency of exposure.

Recognition of Faces

Anderson and Paulson (*11*) attempted to see if similar phenomena could be demonstrated with pictorial material such as faces constructed by arbitrarily combining the four features of hair, chin, mouth, and eye types (Fig. 4). Subjects were trained to identify each face as a person belonging to a particular profession. After reaching the criterion in this training task, subjects judged whether various faces had been studied (recognition phase). Foil faces were created by recombining the original features into novel combinations. Again we measured latency of recognition judgment, which we took to indicate the speed with which the relevant information could be retrieved from long-term memory.

A fan manipulation was produced by having the features contained in various faces. For instance, the two faces on the right in Fig. 4 share eyes, mouth, and chin. For any particular face, two of the four features were what we called null features, which appeared in six different study faces. The other two features appeared in either one or three faces. The faces—target or foil—could be classified by the number of faces associated with these last two features. The possible fan conditions are 1-1, 3-1, and 3-3. Thus, this experiment is analogous to the verbal experiments, with faces corresponding to sentences and physical features to words or concepts.

Figure 5 illustrates the hypothesized network consequences of our manipulation. Nodes for each face are attached to the features and to the professions. Features that appear in multiple faces have multiple arrows associated with them. The sailor in this figure is an example of a 3-1 face. Eyes a and chin g are the null features, hair d occurs in three faces, and mouth i in just that face.

We assumed that when a face was presented for recognition, its features would be identified and activated in memory. Activation would flow down the paths leading from these features. A face would be recognized when sufficient activation had accumulated at one of the face nodes. Thus, by having a feature appear in more faces we increased its network fan and reduced the amount of activation that could accumulate at any face. We predicted that face recognition time should increase as a feature was used in more faces. Judgment time for both targets and foils did increase with the fan of the critical features in the face, supporting the idea that the same retrieval processes that have been documented with verbal material also extend to pictorial material.

The Paradox of the Expert

These results are paradoxical in that they imply that the more one knows about a concept, the longer it will take to retrieve any particular fact about the concept. Therefore, an expert in a domain should be slowest to retrieve facts about that domain. Although some situations in everyday life seem to correspond to this implication (the more times one parks a car in a lot, the harder it is to remember its location), clear cases seem to contradict it.

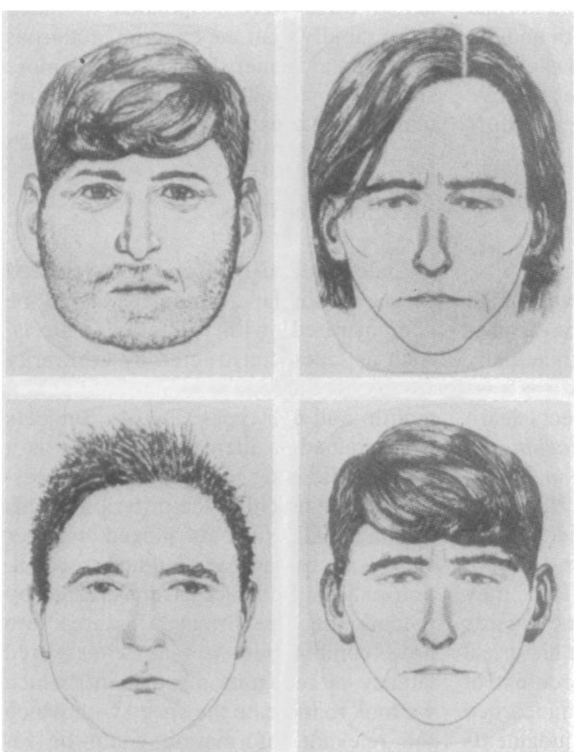


Fig. 4. Examples of the faces used in the study by Anderson and Paulson. [Adapted from figure 1 in (*11*)]

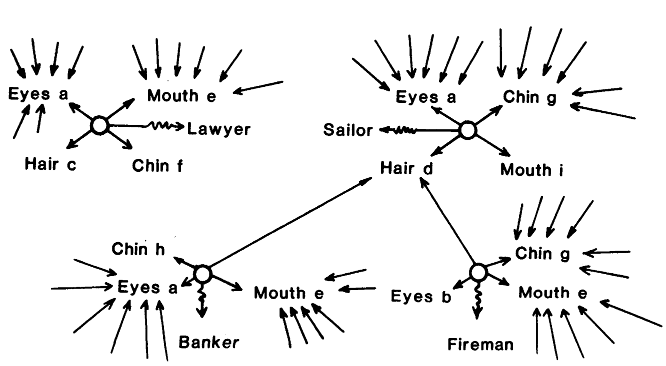


Fig. 5. A schematic representation of what a network encoding might be for the face material used by Anderson and Paulson [figure 2 in (*11*)]. Different letters after the features (eyes a and eyes b, for example) signify different specimens of those features.

I contrasted subject's speed of recognizing facts about familiar people (*Ted Kennedy is a senator*) with that of recognizing corresponding facts about less familiar people (*Birch Bayh is a senator*—true in 1975 and known to the subject population) (*5*). Subjects verified the fact about the more familiar concept more rapidly (in this case, 2.16 seconds for Kennedy versus 2.50 seconds for Bayh). These naturalistic materials lacked controls, but subsequent research (*12, 13*) has produced the same result with better controls. Thus, sometimes we are faster to judge facts about concepts with which we have greater familiarity. Much of the recent research in the fan procedure has been devoted to understanding how people overcome interference with familiar concepts.

Effects of Degree of Practice

The amount of practice that subjects have with experimental facts in a typical fan experiment is much less than the frequency with which the typical person has encountered explicitly or implicitly the fact that *Ted Kennedy is a senator*. Also, such a real-world fact has been studied over a more spaced interval. In a recent experiment we had subjects commit to memory subject-verb-object sentences such as *The doctor hates the lawyer* [see also (*6*)]. Subjects learned eight such sentences in the no-fan condition and eight in the fan condition. In the no-fan condition, the subject, verb, and object occurred in just that sentence, and in the fan condition, each term also occurred in some other sentence. Subjects learned this material to a criterion of perfect recall.

Subjects were then tested for their recognition of the sentences in a procedure that was repeated 5 days a week for 5 weeks. Each day they were presented with each sentence 24 times for recognition (mixed with foils). By the end of the experiment the subject had encountered each sentence 600 times. As with most practice data in a wide variety of domains (*14*), a power function describes recognition times for targets and foils as a function of days of practice (Fig. 6). The best-fitting line through the observed recognition times can be described as

$$T = 0.36 + 1.15(P - 0.5)^{-0.36}$$

for the no-fan condition and

$$T = 0.36 + 0.77(P - 0.5)^{-0.36}$$

for the fan condition, where T is recognition time in seconds and P is number of days of practice. That the two functions have the same exponents indicates that they are being equally benefited by the practice. Also, since the two functions are moving to the same asymptotic recognition time (0.36 second), with enough practice the fan effect would disappear. Hayes-Roth (*6*) has come to a similar conclusion.

The implication of this research is that with practice, nodes increase in their capacity for spreading activation. Given evidence about the impact of practice on increasing efficiency of neural transmission (*15*), this conclusion is not surprising. This means that familiar concepts should be able to support rapid retrieval of more facts.

Consistency Judgments

In the real world, one is rarely required to judge whether a particular fact was studied, but rather whether a fact is true or consistent with what one knows. When asked whether Darth Vader is good or evil one does not have to retrieve the specific proposition "Darth Vader is evil." The implications of fanning can be quite different when we look at judgments of consistency.

Smith *et al.* (*8*) had subjects in one condition study a pair of facts about Marty such as:

Marty produced the sour note.
Marty realized his seam was split.

In another condition subjects studied these two facts plus a third unrelated fact such as:

Marty painted the old barn.

Replicating past results they found subjects slower in the three-fact than in the two-fact condition. However, in a third condition, subjects would study the first two facts plus

Marty played the damaged bagpipes.

This third fact serves to integrate three sentences into a thematically related whole. Smith *et al.* found no difference between the two-fact condition and the integrated three-fact condition, that is, no fan effect was obtained. A similar result has been reported by Moeser (*16*).

Reder and Anderson (*17*) suggested that subjects in these experiments might be deciding about integrated facts on the basis of a consistency judgment. Rather than judging whether they had studied a particular fact, subjects may have been judging whether the fact was consistent with what had been studied. To test this hypothesis we tested subjects on their

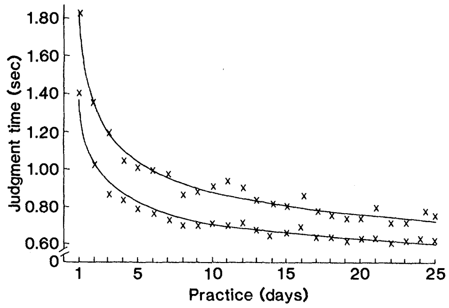


Fig. 6. Recognition times for sentences as a function of practice and fan.

memory for thematically related material under two conditions. Suppose the subject studied the following three facts:

The doctor went to the train station.
The doctor heard the conductor say "All aboard."
The doctor got off at Grand Central Station.

In the condition which tested with unrelated foils, subjects had to discriminate targets they had studied from foil sentences like

The doctor put soap in the washer.

when they might have studied *put soap in the washer* about the lawyer. In this condition it should be possible for subjects to discriminate between targets and foils by means of a consistency judgment. This is like the thematically related condition of Smith *et al.* with unrelated foils. In the other condition subjects were tested with related foils such as

The doctor got on the train.

when the predicate *got on the train* had been studied about some other individual. Subjects could not reject such foils by means of a consistency judgment and would have both to accept targets and to reject foils by a recognition judgment. We replicated the Smith *et al.* result of no-fan effect in the unrelated foil condition. In contrast, in the related-foil condition, recognition times were much longer and there was a large fan effect.

In the Reder and Anderson experiment (*17*), the subject learned either one or two themes about an individual [see also (*18*)]. In the two-theme condition, they might also learn about the doctor:

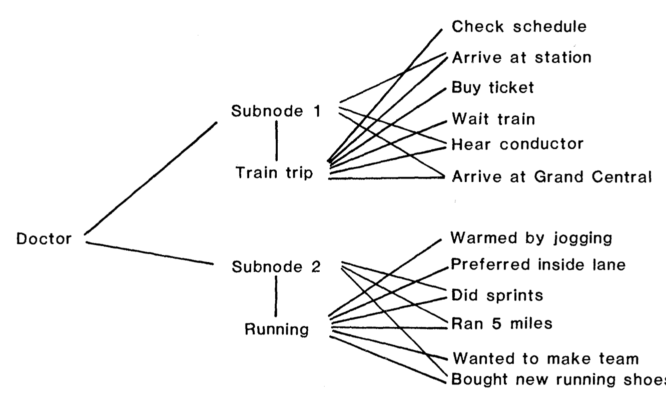
The doctor bought a new pair of running shoes.
The doctor ran 5 miles each day.
The doctor sprinted on the track.

Independent of whether subjects had to discriminate targets from related or unrelated foils, they took longer to make their judgments if two themes were associated with the doctor rather than one. We also looked at the factor of irrelevant fan. When subjects were judging a fact about one theme (for example, a fact from the running theme about the doctor), irrelevant fan refers to the number of facts studied about the other theme studied with the doctor. Although subjects were slowed by the presence of a second theme, they were not affected by the number of facts in the irrelevant theme. Again this result held whether the subjects were making their judgments in the presence of related or unrelated foils.

To account for this pattern of data, we proposed that, when encoding the information, subjects create a network representation (Fig. 7). In the presence of two themes, two subnodes, one associated with the general train-trip concept and the other with the general running concept, are created and attached to the doctor concept. These general concepts are attached to all the predicates that might be studied about an individual in the experiment. The subnodes are connected to the predicates actually studied about that individual.

When the subject hears a sentence like *The doctor arrived at Grand Central Station*, activation will spread both from *doctor* and from the concepts in the predicate. If that theme was studied about the person, there will be a large intersection of activation at that sub-

Fig. 7. Network structure proposed (*20*) for the Reder and Anderson experiment (*17*).



node. If the predicate is unrelated to the theme (that is, an unrelated foil) there will not be a strong intersection of activation. Thus, the subject can use the intersection of activation to judge whether it is plausible that the fact was studied in the presence of unrelated foils. The amount of activation arriving at the node will be affected by the presence of one or two themes, since the second theme causes activation to shift to an unrelated part of the network (the second subnode). It should not matter how many facts are attached to the relevant or the irrelevant subnode, since activation does not have to spread from the subnode to the specific predicates. Thus, the number of themes but not the relevant fan (number of facts about tested theme) or irrelevant fan (number of facts about nontested theme) should have an effect. This was observed in the presence of unrelated foils.

In contrast, in the presence of related foils it will not suffice to make a decision on the basis of activation of the subnode. Rather, the subject will have to retrieve the subnode-predicate combination. This decision will be affected by the amount of activation that spreads from subnode to predicate and hence by the relevant fan from the subnode. Thus, in the presence of related foils, number of themes and relevant fan, but not irrelevant fan, should have an effect.

Reder and Ross (*19*) have explicitly instructed subjects to judge whether test sentences were consistent with what they had studied rather than whether the actual sentence had been studied. They presented both targets and related statements not studied; subjects had to discriminate both of these from unrelated foils, which they were to judge as inconsistent. They found that fan decreased judgment time—a negative fan effect—and suggested that in the unrelated foil condition of Reder and Anderson (*17*) and in Smith *et al.* (*8*) there was a mix of consistency judgments, which produced a negative fan effect, and recognition judgments, which produced a positive fan effect. The net result was to produce no effect of fan.

In a recent simulation of the spreading activation process, I discovered that the amount of activation to spread to the subnode increases with the fan of facts out of the subnode, because these provide additional paths by which activation can converge on the subnode. For instance, in Fig. 7 activation can spread from *hear conductor* to *train trip* to *arrive at station* to subnode. Thus, for a consistency judgment based on subnode activation, subjects should be faster the greater the relevant fan.

Future Research

The theory developed in this article has been concerned with the processes leading from presentation of a probe to levels of activation in the network. The working assumption has been that level of activation maps onto latency of memory performance, but I have not set forth the processes mapping activation of network structure into performance. This is one of the frontier areas in the development of a spreading activation model. A number of recent proposals have been offered (*5, 20–22*) for a production system architecture for control of cognitive behavior. In such a scheme the steps of cognition are determined by condition-action rules. The condition of a rule specifies some pattern that should be active in working memory, and the action specifies some cognitive or external operation that will be performed if the pattern is matched. While these schemes agree on this basic architectural principle, they differ widely in the role they see for activation in determining the speed of production application. I expect that future research will focus on selecting from among these proposals.

References and Notes

1. F. I. M. Craik, *Annu. Rev. Psychol.* **30**, 63 (1979).
2. J. R. Anderson, *Cognit. Psychol.* **5**, 451 (1974).
3. ——— and G. H. Bower, *Human Associative Memory* (Winston, Washington, D.C., 1973).
4. D. A. Norman and D. E. Rumelhart, *Explorations in Cognition* (Freeman, San Francisco, 1975).
5. J. R. Anderson, *Language, Memory and Thought* (Erlbaum, Hillsdale, N.J., 1976).
6. B. Hayes-Roth, *Psychol. Rev.* **84**, 260 (1977).
7. D. R. W. King and J. R. Anderson, *J. Verbal Learn. Verbal Behav.* **15**, 587 (1976).
8. E. E. Smith, N. Adams, D. Schorr, *Cognit. Psychol.* **10**, 438 (1978).
9. S. Glucksberg and M. E. McCloskey, *J. Exp. Psychol. Hum. Learn. Mem.* **7**, 311 (1981).
10. C. H. Lewis and J. R. Anderson, *Cognit. Psychol.* **7**, 311 (1976).
11. J. R. Anderson and R. Paulson, *ibid.* **10**, 178 (1978).
12. J. R. Anderson, *Mem. Cognit.* **9**, 237 (1981).
13. J. M. Keenan and S. D. Baillet, in *Attention and Performance*, R. S. Nickerson, Ed. (Erlbaum, Hillsdale, N.J., 1980), vol. 8.
14. A. Newell and P. Rosenbloom, in *Cognitive Skills and Their Acquisition*, J. R. Anderson, Ed. (Erlbaum, Hillsdale, N.J., 1981).
15. J. C. Eccles, in *Brain and Human Behavior*, A. G. Karyman and J. C. Eccles, Eds. (Springer-Verlag, New York, 1972).
16. S. D. Moeser, *J. Exp. Psychol. Hum. Learn. Mem.* **5**, 125 (1979).
17. L. M. Reder and J. R. Anderson, *Cognit. Psychol.* **12**, 447 (1980).
18. M. McCloskey and K. Bigler, *Mem. Cognit.* **8**, 253 (1980).
19. L. M. Reder and B. H. Ross, paper presented to the Cognitive Science Society, Berkeley, Calif., August 1981.
20. J. R. Anderson, *The Architecture of Cognition* (Harvard Univ. Press, Cambridge, Mass., in press).
21. P. C. Langley and R. Neches, "PRISM's User's Manual" (Carnegie-Mellon University, Pittsburgh, 1981).
22. A. Newell, in *Perception and Production of Fluent Speech*, R. Cole, Ed. (Erlbaum, Hillsdale, N.J., 1980).

23. Supported by NSF grants GB 40298, BNS 76-00959, and BNS 78-17463. I thank M. Lewis, L. Reder, and B. Reihle for their comments on this article.