

Retrieval of Propositional Information from Long-Term Memory¹

JOHN ROBERT ANDERSON
University of Michigan

Three experiments are reported in which subjects learn propositions like *A hippie is in the park*. The experiments manipulate the number of such propositions involving a particular person (e.g., hippie) or a particular location (e.g., park). After learning the material, subjects are asked to judge whether particular probe propositions are from the study set. Times to make these judgments about probe propositions increase with the number of study propositions involving the person or location used in the probe proposition. A model is presented which assumes a subject simultaneously accesses memory from all concepts in a probe proposition and serially searches through all study propositions involving each concept. Search of memory terminates as soon as one search process from a concept finds the probe proposition or exhausts the study propositions attached to that concept.

This paper is concerned with how propositional information is retrieved from long-term memory. Three variations will be reported on a basic experimental design in which a subject commits simple facts to memory and then is asked to retrieve them. These experiments are similar in some respects to the semantic memory experiments of Collins and Quillian (1972), Landauer and Meyer (1972), Loftus (1973), Meyer (1970), Rips, Shoben, and Smith (1973), and Schaeffer and Wallace (1970). In a typical semantic memory experiment the subject is presented with a proposition (for example, *A canary has skin*) and must determine whether it is true. One reasonable model for such a task would analyze it as involving an implicit syllogism. That is, the subject retrieves relevant propositions from memory (i.e., *A canary is an animal* and

¹Experiments 1 and 2 were supported by Grant MH-13950 from the National Institute of Mental Health to Professor Gordon Bower of Stanford University. Experiment 3 was supported by Yale Laboratory funds. Peter Moore assisted in the execution of Expts 1 and 2 and Jane Yamaguchi with Expt 3. I wish to thank Gordon Bower for his general support of this empirical and theoretical enterprise. Discussions with Edward Smith and Perry Thorndyke helped me in formulating the theory. Lynne Reder greatly helped me in the writing of the paper. Requests for reprints should be sent to John Anderson, Human Performance Center, 330 Packard Road, Ann Arbor, Mich. 48104.

Animals have skin) and with these he logically determines the truth of the original assertion. This paper looks at the simpler task of answering the question of whether a test proposition resides in memory rather than of whether a test proposition is implied by other propositions residing in memory.

Recently, there has been a surge of interest in how propositional information is represented in long-term memory (i.e., Anderson & Bower, 1973; Kintsch, 1972; Quillian, 1969; Rumelhart, Lindsay, & Norman, 1972). A proposition is defined as an abstract memory representation which is structured according to certain rules of formation and which has a truth value. Although propositions are asserted by English sentences, propositions are not sentences nor are they made out of words. Rather, a proposition is a more abstract entity composed of concepts referenced by words. This distinction between propositions and concepts on the one hand and sentences and words on the other is at the heart of all propositional models of long-term memory.

Subjects in these experiments learned simple sentences like *A hippie is in the park*. It is assumed subjects store the sentences as propositions and not as strings of semantically unrelated words. There is no hard experimental evidence to support this assumption, but all subjects claimed to treat the sentences as meaning-bearing entities. There were frequent reports of imagery and semantic elaborations of the propositions. Most subjects claimed to have created for themselves a fictional world of people and places. More than one subject joked after an experimental session that he felt he was leaving old friends.

First, I will present the details of the three experiments and their results. In these experiments some important trends will appear. The three experiments were performed to test how reliable these trends were under various relevant changes of experimental conditions. After presenting the experiments, I will present a retrieval model that predicts the principal trends in the data.

EXPERIMENT 1

In these experiments subjects learn a large number of facts of the form:

- A hippie is in the park.
- A hippie is in the church.
- A policeman is in the park.
- A sailor is in the park.

They are then asked whether a particular proposition is among the designated set of facts. This situation raises interesting questions. For instance,

will the time to verify a particular proposition about a concept like *hippie* increase with the total number of propositions the subject knows about *hippie*? If so, should the same effect be obtained for other concepts in the proposition like the location *park*? The parallel activation model of Quillian (1969) would predict no effect of number of propositions for any concept. In this model, retrieval time is independent of the number of alternative facts that must be considered about a concept. Another plausible retrieval scheme is that subjects serially search through all propositions about the subject of a sentence looking for the target proposition. This model would predict an effect of the number of propositions involving the subject but not of the number of propositions involving other elements such as the locations in the above sentences. These are fundamental issues bearing on how long-term memory is organized and accessed. Experiment 1 was performed to get some preliminary data to help make a choice between alternative models.

Method

Materials. Subjects first committed to memory a data base of 26 sentences of the form *A person is in the location*. The sentences could be classified according to the total number of sentences involving the person (1, 2, or 3) and the number of sentences involving the location (1, 2, or 3). Sentences were created to realize all nine combinations of these two dimensions. The 3×3 matrix at the top of Table 1 indicates the assignment of the 26 sentences to conditions. Each sentence is represented by two letters, a lower-case letter for the person, and an upper-case letter for the location. Repetition of the same letter in a number of letter pairs means the same word occurred in a number of sentences. The persons (lower case letters) occur uniquely in the sentences of column 1, twice in column 2, and three times in column 3. A similar pattern of repetitions occurs for locations across the three rows.

After learning these sentences, subjects entered a reaction-time phase of the experiment in which they were asked to judge whether particular probe sentences came from the designated set of 26 sentences. Sentences in the designated set are called *true* and those not in the designated set are called *false*. False probe sentences were created by re-pairing persons and locations. Just as the trues, false sentences could be classified into nine conditions depending on how many originally presented (true) sentences involved the person (1, 2, or 3) used in the probe, and how many involved the location (1, 2, or 3) used in the probe. The physical sentences serving particular true or false conditions were counterbalanced over subjects so that each sentence served in all nine conditions. Because

TABLE 1
Schematic Structure of Study Materials: Sentences in Various Conditions^a

		Number of propositions per person		
		1	2	3
Number of propositions per location	1	aA bB cC	dD eE fF	gG hH iI
	2	jJ kK lL	eK rR	gJ hR iL
	3	mM nN oO	dM rN fO	gM hN iO
Condition	Trues	Falses		
1-1	aA	aD and aG		
1-2	jJ	jK and jR		
1-3	mM	mN and mO		
2-1	dD	dA and dG		
2-2	eK	eJ and eR		
2-3	rN	rM and rO		
3-1	gG	gA and gD		
3-2	hR	hJ and hK		
3-3	iO	iM and iN		

^a Lower-case letters indicate persons and upper-case letters locations.

materials vary with subjects, generalizing statistically across subjects will not involve the language-as-fixed-effects fallacy (see Clark, 1973).

In the study sentences, some words had to occur more frequently than others. To avoid having this frequency confound carry-over into the reaction-time phase, a set of nine of the sentences were chosen for special testing. Within that set of nine there were nine unique persons and nine unique locations. The nine sentences used are displayed schematically at the bottom of Table 1. From these nine trues, 18 falses were created by re-pairing the nine persons and nine locations. These falses are also displayed at the bottom of Table 1. Note that one true and two falses can be classified into each of the nine conditions defined by number of propositions per person and number of propositions per location. These conditions are referred to by a pair of numbers *i-j*, where *i* is the number of propositions per person and *j* is the number of propositions per location.

The reaction-time tests consisted of three blocks of 50 trials or 150 trials in all. In each block of 50, the 9 trues were tested twice each and the 18 falses once. Thus, the nine persons and nine locations occurred four times, twice in trues and twice in falses. They did not occur in any of the remaining 14 filler sentences in a block of 50. These sentences were composed by randomly re-pairing the seven persons and the seven locations not used in the basic trues and falses. The purpose of these filler sentences was to help keep active the subject's memory for all the material. The reaction times from these filler sentences were not analyzed. The order of sentences in each block of 50 was randomized.

The counterbalancing of sentences across conditions was accomplished by assigning different physical sentences to the schemes for different subjects. This was done by changing some of the sentences a subject studied. So, *A hippie is in the park* might fulfill the *aA* schema for one subject, the *jJ* schema for another subject, and so on. For the first subject, *park* would be *A* and would only serve in one proposition; but for the second subject, *park* would be the *J* and serve in two propositions (i.e., *jJ* and *gJ*). Note that by manipulating assignment of true sentences to schemata, assignment of false sentences to schemata is also manipulated. Nine distinct study sets were created to complete the counterbalancing.

Procedure. The experiment consisted of two principal phases: a learning phase in which the 26 sentences were committed to memory and a reaction-time phase in which subjects judged whether probe sentences were members of the designated set. During the learning phase, the subjects first studied the sentences at a self-paced rate. They were told to "try to understand fully the meaning of these sentences because this is the best way to remember them." After studying the sentences, the subject's memory was checked by having him recite answers to questions of the form *Who is in the park?* and *Where are the hippies?* One question was constructed for each of the 16 persons used in the data base and one was constructed for each of the 16 locations. So all together there were 32 questions which were presented to the subject in random order. The sentences were very imperfectly learned after the initial study and subjects made many errors. When they made errors in this training phase, they were told the correct answer. Questions to which the subject could recall the correct answers were dropped out of the set. After one pass through the set, subjects were cycled through those questions that remained. This continued until there were no questions left. Then all 32 questions were randomized and the subject started all over. This continued until the subject could correctly answer all 32 questions in a row. This training session ranged from 20 min to 1 hr.

After the subject could recite answers perfectly to these questions,

the reaction-time (RT) test phase of the experiment commenced. The subject was shown a sentence of the form *A person is in the location* and he had to judge whether it was *true* or *false*. The sentences were back-projected on a screen before the subject by a Carousel slide projector with an electronic shutter. Reaction times were measured from the opening of the shutter (projecting the sentence) to the depression of one of two response buttons. The subjects were instructed to respond as fast as was compatible with high accuracy. They were given feedback as to their correctness after each response.

Subjects. Eighteen subjects served in this experiment as a partial fulfillment of a requirement for the introductory psychology course at Stanford University.

Results

Table 2 displays the mean reaction times and error rates for the various conditions in this experiment. The errors were excluded in calculating the mean reaction times. There are a total of 108 (18 subjects \times 6 trials) observations per condition. An analysis of variance was performed using as data each subject's mean reaction time for each condition. The analysis was a factorial design with three fixed effects: truth, number of propositions per location, and number of propositions per person. The subject factor was a random effect and was used as the source of the error estimate. The three fixed effects were all significant—truth ($F(1,17) = 6.57, p < .025$), location ($F(2,34) = 6.95, p < .001$), and person ($F(2,34) = 3.54, p < .05$). None of the interactions reached statistical significance.

Of particular interest is the effect of number of propositions per concept. Averaging over trues and falses and collapsing the person and location effects together, subjects are 94 msec slower to judge a probe proposition when the concept (either person or location) is involved in three rather than one study proposition. This clearly suggests a search in long-term memory through the propositions attached to the person and location concepts. Although there are some reversals in Table 2, the general trend in going down rows or across columns is a systematic increase in reaction times. There are 36 possible comparisons and 32 are in the correct direction. The effect of one vs three propositions is actually larger for trues (119 msec) than for falses (69 msec) although this difference is not significant. The fact that trues show a greater effect is difficult to reconcile with a self-terminating search model such as the one to be presented later in this paper.

Another result of interest is the fact that the effect of number of propositions is approximately equal for the person dimension (93 msec) as

TABLE 2
 Mean Reaction Times (in Seconds) and Error Rates (Proportions in Brackets) for Expt 1

		Trues				Falses			
		Number of propositions per person				Number of propositions per person			
		1	2	3	Mean	1	2	3	Mean
Number of propositions per location	1	1.178 (.028)	1.241 (.046)	1.268 (.046)	1.229 (.040)	1.256 (.000)	1.271 (.046)	1.316 (.019)	1.281 (.022)
	2	1.278 (.074)	1.245 (.046)	1.321 (.046)	1.281 (.056)	1.335 (.019)	1.319 (.028)	1.324 (.056)	1.326 (.034)
	3	1.191 (.046)	1.240 (.028)	1.514 (.056)	1.315 (.043)	1.361 (.056)	1.389 (.037)	1.407 (.028)	1.386 (.040)
	Mean	1.216 (.049)	1.242 (.040)	1.368 (.049)	1.275 (.046)	1.317 (.025)	1.326 (.037)	1.349 (.034)	1.331 (.032)

for the location dimension (96 msec). The next two experiments are designed to investigate under what conditions this symmetry could be maintained. This question is important in deciding between two classes of search models for long-term memory: the single-access model and the multiple-access model. The single-access model claims that the subject only accesses memory from one concept (either the person or location in this experiment) and searches memory from that concept. In contrast, the multiple-access model assumes the subject enters memory from all concepts and searches simultaneously from all concepts. This multiple-access model is suggested by the symmetric increase in reaction time with number of propositions per person or per location. Clearly, the multiple-access model predicts reaction time is equally a function of amount of search from the person and from the location. However, a single-access model could produce this symmetry if it assumed subjects vary whether they enter from the person or the location, producing average effects that are equal.

EXPERIMENT 2

To account for Expt 1, the single-access model must assume that the subject was equally likely to access memory from the person or location. This is not an implausible assumption for the training procedures of Expt 1. The subjects were probed with questions both of the form *Who's in the park?* and *Where are the hippies?* The subject went on to the verification phase of the experiment only after he could consistently answer these questions correctly. As a result the subject was forced to build up a memory structure that he could access from either the person or the location.

The single-access model would expect different results if the subject had been trained only with questions of the form *Where are the hippies?* Subjects should be biased to access memory from the person. Therefore, the more potent variable in the data should be number of propositions per person. By similar reasoning, the location dimension should be more potent if the subject were trained to retrieve from location probes such as *Who is in the park?* In contrast, the strong version of the multiple-access model expects no effect due to training procedure because memory is always accessed from all concepts in the proposition.

Experiment 2 was conducted to test for this interaction between training procedure and the relative potency of the person vs location dimension. A different group of nine subjects served in each of the two training procedures to partially fulfill a requirement of the introductory psychology course at Stanford. The experiment in which they participated was identical in procedure to Expt 1 with two exceptions. First, there

were the new training procedures described above. Second, subjects were tested on the same sequences of 150 reaction-time trials twice. Thus each subject contributed 12 observations to each cell in a data matrix like Table 2.

Results

The reaction times and error rates are presented in Table 3 classified according to type of training (recall to person probe only or to location probe only), truth, and number of propositions per person and location. There are a potential of 108 observations (9 subjects \times 12 observations) contributing to each cell. An analysis of variance was performed on these data similar to the one used in Expt 1. The only difference was that training procedure was a between-subjects variable. Again, there were significant main effects of truth ($F(1,16) = 8.74, p < .01$), number of propositions per person ($F(2,32) = 9.63, p < .001$), and number of propositions per location ($F(2,32) = 33.47, p < .001$).

The training variable was not significant in its main effect or in any of its interactions. The effects of the two dimensions are about equal under either training procedure. The effect of one vs three propositions per person is 169 msec for subjects trained with the person probe. For the same subjects, the location effect is actually slightly greater—172 msec. For location-trained subjects, the effect of the person dimension is 92 msec and the effect of the location dimension is 116 msec. So location-trained subjects show about 70 msec less of an effect for either dimension, but there is no apparent interaction between training procedure and potency of the dimensions. The interaction between training conditions and location dimension is not significant ($F(2,32) = .78$) nor is the interaction between training and the person dimension ($F(2,32) = 1.37$). The reason approximately 70-msec effects are not significant is that they involve a between-subjects contrast. Thus, the apparent difference in training procedures reflects the variability between individual subjects with respect to how much they are affected by an increase in number of propositions per concept. The failure of the training procedure to have any affect on the potency of the two dimensions is quite important. It indicates that both storage and retrieval of propositional information is largely independent of the procedures under which the information is learned. This fact will have to be incorporated into a final model for this paradigm.

The interaction between the person and location dimension reached statistical significance ($F(4,64) = 2.82, p < .05$). This reflects the fact that the effects of one variable are larger for larger values of the other variable. When one variable (either number of propositions per person

TABLE 3
Mean Reaction Times (in Seconds) and Error Rates (Proportions in Brackets) for Expt 2

	Trues				Falses				
	Number of propositions per person				Number of propositions per person				
	1	2	3	Mean	1	2	3	Mean	
Person probe Number of propositions per location	1	1.050 (.065)	1.081 (.019)	1.043 (.034)	1	1.097 (.019)	1.112 (.046)	1.151 (.065)	1.120 (.043)
	2	1.030 (.074)	1.070 (.074)	1.165 (.093)	2	1.140 (.000)	1.447 (.065)	1.192 (.074)	1.260 (.046)
	3	1.046 (.065)	1.090 (.046)	1.393 (.056)	3	1.117 (.009)	1.421 (.120)	1.462 (.019)	1.333 (.049)
	Mean	1.024 (.068)	1.070 (.046)	1.213 (.056)	Mean	1.118 (.009)	1.327 (.077)	1.268 (.053)	1.238 (.046)
Location probe Number of propositions per location	1	1.052 (.056)	1.086 (.056)	1.069 (.059)	1	1.130 (.019)	1.109 (.056)	1.150 (.019)	1.130 (.031)
	2	1.038 (.028)	1.058 (.028)	1.134 (.028)	2	1.115 (.019)	1.143 (.019)	1.146 (.009)	1.135 (.016)
	3	1.095 (.046)	1.125 (.056)	1.176 (.056)	3	1.158 (.074)	1.275 (.046)	1.464 (.120)	1.299 (.080)
	Mean	1.062 (.043)	1.090 (.047)	1.126 (.050)	Mean	1.134 (.037)	1.176 (.040)	1.253 (.049)	1.188 (.042)

or number per location) is held at 1 and the other varied from 1 to 3, the mean effect is 40 msec, when held at 2 the mean effect is 108 msec, and when held at 3, the mean effect is 265 msec. A contrast testing for this trend is quite significant ($F(1,64) = 7.86, p < .01$) while the residual variance in the person-by-location interaction not accounted for by the contrast ($F(3,64) = 1.14$) is not. There had been a similar effect in Expt 1 with mean effects of 67, 37, and 177 msec, but the trend did not reach statistical significance. This trend to larger effects of one variable at larger values of another will prove to be an important piece of information in guiding the choice of a retrieval model.

The other significant interaction in this experiment was between truth and number of propositions per location ($F(2,32) = 4.14, p < .05$). The effect of 1 vs 3 propositions per location for falses was much larger (191 msec) than for trues (98 msec). A contrast testing for this difference is quite significant ($F(1,32) = 8.26, p < .01$). This nearly two-to-one difference suggests a self-terminating search. However, the effect of trues vs falses on the person dimension was not nearly so large. The mean effect for falses was 135 msec and for trues it was 127 msec.

Since the same block of 150 trials was repeated twice, it is possible to compare performance in the first vs second half of the experiment. There is a highly significant speed-up with mean reaction time 1277 msec in the first half and 1033 msec in the second half ($F(1,16) = 14.50, p < .005$). The general relationships between conditions were maintained in both halves of the experiment. In particular, the effect of number of propositions was obtained in both halves although smaller in the second half. The mean effect of 1 vs 3 propositions was 172 msec in the first half and 100 msec in the second half.

It is informative to compare the data for the false sentences in the two blocks of 150 trials. Each false sentence was repeated three times in a block, so one might suspect that the false sentences were becoming "familiar" by the second block. The false sentences behaved like the trues. The falses showed a 218-msec decrease in mean reaction time from block 1 to block 2. The effect of 1 vs 3 propositions per concept decreased from 197 to 131 msec. These facts rule out the possibility that subjects are using a strength model for memory search of the variety proposed by Atkinson and Juola (1972). According to that sort of model, subjects will accept a sentence as true if it is sufficiently familiar, search memory if it is less familiar, and reject the sentence as false if it is sufficiently unfamiliar. This model expects as falses are repeated that reaction times will increase and also the effect of number of propositions will increase. This is because the familiarity is increased, making the probability greater that subjects will have to search memory. While this sort of strength

model is ruled out, the general decrease in search rates suggests there are strength effects of a different variety operating in this situation.

EXPERIMENT 3

Experiment 2 had indicated that retrieval of information was independent of the conditions that existed during the acquisition of the information. Specifically, it was not possible to bias the relative potency of the person and location dimensions by use of training procedures. The experiment had been performed to test the single-access model which seemed to expect subjects would access memory only from the concept they had been trained with. The multiple-access model had expected subjects would access memory from both concepts.

Experiment 3 was an attempt to manipulate the relative potency of the two dimensions in such a way that the multiple-access model would expect an effect. The experiment permitted the subject to access the location concept before the person concept or vice versa. If one concept could be accessed in memory before another, search could begin from that concept before the other and reaction time would be more a function of the number of propositions attached to the head-start concept.

The training procedures in this experiment were identical to those of Expt 1. The difference concerned the reaction time testing phase of the experiment. Approximately 1 sec before the probe sentence appeared on the screen, the experimenter told the subject the person or the location, or said the word *none*. Thus, if the subject were to verify *A hippie is in the park*, he would hear one of *hippie*, *park*, or *none*. The expectation was that if the subject were precued with the person, he would be able to access that concept in memory and ready a search process from that concept before the sentence appeared. Upon presentation of the sentence, the search from this person node could begin before the search from the location node. In this way, the subject is given a "head start" from one concept. Consequently, when precued with the person, time to search memory should be more influenced by the number of propositions leading from that concept. For the same reason, when precued with the location, there should be a greater effect of the location dimension. In the control condition in which the subject is precued with *none*, there should be an equal effect of both dimensions. Further, verification times should be much slower in the control condition because it does not have the advantage of a head start in its search from one of the concepts.

The process model underlying these predictions assumes that the subject's verification time can be divided into a number of additive components. These components are (a) a time to read the sentence, (b) the time to access a concept in memory, (c) the time to search from that

concept, and (d) the time to generate a response. Component (b) is eliminated for the precued concept. As soon as the sentence is read, search can begin from that concept. In this way the search from the precued concept has a head start over the search from the noncued concept which can only begin after that concept is accessed in memory.

The precuing variable was manipulated within subjects. Each subject went twice through the sequence of 150 test sentences used in Expts 1 and 2. For each subject, one-third of the propositions were precued with person, one-third with location, and one-third with *none*. Three separate cueing sequences were constructed to counterbalance over subjects the assignment of cueing conditions to sentences. There were 54 experimental conditions in a completely within-subject, factorial design—three levels of the person dimension, three of the location, two values of truth, and three cueing conditions. Each subject contributed four observations to each condition. There were 92 filler sentences. Twenty-seven subjects participated in this experiment as partial fulfillment of a requirement in the introductory psychology course at Yale University.

Results

The mean reaction times in milliseconds for the 54 conditions of the experiment are displayed in Table 4. An analysis of variance, similar in design to that for Expt 1, was performed on the data. Again, there were significant main effects of truth ($F(1,26) = 25.38, p < .001$), the person dimension ($F(2,52) = 20.12, p < .001$), and the location dimension ($F(2,52) = 7.00, p < .025$). Precueing produced a very significant ($F(2,52) = 162.04, p < .001$) speedup in subject reaction times. Subjects were on the average 254 msec faster when precued. This is similar to the 250-msec advantage found by Freedman and Loftus (1971) when they precued with a category.

The expected interaction between cueing condition and the location dimension was only marginally significant ($F(4,104) = 2.35; p < .10$). The effect of one vs three propositions per location is largest when precued with location (120 msec), next largest when precued with the person (89 msec), and least in the control condition (78 msec). Although the interaction is not significant ($F(4,104) = 1.05$) between cueing conditions and the effect of the person dimension, the ordering of the effect across conditions is as expected. The biggest effect occurs when subjects are precued with person (110 msec), next biggest in the control condition (78 msec), and least with location precueing (60 msec). Averaging together the effects of location and person precueing, the average effect for the cued dimension is 115 msec, and for the noncued dimension it is 75 msec. The significance of this 40-msec difference can be evaluated

TABLE 4
Results from Expt 3

A. Precue with none

		Trues				Falses			
		No. of propositions per person				No. of propositions per person			
		1	2	3	Mean	1	2	3	Mean
No. of proposi- tions per location	1	1231	1312	1406	1316	1314	1376	1424	1371
	2	1335	1422	1259	1339	1407	1490	1488	1462
	3	1282	1406	1366	1352	1403	1574	1494	1490
	Mean	1283	1380	1344	1336	1375	1480	1469	1441

B. Precue with person

		Trues				Falses			
		No. of propositions per person				No. of propositions per person			
		1	2	3	Mean	1	2	3	Mean
No. of proposi- tions per location	1	973	1059	1119	1050	1036	1114	1205	1118
	2	1115	1120	1158	1131	1108	1163	1254	1175
	3	1048	1203	1137	1129	1192	1200	1258	1216
	Mean	1045	1127	1138	1103	1112	1159	1239	1170

C. Precue with location

		Trues				Falses			
		No. of propositions per person				No. of propositions per person			
		1	2	3	Mean	1	2	3	Mean
No. of proposi- tions per location	1	954	1042	1083	1026	1044	1105	1031	1060
	2	1121	1202	1128	1150	1221	1160	1284	1222
	3	1066	1128	1168	1121	1132	1285	1201	1206
	Mean	1047	1124	1126	1099	1132	1183	1172	1163

using the subject \times person \times location \times cueing interaction as an error term. The difference is only marginally significant ($F(1,208) = 3.10$, $p < .10$). It is puzzling how precueing can have such a large effect on overall reaction times but such a weak effect on search rates.

None of the other interactions approached statistical significance. Nonetheless, it is worthwhile to note some of the nonsignificant comparisons. The effect of the location dimension (96 msec) was only slightly larger than the effect of the person dimension (83 msec). The average effect of number of propositions for falses (103 msec) is larger than for trues (75 msec). The other interesting comparison concerns the effect of varying one dimension (either person or location) from one to three propositions for different values of the other dimension. The effect of one dimension is 107 msec when the other dimension is held at one, 89 msec when it is held at two, and 71 msec held at three. Thus, the trend, while very weak, is towards lesser effects of one dimension for higher values of the other. This is just the opposite of strong trends obtained in Expts 1 and 2. In Expt 2, the effect had been quite significant.

A RETRIEVAL MODEL

Having now examined the particulars of each experiment, a model will be presented that captures the main trends in the experiments. To facilitate an evaluation of this model, Table 5 presents the means of the reaction times obtained in Expts 1 and 2 and in the control condition of Experiment 3.² Overall there are 63 subjects contributing 432 observations to each of the 18 conditions displayed in Table 5.

The following model is proposed for the retrieval processes that subjects were performing in these experiments:³ It is assumed that in order to determine whether a proposition is stored in memory, the subject accesses memory from all the concepts in the proposition. To determine whether the target proposition is attached to a concept, the subject must serially search through the propositions attached to the concept. This assumption of a serial search through propositions is consistent with the strongest and most consistent trend in the data—the increase in reaction time with number of propositions attached to person or location. This increase with number of propositions per concept is highly stable across subjects. Since sentences are assigned to different conditions for different subjects, the trends are also highly stable across sentences. For instance, for 58 of the 63 subjects, the [1,1] cell is faster than the [3,3]. This

² For sake of brevity, I will not present a model for the pre-cue conditions of Expt 3. However, such a model is available and the interested reader should write to me.

³ This model was first developed by Anderson & Bower (1973; Chap. 12) as part of our theory of long-term memory, HAM. These three experiments were designed to test certain aspects of that theory. Other tests appear in Thorndyke & Bower (1974). Both Anderson & Bower and Thorndyke & Bower are concerned with the details of how a proposition is represented and searched. Such details will not be considered here, however, because these experiments are not sensitive to the details of representation. Rather, whole propositions will be considered as the units of analysis.

TABLE 5
Mean Verification Times and Error Rates for Trues and Falses

		Trues				Falses			
		Number of propositions per person			Mean	Number of propositions per person			Mean
		1	2	3		1	2	3	
Number of propositions per location	1	1.111 (.051)	1.174 (.042)	1.222 (.046)	1.169 (.046)	1.197 (.019)	1.221 (.042)	1.264 (.030)	1.227 (.030)
	2	1.167 (.065)	1.198 (.056)	1.222 (.060)	1.196 (.060)	1.250 (.014)	1.356 (.037)	1.291 (.044)	1.299 (.032)
	3	1.153 (.063)	1.233 (.044)	1.357 (.054)	1.248 (.054)	1.262 (.042)	1.471 (.079)	1.465 (.051)	1.399 (.057)
	Mean	1.144 (.058)	1.202 (.048)	1.267 (.054)	1.204 (.053)	1.236 (.025)	1.349 (.053)	1.340 (.042)	1.308 (.040)

orderly increase in verification times reduces the credibility of a Quillian-like model (see Quillian, 1969) which posits that search proceeds in parallel along all paths from a concept, unaffected by the number of paths. Rather, increasing the number of paths (propositions) out of a concept results in marked increases in verification time. The overall effect in Table 5, comparing the [1,1] and [3,3] cells, is of the order of 250 msec, which is very large on a base reaction time of just over 1 sec. While this effect is not compatible with a Quillian-like parallel model, it is compatible with a parallel model whose search rate is slower in proportion to the number of paths searched (see Townsend, 1971).

While search is serial from a particular concept, it is assumed that a subject can access memory in parallel from a number of concepts. There were four principal sources of support for this assumption in these experiments. First, there is the symmetric effect obtained of the location and person dimensions. Second, there is a failure of Expt 2 to find any effect of training on the relative potency of the two dimensions. Third, there is the result obtained in Expt 1 and 2 (but not Expt 3) that the effect of one dimension was greater for larger values of the other dimension. To appreciate why the multiple-access model expects this result, consider the fact that there are two search processes proceeding in parallel and that the search time will be determined by the fastest. Clearly, the search time will be principally (but not completely) affected by the process that has the fewest proportions to search. Therefore, when the number of propositions is held at one, there can be little effect of varying the number of propositions for the other concept from one to three. There is much more "room" for an effect when one concept is being held constant at three propositions.

It is unfortunate that a reverse effect was obtained in Expt 3. That reversal was small and insignificant and should probably be regarded as a random deviation from an effect in the opposite direction. That is, the effect of 1 vs 3 propositions for one dimension was 36 msec less when the other dimension was held at three than when it was held at one. The .05 confidence limits for the effect are from +65 to -137 msec. Therefore, we cannot reject the hypothesis that the true effect in Expt 3 was positive. Similarly, Expt 1 found a positive effect of 110 msec but could not reject the hypothesis that the true effect was zero or negative. Only in Expt 2, where the obtained effect was +215 msec, could we reject (at the .01 confidence level) the hypothesis that the true effect was zero or negative. Thus, the only possible conclusion consistent with all three experiments is that the true effect is positive.

The multiple-access model conceives of the overall retrieval process as a race between the person and location search processes. The one way

to affect the relative potency of the person and location dimensions would be by giving one search process a head start in the race. This is what was done in Expt 3 by precueing, and the desired effect was obtained: The precued dimension showed a greater effect of number of propositions. This is the fourth and final piece of evidence indicating the multiple-access model. Unfortunately, this effect was only of marginal statistical significance.

The obvious alternative to the multiple-access model is the single-access model, which assumes memory is searched from just one concept. This model can only deal with the symmetric effects of the person and location dimensions by the ad hoc assumption of an equal probability of accessing memory from either concept. The failure of the training procedure in Expt 2 to affect this symmetry strains the credibility of that assumption. Finally, the single-access model has no way to explain the increasing effect of one dimension at higher values of the other. This interaction is quite strong in the average data of Table 5: When one variable is fixed at one proposition, a change in the other from one to three produces a 71-msec increase in reaction time; when fixed at two, the increase is 101 msec; and when fixed at three, it is 186 msec. Such an interaction embarrasses the single-access model because it assumes that on any particular trial a subject only searches from one concept.

The final theoretical decision is whether subjects terminate search as soon as they find the target proposition or whether they search through all the propositions attached to a concept. This is the classical question of exhaustive versus self-terminating search (see Sternberg, 1969). A self-terminating model expects a lesser effect of number of propositions for trues than for falses. The average effect in Table 5 of one vs three propositions for trues is 102 msec, while for falses it is 138 msec. Although there is less of an effect for trues, the ratio of effects is 1.35 to 1, which is considerably less than the 2 to 1 effect commonly expected for self-terminating models.

One should note, however, that there is a strong positive correlation ($r = .83$) between reaction times and error rates for the false data whereas there is no correlation for the true data ($r = -.12$). This difference between true and false data becomes important if one considers the speed-accuracy tradeoff (see Pachella, 1974) which appears to be ubiquitous in reaction-time experiments. The phenomenon is that subjects' reaction times decrease whenever their error rates increase. With generally low error rates such as those obtained in this experiment, small increases in error rates can produce marked decreases in reaction-time. Since longer reaction times in the false data were correlated with higher error rates, this means that the reaction times would be even

larger in long reaction time conditions if subjects were working at the same accuracy as they did in the conditions that produced shorter reaction times. The consequence of this speed-accuracy tradeoff and the differing correlations for trues and for falses is that the obtained ratio of effects for trues vs falses is underestimated.

There are no adequate methods for dealing with the speed-accuracy tradeoff except to develop a model of how errors are made. The data here are inappropriate for developing such a model. It would have been necessary to run subjects under different deadline conditions. So I will simply provide separate fits of this model to trues and to falses. A self-terminating model will be assumed for trues because intuitively this is a more reasonable search algorithm. Given the different correlations between errors and reaction times, this decision to provide separate fits is clearly justified. However, because of it, the question of self-terminating vs exhaustive search is left empirically undecided. Research is now underway to obtain data more suitable to testing a model of errors which, as a consequence, will permit a decision on whether long-term memory is searched in an exhaustive or self-terminating fashion.

To summarize the search model: The subject simultaneously accesses and searches memory from all concepts in the proposition. However, the search from one concept is a serial search of all propositions involving that concept, looking for the target proposition. The search terminates as soon as any search process uncovers the target or exhausts the propositions attached to it. Thus search time is determined by the fastest of a number of racing processes.

THE MATHEMATICAL MODEL

I have now described verbally the search model and indicated how it accounts for the main trends in the data. What remains to be done is to formalize mathematically that model's predictions for these experiments. Then one may obtain a quantitative comparison of the model's predictions and the actual data. Suppose that the mean time to search a proposition is α and suppose i propositions must be searched from the person and j propositions from the location. The mean time to search from the person concept alone would be $i\alpha$ and from the location alone $j\alpha$. However, since searches are proceeding from both in parallel and search time is determined by the fastest, we would expect mean search time to be less than either $i\alpha$ or $j\alpha$. In order to determine the mean time of the fastest of a number of processes, a characterization is needed of the variability in the times for each process. Therefore, the following model is proposed for the variability of search times from a concept. It has three

features to commend it: (a) It is not unreasonable *a priori*; (b) It is mathematically tractable; and (c) it predicts the data.

In understanding the following development, the reader should keep in mind that there are three levels of analysis: First, there is the question of the variability in times to examine one proposition. Second, there is the question of how search times for individual propositions become compounded together to yield a model of the variability in times to search n propositions attached to a concept. Finally, a number of search processes are proceeding in parallel from all concepts in a test proposition. A characterization is needed of the variability in the "winning" time of these processes.

The natural assumption is that the time, t , to search one proposition has an exponential probability density, i.e.,

$$f(t) = (1/\alpha)e^{-t/\alpha}. \quad (1)$$

The density in Eq. (1) has mean α which is the time to search one proposition. How should one proceed from Eq. (1) to an equation giving the density of times to search n propositions? There are two obvious density functions—one assumes complete independence and the other complete dependence among the search times for individual propositions. Suppose that, *within* the search from a concept, time to search one proposition is independent of the times to search the other propositions. Then the density of times to search n propositions would be a gamma distribution with parameters n and α . There is no tractable characterization of the fastest of an arbitrary number of gammas which discourages the assumption that times to search individual propositions are independent.

In contrast, a complete-dependence assumption leads to a workable model. Assume that a particular search process examines propositions at a constant rate but that the rates vary across applications of search processes to particular concepts. Let r be the search rate, i.e., the number of propositions searched per millisecond. It follows from Eq. (1) that r has the following density:

$$g(r) = (1/\alpha r^2)e^{1/r\alpha}. \quad (2)$$

From Eq. (2) we can derive the density of t_n , the time to search n propositions. Using the fact $t_n = n/r$ and computing a change of variable, Eq. (3) may be obtained from Eq. (2):

$$f(t_n) = (1/\alpha n)e^{-t_n/\alpha n}. \quad (3)$$

Equation (3) describes a family of exponential densities. For each value of n , there is a different exponential with the mean $n\alpha$. Equation (1) is a

special case of Eq. (3) where $n = 1$. Equation (3) could have been asserted directly, but the derivation through Eq. (2) makes two features salient: (a) The *sample space* for the density in Eq. (3) is searches from concepts, and (b) within a search, the times to search individual propositions are identical.

It is easy to characterize the fastest of an arbitrary number of exponentials. The fastest of i exponentials with rates r_1, r_2, \dots, r_i is also an exponential, but with rate $r_1 + r_2 + \dots + r_i$. Using this fact, Eq. (4) gives the mean time, t_M , of the fastest of i process searching n_1, n_2, \dots, n_i propositions:

$$t_M = \frac{1}{\frac{1}{\alpha n_1} + \frac{1}{\alpha n_2} + \dots + \frac{1}{\alpha n_i}} \quad (4)$$

Equation (4) is central to the predictions of the model for verification data.

To review the significant features of this search model:

(a) The search involves a mixture of *serial* and *parallel* processes. Within the search from one concept, search is serial, examining one proposition after another. However, different search processes proceed in parallel from all concepts of a probe proposition.

(b) The model is self-terminating in two senses. First, a particular search process terminates immediately upon finding a matching memory structure. Second, the search of memory terminates as soon as the first of the racing processes comes to a conclusion.

(c) The model involves a mixture of independent and dependent processes. The times for searching individual propositions from a concept are totally dependent. The times for processes from different concepts are totally independent. The assumption of exponential distributions of search times from one concept is not that critical. I have worked through the assumption of the gamma distribution and obtained similar predictions. However, the theoretical expressions are terribly complicated and specific to this experiment. In contrast, Eq. (4) is general and can be used for any retrieval experiment.

PREDICTIONS

With this development of the race model, it is now possible to obtain predictions for the data in Table 5. It is easier to work out the predictions for the falses than for the trues. Suppose that the subject must exhaust i propositions from the person concept or else j from the location to determine that the sentence is false. Derived from Eq. (4), the following equation gives the expected time to falsify an i, j proposition:

$$F_{i,j} = K_F + \frac{1}{\frac{1}{i\alpha_F} + \frac{1}{j\alpha_F}} = K_F + \frac{ij}{i+j} \alpha_F. \quad (5)$$

In Eq. (5) K_F represents the constant time associated with encoding the sentence and generating a false response. It does not involve memory search and hence is independent of i and j . The parameter α_F is the time to search a proposition for falses. Because of the pattern of correlations between errors and reaction times, separate α parameters are being estimated for trues and for falses.

In deriving predictions for trues, it is necessary to consider the order in which the propositions involving a concept are considered and the position of the target proposition in that ordering. If the target proposition is considered m th among the i person propositions and n th among the j location propositions, then the search time for the target proposition will be the same as for a $[m,n]$ false. This is because search from a concept is self-terminating. As a simplifying assumption, suppose all combinations of m and n (i.e., all combination of positions in the two orderings) up to i and j are equally likely. From this assumption a prediction for the true sentences can be derived:

$$T_{i,j} = K_T + \frac{1}{ij} \sum_m^i \sum_n^j \frac{mn}{m+n} \alpha_T. \quad (6)$$

The constant K_T in Eq. (6) represents the time to encode the sentence and to generate a true response. The parameter α_T is the estimate for trues of time to search a proposition.

Figure 1 illustrates the predictions of the model together with the data from Table 5. In Fig. 1, i gives the number of propositions per person and j the number of propositions per location. The parameters were obtained by means of multiple regression analysis, and their values are $K_T = 844$ msec, $K_F = 1053$ msec, $\alpha_T = 518$ msec, and $\alpha_F = 280$ msec. The model clearly captures the main trends in the data. The correlation between prediction and data is .95. However, there are a number of seriously discrepant points. Each of these points is associated with a reversal in Table 5 of the ordering of reaction times down rows or across columns. There are four such reversals: In the true matrix, the (1,2) point is 14 msec larger than the (1,3) point, and the (3,1) point is equal to the (3,2) point. In the false matrix, the (2,2) point is 65 msec larger than the (3,2) point, and the (2,3) point is 6 msec longer than the (3,3) point. Thus, the source of these discrepancies is not in the multiple-access assumption or in the self-terminating assumption. Any model that assumed an orderly increase in reaction time with number of

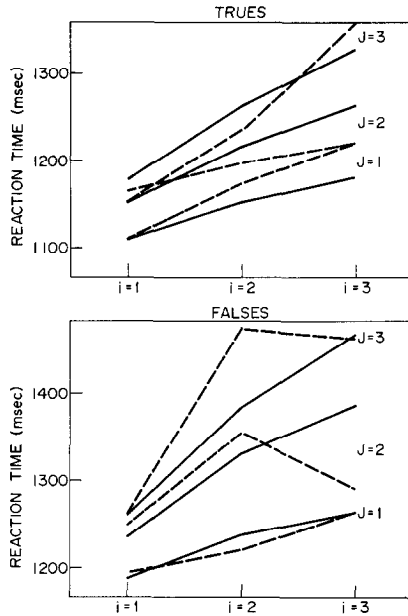


FIG. 1. Comparison of the model's predictions with the average data in Table 5. The solid lines are the predicted times and the dotted lines the observed. The variable i indexes the number of propositions per person and j the number of propositions per location.

propositions would face these difficulties. However, the overall trend in favor of this assumption is very strong, and it seems unwise to give up the assumption on the basis of four discrepant comparisons. The hypothesis is confirmed by the 32 other possible comparisons in Table 5 between pairs of reaction times (along rows or down columns of the matrices).

CONCLUSIONS

The general force of these experiments has been to support two assumptions of the Anderson and Bower (1973) model of retrieval from long-term memory. These assumptions are the serial search of propositions leading from any one concept and the multiple-access in parallel to memory from all concepts in the proposition. The correlation between reaction times and error rates vitiates any conclusions about self-terminating search. There were disturbing inconsistencies of detail from experiment to experiment and perplexing misorderings of conditions. In part, these are chance perturbations that can be expected across experiments with large numbers of conditions. The general trends are

clear, however, and consistent with the data in Anderson and Bower (1973; Chapter 12), Thorndyke & Bower (1974), and more recent unpublished data of my own. These general findings, then, are ones that will have to be reckoned with by any theory of long-term memory.

REFERENCES

- ANDERSON, J. R., & BOWER, G. H. *Human associative memory*. Washington: Winston and Sons, 1973.
- ATKINSON, R. C., & JUOLA, J. F. Factors influencing speed and accuracy in word recognition. In S. Kornblum (Ed.), *Attention and performance*. Vol. IV. New York: Academic Press, 1972.
- CLARK, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 335-359.
- COLLINS, A. M., & QUILLIAN, M. R. Experiments on semantic memory and language comprehension. In L. Gregg (Ed.), *Cognition and learning*. New York: Wiley, 1972.
- FREEDMAN, J. L., & LOFTUS, E. F. Retrieval of words from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 107-115.
- KINTSCH, W. Notes on the semantic structure of memory. In E. Tulving and W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press, 1972.
- 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. Category dominance, instance dominance, and categorization time. *Journal of Experimental Psychology*, 1973, 97, 70-74.
- MEYER, D. On the representation and retrieval of stored semantic information. *Cognitive Psychology*, 1970, 1, 242-300.
- PACHELLA, R. G. An interpretation of reaction time in information processing research. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. New York: Lawrence Erlbaum Associates, 1974.
- QUILLIAN, M. R. The teachable language comprehender. *Communications of the Association for Computing Machinery*, 1969, 12, 459-476.
- RIPS, L. J., SHOEN, E. J., & SMITH, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 1-20.
- RUMELHART, D. E., LINDSAY, P. H., & NORMAN, D. A. A process model for long-term memory. In E. Tulving and W. Donaldson (Eds.), *Organization and memory*. New York: Academic Press, 1972.
- SCHAEFFER, B., & WALLACE, R. The comparison of word meanings. *Journal of Experimental Psychology*, 1970, 86, 144-152.
- STERNBERG, S. Memory-scanning: Mental processes revealed by reaction-time experiments. *Acta Psychologica*, 1969, 30, 276-315.
- THORNDYKE, P. W., & BOWER, G. H. Storage and retrieval processes in sentence memory. *Cognitive Psychology*, 1974, 6, 515-543.
- TOWNSEND, J. T. A note on the identifiability of parallel and serial processes. *Perception and Psychophysics*, 1971, 10, 161-163.

(Accepted January 21, 1974)