Memory & Cognition
1981, Vol. 9 (3), 237-246

Effects of prior knowledge on memory for new information

JOHN R. ANDERSON
Carnegie-Mellon University, Pittsburgh, Pennsylvania 15313

Three experiments examined the effects of prior knowledge on the learning and retention of new information. Subjects learned varying amounts of prior knowledge about individuals referred to by first name/last name combinations. In the first two experiments, subjects more rapidly learned new information about individuals for whom they were given prior knowledge, retrieved this information more slowly, and showed smaller interference (fan) effects. This complex of results is predicted by a spreading activation model in which we assume subjects try to integrate the prior knowledge with the new information in a network fashion. The third experiment, in addition to including conditions of prior vs. no prior experimental knowledge, included well-known names like Ted Kennedy, about which subjects have a great deal of preexperimental prior knowledge. The relationships between conditions of experimental prior knowledge and no prior knowledge obtained in Experiments 1 and 2 were replicated, but the well-known names did not behave simply like the extreme of experimental prior knowledge. In particular, subjects showed the fastest verification of new facts learned about the well-known names, rather than the slowest, as predicted from the spreading activation network model.

Many interference analyses of memory imply that acquiring new information about a concept should become more difficult with each new fact learned about that concept. This interference should manifest itself in many ways. First, the learning of each new fact should suffer increasing negative transfer from the earlier facts. Second, a fact's retention, once learned, should be proactively interfered with by the earlier facts. Third, its retention will suffer retroactive interference from subsequently learned facts. These interference effects have traditionally been documented by looking at a percent recall measure. More recently, these interference phenomena have been studied using a reaction time methodology. It has been repeatedly demonstrated that, the more other facts the subject has learned about a concept, the slower he is to recognize an experimentally learned fact involving that concept. This phenomenon has been interpreted (Anderson, 1976) within a spreading activation network theory called ACT. Each concept is considered to be a node in a memory network, and the facts learned about the concept are encoded by network paths leading from the node. Recognition depends upon activating the path that encodes the fact. By increasing the number of facts (paths) leading from a node, the rate at which activation can spread down a path is decreased. This interference in reaction time is referred to as the "fan effect," because of the fan of paths leading out of the node.

This research was supported by Grant BNS78-17463 from the National Science Foundation. Thanks are due to Dick Proudfoot for programming Experiments 1 and 2 and to Takashi Iwasawa for programming Experiment 3 and helping in all aspects of data analyses. Lynne Reder was also very helpful with her comments on an earlier version of the manuscript.

Some recent research has uncovered exceptions to the principle of more information, more interference. Hayes-Roth (1977) has demonstrated that highly overlearned material appears to be immune to such interference effects. Moeser (1979), Reder and Anderson (1980) and Smith, Adams, and Schorr (1978) in a series of studies have shown that facts that are highly thematically consistent do not interfere with each other in a recognition paradigm if they have to be discriminated from facts that are inconsistent with the theme.

Counterintuitive Aspect of Interference

It is quite counterintuitive to claim that the more people know about a concept, the slower they will be to retrieve a fact about the concept. This would seem to predict, for instance, that subjects would be faster to verify "Birch Bayh is a senator" than "Ted Kennedy is a senator." In a pilot experiment, I found that subjects took 2,156 msec to verify predicates about four well-known figures, such as the Kennedy fact above, but 2,501 msec to verify the same predicates of four less well-known people, such as the Bayh fact. Unfortunately, there are a number of possible explanations of such a result with uncontrolled natural material: (1) Subjects are faster at encoding graphemes like "Ted Kennedy" because of greater exposure. (2) The strength of the "backpath" connecting "senator" to "Ted Kennedy" is much stronger than the backpath from "senator" to "Birch Bayh." That is, these materials cannot control for the strength of connection from predicate to subject. (3) The degree of learning of "Ted Kennedy is a senator" may be so high that it is

  0090-502X/81/030237-10$01.25/0

no longer susceptible to interference (Hayes-Roth, 1977). (4) In one way or another, the node in the memory corresponding to Ted Kennedy has greater capacity. In the ACT framework, this might be greater capacity for spreading activation.

One function of the present experiments will be to sort out these possible explanations. The other function will be to test the predictions of the ACT theory about the effects of prior knowledge about a concept on the learning and retrieval of new knowledge about the same concept. As will be shown, the ACT theory predicts that prior knowledge has both interfering and facilitating effects. To obtain a firm manipulation of amount of prior knowledge, the first two experiments will provide subjects with differing amounts of prior knowledge. The third experiment will see how these results extend to prior knowledge acquired outside the laboratory.

Experimental Paradigm

Over the course of the experiment, subjects were introduced to 20 individuals referred to by first and last name. These 20 individuals could be categorized into four groups of five according to amount of prior information learned about them. The experiment started with the subject's being introduced to individuals in two of these groups. For individuals in one group, subjects read a rather elaborate description of one aspect of their lives. An example is reproduced below.

"Carol Norman owns a used furniture and antique shop on State street. She is a lawyer by profession and her shop contributes to her income only by being a tax loss. Carol keeps the shop because it provides a convenient basis for her primary hobby which is collecting and refurbishing old furniture and mechanical devices such as sewing machines and clocks. Carol has a particularly fine collection of grandfather clocks. Many of these clocks were bought by her at tag sales and auctions which she frequents. Each new grandfather clock that she buys provides the dual opportunity both to repair and refinish the clock housing and to tinker with the clock mechanism. Carol is planning to have a display of the best of her clocks at next year's antique show in the New Haven coliseum."

Subjects read a single sentence about individuals in the second group. An example is: "Henry Caputo bought a season ticket to the New Haven Night Hawks hockey games." These two conditions were referred to as the "paragraph" and the "sentence" conditions, respectively.

There was a third set of five names about which the subject learned no prior facts. Subjects were told the names of these five individuals before the experiment, however. This is referred to as the "name" condition. To equate these three conditions (paragraph, sentence, and name) with respect to ease of encoding the name graphemes, a recognition training phase was instituted in which subjects were drilled until they were equally fast at recognizing the names in all three conditions. There was a fourth group of five names to which subjects were not given prior exposure, nor were they given any encoding training. Therefore, this final set of names is said to be in the "unfamiliar" condition.

All 20 names then appeared in a location learning phase in which subjects learned that the people were in various locations. The locations were unrelated to facts previously learned about the individuals. Some of the individuals were in just one location and some were in two locations, to give a manipulation of fan. Thus, the material can be classified by the combination of four conditions of differential prior experience and the two fan conditions.

There are two dependent measures of interest. One was the time to learn the locations. The second was the time to recognize the fact that an individual was in a location, once this fact was learned. This learning rate measure is traditional in interference research, whereas the recognition time measure is the measure of primary interest in work on the fan effect.

ACT's Predictions

The materials can be ordered on a dimension of amount of prior knowledge from paragraph to unfamiliar. It is easiest to develop ACT's predictions for the contrast between the paragraph and name conditions. Figure 1 provides a schematic network representation for the name condition (Part a) and the paragraph condition (Part b). Dark solid lines represent network structure before the location learning phase. Light

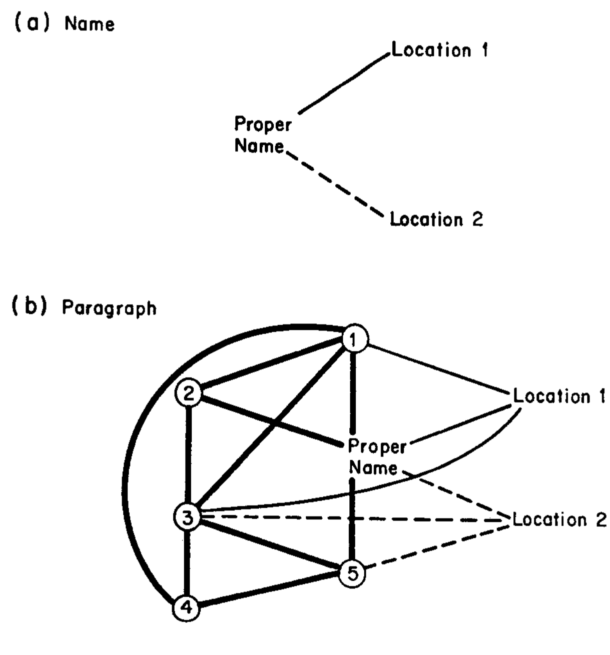

Figure 1. Schematic network representation of knowledge structure in name condition (Part a) and paragraph condition (Part b). Thick lines represent encoding of prior knowledge, solid lines represent the encodings of the first location, and dotted lines represent the encoding of the second location.

solid lines represent information used to encode the first location, and dotted lines represent information added to encode the second location (if there was one). In both conditions, there is a connection between the name and a node standing for the location. However, in the paragraph condition, there is also a large amount of additional prior structure encoding the paragraph.

The ACT theory expects that the location information will be differently encoded in the paragraph condition. This is because of the existence of an elaborative process (Anderson, 1976; Anderson & Reder, 1979) that adds to the encoding of a proposition by making connections to other facts known about the concept and by using these facts as a basis for embellishing the to-be-learned material. For instance, if Carol Norman is the tinker and is associated with being in the bank and the church, one might elaborate that she went to the bank to borrow money to buy an antique clock and she went to the church to pray that she would win at the New Haven antique show. Thus, the location propositions in the paragraph condition (Part b of Figure 1) are represented with integrating propositions connecting them to the prior knowledge about the concept. In contrast, the location propositions are directly encoded only in the name condition.

On the basis of these representations, a number of predictions follow about rate of location learning and recognition times. First, material should be more rapidly learned in the paragraph condition because there are redundant paths by which the location information can be inferred if it cannot be directly retrieved. Thus, even if the subject fails to encode that Carol Norman is in the bank, he might be able to infer that this was the location if he could recall the elaboration that "Carol Norman borrowed money from the bank to buy an antique clock." Thus, in terms of a learning measure, we should see positive transfer.

A second prediction is that despite the faster learning in the paragraph condition, subjects should show slower reaction times to recognize the material in the paragraph condition. This is because all the paragraph paths should dissipate activation from the paths encoding the location. There is no dissipation in the name case. Thus, we predict a dissociation between learning rate and reaction time measures. The former should show facilitation and the latter, interference of prior knowledge.

The third prediction of the ACT model, concerning the fan effect, is more complex. Note in Part b of Figure 1 that, because there are paths to the location through the prior knowledge, activation that spreads to the prior knowledge will eventually spread to the locations. Thus, some proportion of the activation that spreads from PROPER NAME to Node 1 in Figure 1b will spread to Location 1. Indeed, some of the activation spreading to Node 1 will get to Location 2 via the intermediate Node 3.

The exact predictions depend on the exact details of the network representation, but the following is an attempt at an approximate analysis[1]: Assume that the total strength of prior paths out of the node is K. Let the strength of each path to a location be s. Then, the strength of all paths out of the proper name is $K + ns$, where n is the number of locations, or fan. Since the ACT theory assumes a fixed amount of activation can spread from a node and the amount spread down any path is proportional to its strength, the amount of activation spread directly to a location is $A[s/(K + ns)]$, where A is the constant of proportionality. Assume that a fraction f of the activation spread to prior knowledge is indirectly spread to the location by the circuitous routes. Then, the indirect activation is $A[fK/(K + ns)]$, and the total activation, indirect and direct, is $A[(fK + s)/(K + ns)]$. According to the ACT theory, retrieval time is a linear function of the inverse of this quantity, or:

$$\begin{aligned} RT &= a + b[(K + ns)/(fK + s)] \\ &= a + [bK/(fK + s)] + [bns/(fK + s)]. \end{aligned} \tag{1}$$

The third quantity, $bns/(fK + s)$, reflects the fan effect, since it involves n in the numerator. Since it involves K in the denominator, the prediction is that the greater the prior knowledge (i.e., K), the smaller will be the fan effect. Thus, for instance, the prediction is a smaller fan effect in the paragraph than the name condition.

The reader may notice that Equation 1 does not universally imply that subjects will be slower overall with more prior knowledge, which was earlier given as an informal prediction of the ACT theory. If n is large enough or f is large enough, or some combination thereof, this function will decrease with increasing K. In particular, if $fn > 1$, the effect of K is reversed. However, it is reasonable in these experiments, in which n only varies up to 2, to suppose we are in the range in which longer RTs would be predicted with more prior knowledge.

In summary, ACT offered a rather intriguing array of predictions about the effects of prior knowledge: With more prior knowledge, there should be more rapid learning and slower reaction times, but weaker fan effects.

EXPERIMENTS 1 AND 2

For various reasons, it seemed necessary to perform two closely similar versions of an experiment. Because of a programming error, the first experiment had failed to gather some important data about learning rates. Also, since the first experiment had a very difficult name learning phase that many subjects did not finish, we were interested in seeing what the results would be like with a less select set of subjects. Finally, there was the desire to have the confirmation of a replication. Because the two experiments are so similar, I will describe them together, noting the few points at which the methodologies differed.

Method

Subjects. Subjects were recruited from the general subject pool at Yale University. This consisted mainly of undergraduates. Thirty-four subjects were recruited for the first experiment. Of these, 22 managed to finish the name learning phase and were included in the final analysis. Twenty-seven subjects were recruited for the second experiment. Of these, 24 managed to finish the name learning phase and were included in the final analysis. The experiment lasted about 3 h, and subjects were paid $7.50 for their participation. The experiment involved a completely within-subjects design.

Materials. The 20 names were composed of a set of 20 common male first names (five or six letters in length) and 20 last names (six or seven letters in length) suggesting a variety of ethnic backgrounds. The first name/last name combinations were created randomly from this stock for each subject. The 20 names were also randomly assigned to the four conditions (paragraph, sentence, name, and unfamiliar), with 5 names/condition.

Five paragraphs and five sentences were used in the prior learning phase. The paragraphs varied from 93 to 131 words in length and the sentences from 8 to 14 words. Eight questions were made up for each paragraph and two for each sentence. They were all of the form "Who did (predicate)?" Half involved predicates true of the character and half involved predicates similar to facts true of the character, but not something asserted of the character. The first type of question required the character's name, but the second type required the answer "Nobody." Examples of the question used for the material illustrated earlier (see ACT's Predictions, above) are as follows: Who frequents tag sales and auctions? (answer = Carol Norman); Who earns a good living from a used furniture and antique shop? (answer = Nobody); Who bought a season ticket to the New Haven Night Hawks? (answer = Henry Caputo).

It is hard to measure the exact difference in amount learned in the paragraph vs. the sentence conditions. In terms of length of original tests, the difference is 10:1. In terms of amount of information guaranteed by the question answering, the difference is 4:1. Presumably, the true ratio is somewhere in between. There were 20 location terms, nouns six to seven letters long, used in the location learning phase. Twelve of these location nouns were randomly selected for the no-fan condition and were randomly paired with 12 of the names (3 from each condition). The other eight location nouns were assigned to the fan condition and were randomly paired with two of the eight remaining names. Thus in the fan condition, each name occurred with two locations and each location with two names. The two fans were correlated to avoid the min effect (see Anderson, 1976). There were 16 facts (8 × 2) to be learned in the fan condition. There were 12 more facts (12 × 1) to be learned in the no-fan condition. The location-name pairings were rerandomized for each subject.

Procedure. The experiment was entirely run at a VT50 terminal controlled by a PDP-11/40 (see Proudfoot, 1978, for a description of the system). The experiments were divided into the following four phrases: prior learning, name familiarization, location learning, and name-location recognition.

Prior learning. The experiment began with the presentation of the five paragraphs and five sentences to subjects in random order. The sentences were presented for 30 sec and the paragraphs, for 60 sec. Following this initial study phase, there was a study-test phase in which the subject had to answer the 40 questions based on the paragraphs and 10 questions based on the sentences. The questions were presented in random order. Subjects typed their answers into the computer and received feedback. If they were wrong, the entire paragraph or sentence was presented again for study and subject could restudy it as long as they liked. Subjects had to answer every question correctly twice before they passed into the next phase of the experiment. This was implemented by a drop-out procedure, in which each question dropped out of the sequence after having been correctly answered once. When all questions were answered correctly once, the drop-out procedure was repeated a second time.

Name familiarization phase. After completion of the above study-test phase, subjects went into a name familiarization phase whose function was to equalize name encoding times for the paragraph, sentence, and name conditions. First the five names from the name condition were presented twice to a subject in random sequence at a 5-sec rate. Then the 15 names from the paragraph, sentence, and name conditions were pooled. Each name was presented twice, along with 30 foil names. The 30 foil names were created by randomly re-pairing the first and last names of studied names. The subject's task was to judge whether or not each name was one that had been studied. From this first pass through the material, an overall mean reaction time was obtained. In Experiment 1, a performance criterion was set at 80% of the mean reaction time in the first block. In Experiment 2, this criterion was reset to 85%, because so many subjects failed to achieve the criterion of Experiment 1. Subjects then went into a phase in which they were tested on the material in each of the three conditions in separate blocks until they achieved the performance criterion for each condition. Each of the five names appeared twice in each block, along with 10 foils. The subjects began by going through one block for each of the three conditions. If subjects did not meet the performance criterion for any condition in the first block, that condition had to be repeated later. If more than one condition needed to be repeated, the condition tested first was the one with the largest mean reaction time. Once a condition reached criterion, the names for that condition were not tested again. Blocks of trials were given to subjects until they reached criterion in all of the conditions or until 1 h had passed. Subjects who failed to reach criterion within 1 h were dropped from the experiment.

Location learning. Upon completion of the name recognition phase, subjects entered the location learning phase. First they were presented 28 sentences asserting the name-location relationships. These were individually presented at a 15-sec rate. Then the subjects passed into the test-study cycle, in which they were presented with a location and had to type back all names (one or two) that occurred with that location. A double drop-out procedure was instituted here, as in the paragraph and sentence learning. That is, subjects had to correctly recall each name before that location was dropped out of the test sequence. Then they had to do this a second time. For each condition (prior learning by fan), the number of errors subjects made in recalling the names was recorded.

Name-location recognition. After finishing the location learning phase, subjects went into the final phase, in which they had to recognize the name-location combinations. This was the phase of most interest. This involved 10 blocks of 56 trials. In each block, the 28 studied name-location pairs and 28 foils were presented in the order of first name, last name, and location. The foils were created by combining the names (first and last together) randomly with locations different from those studied. There were no "fake" names (i.e., re-pairings of first and last name) used in this phase. In each set of 28 foils, there were two tokens of each name or location from the fan conditions. This meant that the names and locations occurred with the same frequency in the foils as in the targets.

Subjects sat before the CRT screen with a left finger on the terminal d and a right finger on the terminal k. They were instructed to press the left key if their recognition decision was "no" and the right key if their recognition decision was "yes." Reaction time was measured from the appearance of the stimulus on the screen to the depression of the response key. Subjects were given feedback as to the correctness of their response. The feedback was displayed on the screen for 1 sec. The interval between disappearance of the feedback from the screen to the next test probe was 1 sec. Thus, the dissappearance of the probe from the screen served as a warning for the next trial. There was an enforced 30-sec rest break between blocks of trials. Subjects could extend this break to however long they desired.

Results

Table 1 presents the results from the name familiarization phases of the two experiments. Part a presents the mean reaction times on the first block of name recognition times, Part b presents the mean number of blocks to criterion in each of the groups, and Part c presents the mean reaction times and error rates on the last block for each condition. In Part a, we see some evidence for an initial advantage of the more familiar conditions. The standard error of the means over the two experiments was 29 msec. Therefore, the difference between the paragraph and name condition was quite significant [$t(98) = 2.82$, $p < .01$] and was significant at the .05 level by more conservative statistical procedures such as Tukey's honestly significant differences (HSD) (Winer, 1971), but neither condition was significantly different from the sentence condition. While there may appear to be large differences among the conditions in Part b of Table 1, there were no significant differences within experiments in number of trials to criterion. The standard error of the mean was .95 for Experiment 1 and .9 for Experiment 2. Combining the two experiments, there was a significant difference between the sentence and the name conditions [$t(88) = 2.05$, $p < .01$], but one has to be suspicious of this unexpected difference, which was not significant with more conservative statistical techniques (e.g., Tukey's HSD). With respect to Part c of Table 1, the standard error of the mean reaction times (based on the Condition by Subject interaction) was 28 msec, with 42 degrees of freedom for Experiment 1 and 27 msec with 46 degrees of freedom for Experiment 2. There was little difference among the three conditions in either experiment. We can thus conclude that the name recognition phase accomplished its assigned task of equalizing encoding times.

Location learning. Due to an error in the program that collected data for Experiment 1, we failed to gather systematic data about the location learning. Therefore, we have the learning data to report only for Experiment 2. Table 2 presents mean number of errors made before reaching criterion of two correct recalls. The table is organized according to fan of the material to be learned and prior experience with the name. The standard error of the means in Table 2 was .25 based on the Condition by Subjects interaction with 161 degrees of freedom. There was clearly no significant difference between the paragraph and sentence condition. There was a significant difference between the mean of the sentence and paragraph conditions and the mean of the name condition [$t(161) = 2.08$, $p < .05$]. This was predicted by the ACT elaboration hypothesis, which claimed that the prior knowledge in the paragraph and sentence conditions would enable more redundant paths to be created to encode the location information. The unfamiliar condition was significantly worse [$t(161) = 3.60$, $p < .05$] than the name condition. This may reflect nothing more than the fact that subjects had to learn the names as well as the locations in this condition.

There was the predictable interference effect in location learning: Subjects made more errors in the high-fan conditions [$t(161) = 3.00$, $p < .05$]. An unexpected outcome was the marginally significant [$F(3,161) = 1.82$, $p > .10$] interaction between fan and amount of prior experience. Basically, the difference among the conditions of prior experience was smaller in the fan condition. This can probably be explained in terms of the recall test used. The prompts for recall were the location. In the fan condition, subjects had to be able to recall both names to the location to be able to count as answering that item correctly. An item had to be answered correctly twice before criterion was reached.

Table 2
Mean Number of Errors Made in Reaching Criterion in the Location Learning Phase

	Condition				
	Paragraph	Sentence	Name	Unfamiliar	Mean
Fan	1.71	1.59	1.79	2.54	1.91
No Fan	.67	.80	1.51	2.55	1.38
Mean	1.19	1.20	1.65	2.55	1.65

Table 1
Results for Name Recognition Phase

	(a) Mean RT During First Block			(b) Blocks to Criterion				(c) Performance on Last Block					
								P		S		N	
	P	S	N	P	S	N		RT	A	RT	A	RT	A
Experiment 1	1138	1177	1238	4.36	7.36	5.73	Targets	904	.99	952	.97	944	.98
							Foils	1120	.95	1150	.91	1174	.90
Experiment 2	1168	1264	1300	5.00	5.39	3.26	Targets	953	.97	904	.97	965	.98
							Foils	1141	.94	1190	.93	1162	.92
Mean	1153	1221	1269	4.68	6.38	4.50		1030	.96	1049	.95	1061	.95

Note—P = paragraph, S = sentence, N = name; RT = reaction time (in milliseconds), A = accuracy.

The two names associated to a location usually were not from the same prior knowledge condition. This meant that an error on an unfamiliar name could delay reaching criterion for a paragraph name. The errors counted in Table 2 were only when the name in that condition was incorrectly recalled. However, there are three ways that an item in the paragraph condition could be slowed down by being paired with an unfamiliar item. First, to the extent that subjects refused to recall either name if they could not recall both names, an inability to recall an unfamiliar name would affect the recall of the paragraph name. Second, if the subject did make an error on the unfamiliar name, he was tested again with that location and had another opportunity to make an error on the paragraph name. Third, subjects may have taken study effort away from the easy paragraph item to focus on its difficult unfamiliar partner.

The upshot of these various observations is that the learning data were somewhat contaminated in the fan condition. The no-fan condition lacked these problems because subjects had only to recall a single name to a location prompt. With the no-fan material, there is quite clear evidence for an effect of prior knowledge.

Location recognition. Table 3 reports the data from the location recognition phase for both experiments. The reaction times reported there are based only on correct trials. The standard error of these reaction times, based on the Subject by Condition interaction, was 35 msec for Experiment 1 with 315 degrees of freedom and 37 msec for Experiment 2 with 345 degrees of freedom. The standard error of the error rates, based on Subject by Condition interactions, was 1.2% for Experiment 1 and 1.2% for Experiment 2. Either error rate effects were in the same direction as the reaction times or there was no effect at all. The correlation between error rate and reaction times was .91 for Experiment 1 and .86 for Experiment 2. I will therefore focus my discussion on the reaction time effects.

In both experiments, the effects of all three variables (fan, degree of prior knowledge, and target vs. foil) were significant at the .05 level. Of particular relevance is that there was a significant effect of prior experience. The ACT theory (Anderson, 1976) expects targets and foils to show the same effects for reaction times, except that foils should take longer. Since there were no significant interactions involving the target-foil variable, I will average over this in further discussion. Averaging the two experiments together, the means were 1,238 msec for the paragraph condition, 1,229 msec for the sentence condition, 1,188 msec for the name condition, and 1,192 msec for the unfamiliar condition. The mean of the paragraph and sentence conditions was significantly different [$t(660) = 2.42$, $p < .01$] from the mean of the name and unfamiliar conditions. There was no other significant effect of amount of prior experience [$F(2,660) = 1.48$]. This result does confirm the ACT prediction of longer reaction time with more prior experience. It is somewhat perplexing, however, why there was not a larger difference between the paragraph and sentence conditions.

Table 3 reports how the fan effect varied as a function of conditions of prior experience. The standard error of the differences between fan and no fan was 35 msec for Experiment 1 and 37 msec for Experiment 2. There was an overall significant variation in these fan effects, although neither experiment displayed a significant interaction by itself. Averaging the two experiments together, the mean fan effect was 163 msec for the paragraph condition, 154 msec for the sentence condition, 190 msec for the name condition, and 235 msec for the unfamiliar condition. The mean of the paragraph and sentence conditions was significantly [$t(660) = 2.11$, $p < .05$] lower than the mean of the name and unfamiliar condition. The remaining variation not accounted for by this contrast was not significant [$F(2,66) = 2.10$]. This result confirms ACT's prediction

Table 3
Reaction Times (RT) in Milliseconds and Error Rates (ER) from Experiments 1 and 2

		Paragraph		Sentence		Name		Unfamiliar		Mean	
		RT	ER	RT	ER	RT	ER	RT	ER	RT	ER
						Experiment 1					
Targets	Fan	1269	.158	1254	.154	1243	.160	1235	.173	1250	.161
	No Fan	1105	.089	1102	.121	1056	.080	987	.067	1063	.089
Foils	Fan	1356	.164	1324	.167	1333	.166	1325	.162	1335	.165
	No Fan	1184	.118	1196	.117	1154	.121	1119	.133	1163	.122
Mean		1229	.132	1219	.140	1197	.132	1166	.134	1203	.134
Fan Effect		168	.058	140	.042	183	.063	227	.068	180	.058
						Experiment 2					
Targets	Fan	1261	.113	1260	.122	1236	.079	1290	.116	1262	.107
	No Fan	1155	.096	1134	.087	1016	.070	1052	.072	1089	.081
Foils	Fan	1390	.145	1388	.165	1317	.139	1394	.189	1372	.160
	No Fan	1180	.112	1178	.107	1143	.116	1142	.087	1161	.105
Mean		1247	.116	1240	.120	1178	.101	1220	.115	1221	.113
Fan Effect		158	.025	168	.047	197	.016	245	.073	142	.040

of a greater fan effect with less experience. However, it is surprising that there was not a significant difference between the paragraph and sentence conditions in size of the fan effect (nor was there in mean reaction time).

It is worth noting that there was a negative relationship between mean reaction times and the size of the fan effect. This is the opposite of the more frequent result of longer effects of an experimental manipulation for larger reaction times.

Discussion

It is informative to apply the ACT analysis, summarized by Equation 1 from the introduction, to these data. Table 4 provides a summary of the data averaged over the two experiments for this purpose. There, we have collapsed the paragraph and sentence conditions together as conditions of prior knowledge and the name and unfamiliar conditions together as conditions of no prior knowledge. Also, target and foil reaction times are collapsed. Equation 1 can be fit to these data, assuming the prior knowledge, K, in the name and unfamiliar conditions is 0 and letting K be a parameter to be estimated for the prior knowledge condition (sentence and paragraph). Also, s can be set to 1 to establish the scale for prior knowledge. Under these assumptions, we come up with the parameter estimates: a (intercept) = 876 msec, b (time to activate a single, uninterfered proposition) = 207 msec, K (strength of prior knowledge) = .75, f (amount prior knowledge is integrated with to-be-learned material) = .40.

With four parameters, four data points, and the data satisfying the qualitative predictions of the ACT theory, the theory also gives a perfect quantitative fit to the data. So the only thing of interest is the parameter values. The value of f = .40 implies that 40% of the activation that goes to the prior knowledge is spread into the target information. This seems the right order of magnitude for integration with prior knowledge. The value of K = .75 for strength of prior knowledge is surprisingly low. This implies that the strength of all the prior knowledge was less than the strength of one target proposition. In the paragraph condition, many propositions were well trained about the individual. Of course, there was little difference between the paragraph and sentence conditions. The small value of K explains why the effects of prior knowledge, while significant, were quite weak. Other research (e.g., Lewis & Anderson, 1976) has also produced rather small values for estimated strength of prior knowledge.

Table 4
Main Results (in Milliseconds) From Experiments 1 and 2

	Prior Knowledge: Paragraph and Sentence	No Prior Knowledge: Name and Unfamiliar
Fan	1313	1290
No Fan	1154	1083

Note–Each cell mean has a standard error of 12 msec.

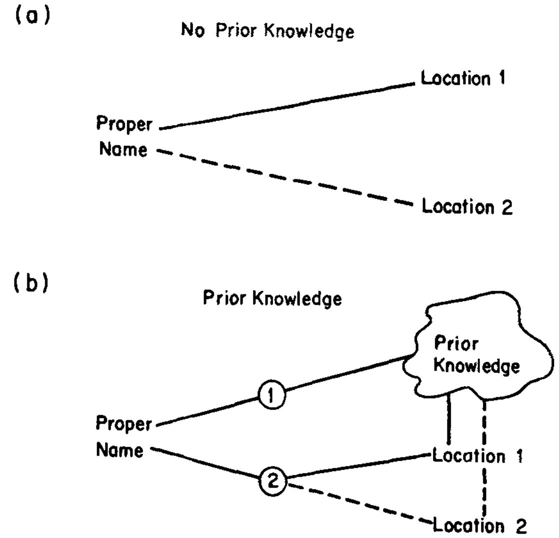

Figure 2. Schematic network representation of knowledge structure with no prior knowledge (Part a) and with prior knowledge (Part b). Subnodes 1 and 2 in Part b represent subnodes that help segregate prior knowledge from experimental knowledge.

It has been suggested (e.g., Anderson, 1976; Anderson & Paulson, 1978; Reder & Anderson, in press) that subjects have some success at filtering out the interfering effects of prior knowledge by use of a subnode structure. This involves creating subnodes of the individual node. Verifying a fact about the individual then involves two stages: selecting the correct subnode attached to the individual and then searching for the target fact attached to that subnode. The interference from prior knowledge is reduced to the relatively simple subnode search, which is not affected by the amount of prior knowledge attached to the subnode. This would explain why there is no difference between the sentence and paragraph conditions. Figures 2a and 2b contrast the knowledge representations under the subnode assumption. Figure 2a for the situation of no prior knowledge is identical to Figure 1a. However, in Figure 2b, there are two subnodes introduced: Subnode 1 collects the prior knowledge and Subnode 2 collects the location facts. We have links between these two sets of facts indicating the potential for integration. In terms of spreading activation, there are two waves of activation. First, activation spreads from the proper name. When Subnode 2 becomes sufficiently activated, it is identified as the one relevant to the current phase of the experiment and it becomes a focus of activation.

Equation 1 given for Figure 1 is not quite accurate as a description of Figure 2b. Here, there is first a process of subnode identification that takes time c. Then there is the process of activation spreading from the subnode. A quantity 1/n (n is fan) will be spread from Subnode 2 to the location, but also, some quantity f will arrive at the location from the prior knowledge (activated from the proper name). Thus the total activation will be proportional to $1/n + f$, and the predicted recognition time for Figure 2b becomes

$$RT = a + c + [bn/(1 + nf)], \quad (2)$$

whereas the recognition time for Figure 2a is simply

$$RT = a + nb. \quad (3)$$

This was fit to the data in Table 4.[2] Again, there are enough parameters to perfectly fit the data (again, given the data are within certain bounds) and the real interest is in the parameter values: c (time to reach subnode) = 89 msec, a (intercept) = 876 msec, b (time to activate a propositions with activation 1.0) = 207 msec, and f (amount of activation spread from prior knowledge) = .094. The low value of c confirms the idea that subnode access is quick. The value of f implied that about 10% of the activation of the prior knowledge subnode is spread through the elaborative connections to a location.

EXPERIMENT 3

The above two experiments were designed to test the ACT predictions about the effects of prior knowledge in a situation in which that prior knowledge is experimentally provided. The experiments generally supported the ACT analysis. This still leaves us with the issue of the apparent advantage of well-known names like Ted Kennedy for which subjects that have a great deal of preexperimental prior knowledge. I suggested four possible explanations in the introduction. The first three are ways of explaining the result without doing violence to the basic assumptions of the ACT theory. The fourth challenges a basic assumption of the theory: that there is a constant capacity of activation for all nodes. The third experiment was an attempt to create a situation that would eliminate the three possible "acceptable" explanations for the advantage of Ted Kennedy facts: differences in grapheme encoding, differential strengths of backpaths, and high degree of overlearning. If well-known names still show their reaction time advantage in this situation, this will be evidence against the ACT analysis and evidence that well-known names enjoy special capacity advantages.

In this experiment, we contrasted the processing of new information about familiar individuals like Ted Kennedy and Jimmy Carter with the processing of comparable information about made-up individuals like Ted Carter or Jimmy Kennedy. We took the first two experiments as providing evidence that with some minimum exposure to these new names, there would not be any differences in the encoding of these names. However, as a protection against a failure of this critical assumption, in the location recognition phase of this experiment, we first presented the name and asked a subject to press a button when he recognized the name. This presumably should guarantee that the name had been encoded in all conditions. Then we presented the person-location pair for judgment. In this way, we hoped to be able to eliminate effects of encoding time, should there be any.

The four conditions in Experiment 3 were paragraph, sentence, name only, and well-known. The ACT predictions for the well-known condition were that it would contrast with the paragraph and sentence conditions in the same way as the last two had contrasted with the name and unfamiliar conditions (i.e., more rapid learning, slower recognition times, and a smaller fan effect). This is assuming that the well-known names would provide a more extreme version of the prior knowledge representation in Figure 1b. On the other hand, if subjects were using a subnode representation like that in Figure 2b, then we might expect no difference between the well-known condition and the paragraph and location conditions with respect to recognition times, just as there has been no difference between the sentence and paragraph conditions.

Method

Subjects. Twenty-eight subjects were recruited from the general subject pool at Carnegie-Mellon University. They were paid $9 for an experiment that lasted less than 3 h. One subject was eliminated because of a very high error rate in the name-location recognition phase.

Materials. The names were derived from the 20 names of well-known individuals in the Appendix. Five of these names were randomly chosen for each subject to be in the well-known condition. The first and last names of the remaining 15 individuals were randomly permuted to create 15 unknown names. Five of the 15 novel names were assigned to the paragraph condition, 5 to the sentence condition, and 5 to the name-only condition. The materials for the paragraph and sentence conditions were essentially the same as those in the previous experiments, but they were slightly edited to introduce references meaningful in Pittsburgh rather than in New Haven.

Procedures. The procedure was very similar to the previous two experiments. Similar programs administered the experiment on Beehive terminals connected to a PDP-11/34 at Carnegie-Mellon University. The prior knowledge phase for the paragraph and sentence material was identical to those in the previous experiments. Rather than the extensive name familiarization phase of the previous experiments, subjects were simply presented at a 15-sec rate the five names in the name-only condition and the five names in the well-known condition. They were given one pass through the name recognition phase for all 20 names, but there was no attempt to bring recognition times to a uniformly rapid rate. This was just to establish some exposure to each name. We were depending on the encoding step in the name-location recognition to eliminate differences in encoding time.

The location learning phase was identical to that in the prior experiments. The name-location recognition phase involved an extra encoding step. The name from a probe appeared on the screen. The subject pressed his finger when he felt he had recognized the name and was ready to see the predicate. Then the full probe appeared on the screen, and the subject judged whether it came from a studied sentence. The time that the word alone appeared on the screen was controlled to never be less than 1 sec or more than 2 sec. If the subject pressed the button before 1 sec, the name-location pair did not appear on the screen until this minimum time had passed. This was to prevent subjects' simply trying to skip over the encoding phase to get on with the experiment. If the subject took longer than the maximum of 2 sec, the name-location pair was presented without waiting for the subject to respond. This was to prevent strategies by which the subject might recall into memory and rehearse the locations. With the addition of this name encoding operation, the person-location trials proceeded identically to the past two experiments. The majority of the subject responses were actually under 1 sec, suggesting the subject was able to encode the word in less than 1 sec.

Results

Table 5 presents the number of errors made before achieving the location learning criterion of two perfect recitals. The standard error of these numbers (based on the Subject by Condition interaction) was .17. The results basically replicate the previous experiments and confirm the ACT predictions with respect to the position of the well-known material relative to the other conditions. The mean errors to criterion decreased with amount of prior knowledge, with mean errors in the well-known condition significantly [$t(182) = 2.89, p < .005$] less than the average of the paragraph and sentence conditions, which was significantly less [$t(182) = 3.97, p < .001$] than the mean errors in the name-only condition. Again, these effects were somewhat attenuated in the fan condition, but, as was explained earlier, this had a somewhat uninteresting explanation.

The recognition times and error rates for the name-location recognition phase are displayed in Table 6. The standard error of the error rates was 1.7%, and for the reaction time means, it was 24 msec. Again, the reaction times and error rates were highly correlated ($r = .902$). However, as we will see shortly, there was one place in which reaction times and error rates point to different conclusions. With respect to the conditions in common with the previous experiment, we obtained a replication. Mean time was faster in the name-only condition than the average of the paragraph and sentence conditions (920 msec vs. 976 msec), and the fan effect was larger (245 msec vs. 193 msec). The difference in mean recognition times was significant [$t(390) = 3.81, p < .001$], but the difference in size of the fan effect did not reach significance [$t(390) = 1.05$] in this experiment. The well-known condition proved to be the opposite of ACT predictions. Rather than have the longest times, it had the fastest by far (823 msec), and rather than have the smallest fan effect, it had an effect (233 msec) that was almost as large as that of the name-only condition. It is worth noting here that the effect of fan on error rates was marginally significantly less [$t(390) = 1.68, p < .1$, two-tailed] for the well-known condition (5.2%) than it was for the other conditions (an average of 8.5%).

Table 7 presents the mean time subjects took to encode the names in the name-location recognition phase. There was very little variation among conditions. However, subjects were slightly faster in the fan condition, reflecting the fact that they studied these names more frequently. They were slightly faster on the well-known names than on the others. However, no effect approached significance. These results confirm, as suggested from the previous experiment, that encoding times do not play an important role in recognition of the name-location facts.

Table 5
Mean Number of Errors Made in Reaching Criterion in the Location Learning Phase

	Condition				
	Well-Known	Paragraph	Sentence	Name Only	Mean
No Fan	.31	.62	.91	1.46	.83
Fan	.88	1.40	1.12	1.74	1.29
Mean	.59	1.01	1.02	1.60	1.06

Table 7
Name Encoding Times (in Milliseconds)

	Well-Known	Paragraph	Sentence	Name Only
Fan	1065	1084	1077	1071
No Fan	1080	1083	1087	1090
Mean	1073	1084	1082	1080

GENERAL DISCUSSION

Experiment 3 indicates that the fast retrieval of facts about well-known individuals is not due to the high degree of overlearning for these facts. In this experiment, we tried to control the degree of learning of facts about different individuals. Indeed, it might

Table 6
Reaction Times (RT) in Milliseconds and Error Rates (ER) for Experiment 3

		Well-Known		Paragraph		Sentence		Name Only	
		RT	ER	RT	ER	RT	ER	RT	ER
Targets	Fan	871	.117	1033	.202	1013	.180	999	.209
	No Fan	678	.098	879	.102	870	.116	780	.110
Foils	Fan	1007	.168	1122	.208	1123	.199	1086	.189
	No Fan	734	.084	875	.109	893	.111	815	.130
Mean		823	.117	977	.155	975	.152	920	.160
Fan Effects		233	.052	200	.100	186	.076	245	.079

be argued that the drop-out procedure gives the advantage of extra learning to the facts about the unknown names, because these names required more trials to criterion. Thus, Explanation 2 (greater strength of backpath) and Explanation 3 (degree of overlearning) in the introduction are ruled out as explanations of why facts about well-known individuals are retrieved rapidly. The use of an encoding phase also makes Explanation 1 (encoding differences) unlikely. The greater speed subjects have in recalling facts about well-known names seems to derive from properties of the name itself, not from properties of the facts attached to that name (Explanation 4 in the introduction).

Experiments 1 and 2 established and Experiment 3 confirmed that experimentally acquired information leads to a rather complex pattern of effects: We see faster learning, slower verification, and smaller fan effects. These results were predicted from simulations with the ACT theory and from our approximate mathematical analysis of the theory. There was not a significant difference between sentence and paragraph conditions, but this was attributed to use of a subnode model (Figure 2). Thus, when we look at experimental prior knowledge, we find benefits (faster learning, smaller fan effect) but also costs (slower verification times). However, when we look at names with high degrees of preexperimental prior knowledge, the striking result is that there appear to be only benefits of prior knowledge. In particular, verification times are least for these names. The fan effect is slightly ambiguous. The size of the effect on reaction time was slightly larger than the mean for the other conditions. On the other hand, the size of the effect for error rates was smaller than the mean of the other conditions. Thus, because of considerations of speed-accuracy tradeoff, one cannot be sure what to judge about the relative size of the fan effect in the well-known condition.

It is clear that these well-known names are not on a simple continuum with our experimental manipulation of prior knowledge. In one way or another, they have greater capacity than do the names we created. This fact raises many questions that we cannot answer with any assurance. For instance, how are the well-known names different? In the framework of the ACT theory, one might assume that the well-known names have greater capacity for spreading activation, but this speculation surely needs converging evidence. Another question is what type of experimental manipulation would convey upon experimental material the power of the well-known names. Is it just a matter of enough exposure? Or is the richness of experience and the multimodality exposure necessary? These are intriguing questions, and someday we may be able to answer them. However, for now, it is sufficient to know that the beneficial effects of well-known names cannot be simply accounted for in terms of the redundancy of their network connections. In knowing this, we know that well-known names have properties different from those produced by prior knowledge learned in 1 h time in a laboratory experiment.

REFERENCES

Anderson, J. R. *Language, memory, and thought.* Hillsdale, N.J: Erlbaum, 1976.

Anderson, J. R., & Paulson, R. Interference in memory for pictorial information. *Cognitive Psychology,* 1978, **10**, 178-202.

Anderson, J. R., & Reder, L. M. An elaborative processing explanation of depth of processing. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory.* Hillsdale, N.J: Erlbaum, 1979.

Hayes-Roth, B. Evolution of cognitive structures and processes. *Psychological Review,* 1977, **84**, 260-278.

Lewis, C. H., & Anderson, J. R. Interference with real world knowledge. *Cognitive Psychology*, 1976, **7**, 311-335.

Moeser, S. D. The role of experimental design in investigations of the fan effect. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, **5**, 125-134.

Proudfoot, R. G. YEPS: Running real-time experiments on a timesharing system. *Behavior Research Methods & Instrumentation*, 1978, **10**, 291-296.

Reder, L. M., & Anderson, J. R. A partial resolution of the paradox of interference: The role of integrating knowledge. *Cognitive Psychology,* 1980, **12**, 447-472.

Smith, E. E., Adams, N., & Schorr, D. Fact retrieval and the paradox of interference. *Cognitive Psychology,* 1978, **10**, 438-464.

Winer, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1971.

NOTES

1. We did try an existing computer simulation that we have of the spreading activation process. The following mathematical analysis is an attempt to capture the highlights of that simulation. Simulating the exact network in Figure 1, we obtained the following time measures for the location structures to be retrieved: no fan, no prior knowledge–1 time unit; fan, no prior knowledge–2 time units; no fan, prior knowledge–1.96 time units; fan, prior knowledge–2.63 time units. As the mathematical analysis indicates, there are main effects of prior knowledge and of fan and an interaction between these two factors.

2. We ran our simulation (see also Footnote 1) on these knowledge representations to confirm the accuracy of the equations. We assumed the structure to be identical to Figure 1 with the addition of subnodes. The obtained values for the simulation were: prior knowledge, fan–2.28 time units; prior knowledge, no fan–1.88 time units; no prior knowledge, fan–2.00 time units; no prior knowledge, no fan–1.00 time unit (assigning c = .80 time units).

Appendix
Famous Names Used in Experiment 3

Barbara Walters	Helen Keller
Shirley Temple	Judy Collins
John Lindsay	Jesse James
Gerald Ford	Richard Burton
Grace Kelley	Dean Martin
Paul Newman	Andrew Young
Benjamin Franklin	Elizabeth Taylor
Carole King	Jerry Brown
Robert Kennedy	George Wallace
Henry Jackson	Isaac Newton

(Received for publication August 7, 1980; revision accepted October 31, 1980.)