

Focused memory search in fact retrieval

MICHAEL McCLOSKEY and KURT BIGLER
Johns Hopkins University, Baltimore, Maryland 21218

Several studies of fact retrieval have shown that the more facts a person learns about a concept, the longer it takes him or her to retrieve any of these facts. This result has been interpreted to mean that retrieval of a fact about a concept involves a search of all facts stored in memory with that concept. In the present study, it is suggested that retrieval involves not an unfocused search of all facts stored with a concept, but rather a focused memory search that examines relevant stored facts and ignores irrelevant information. This argument is supported by three experiments in which subjects first learned simple facts (e.g., "The banker likes horses") and then made speeded true-false decisions for test probes (e.g., "The banker likes elephants"). Specifically, results suggest that facts stored with a concept may be organized into subsets. For example, a person's knowledge about Richard Nixon might be organized into subsets concerning Nixon's resignation, his trips to China, his family, and so on. The data further suggest that a person attempting to retrieve a fact about a concept (e.g., the name of Nixon's wife) may simply decide which subset is most likely to contain the desired fact (e.g., the subset concerning Nixon's family) and search that subset. If the sought-for fact is found in this subset, the search process terminates. If, however, the desired information is not located, other subsets of facts may be searched before the retrieval attempt is given up. The notion that memory search focuses on relevant stored facts and ignores irrelevant information may help to explain why experts (i.e., people who know a large number of facts about a topic) do not experience great difficulty in retrieving facts in their areas of expertise.

In fact-retrieval experiments (e.g., Anderson, 1974; Thorndyke & Bower, 1974), subjects are typically asked to study a set of simple facts (e.g., "The banker touched the prisoner," "The student pushed the gambler") until they can recall the facts perfectly. Test probes (e.g., "The banker hit the priest") are then presented, and the subject's task is to respond "true" if the probe matches one of the learned facts or "false" if it does not. The dependent variable of interest is reaction time (RT) to make the true-false decisions.

Several fact-retrieval studies have shown that the more facts a person learns about a concept, the longer it takes him or her to respond to probes involving that concept (e.g., Anderson, 1974; Anderson & Bower, 1973; Thorndyke & Bower, 1974). For example, people take longer to make a true-false decision for a probe like "The banker hit the priest" when they have learned three facts about the banker than when they have learned only one fact concerning the banker. This result is referred to as the fan effect (e.g., Anderson, 1976).

The fan effect may be interpreted by assuming that a person attempting to retrieve a fact about a concept searches through all of the facts stored in memory with

that concept (e.g., Anderson, 1974, 1976; Anderson & Bower, 1973; Thorndyke & Bower, 1974). For example, a person trying to remember whether or not the banker hit the priest would search through information stored with the concepts "banker," "hit," and "priest." The greater the number of facts stored with these concepts, the longer the search process should take to retrieve the fact "The banker hit the priest" or to determine that this fact is not among those stored in memory.¹

As Smith, Adams, and Schorr (1978) have pointed out, this interpretation implies that a person who knows a large number of facts about a concept (i.e., an expert) should be very slow in retrieving any of these facts. However, informal observation suggests to the contrary that experts on a particular topic are able to retrieve facts concerning that topic at least as quickly as nonexperts.

Smith et al. (1978) have suggested that the expert may have less difficulty than expected in retrieving facts because his or her knowledge is "integrated" in some way. In support of this view, Smith et al. (1978; see also Moeser, 1979) have shown that the magnitude of the fan effect is reduced if an integrating theme unites the facts learned about a concept. How integrating functions to reduce the fan effect is not, however, entirely clear.

In the present study, we take a different approach to the question of how experts manage to retrieve facts without apparent difficulty. Specifically, we suggest that retrieval of a fact about a concept usually involves

This research was supported by Public Health Service Research Grant MH23401, Sam Glucksberg, principal investigator, and by Biomedical Research Support Grant 5507RR07041-12, Michael McCloskey, principal investigator. We would like to thank Susan Anderson and Annette Bicker for testing subjects. We would also like to thank Alfonso Caramazza, Howard Egeth, and Jeffrey Santee for their helpful comments.

not an unfocused search of all facts stored with the concept, but rather a focused search that examines relevant stored facts and ignores irrelevant facts. For example, a person attempting to retrieve the name of Richard Nixon's wife may not search through all stored facts about Richard Nixon. Rather, he or she may focus the memory search on facts concerning Nixon's family and ignore facts about Nixon's foreign policy views, physical appearance, and so forth. If memory search processes can focus on relevant facts and ignore irrelevant information, then knowing a large number of facts about a concept need not lead to very slow retrieval.

The focused memory search notion is not a new one. For example, Oldfield (1966) suggested that retrieval of lexical information (e.g., names of objects) is accomplished by a focused search of the lexicon. In addition, Anderson and Paulson (1978) have employed the idea of focused memory search to interpret performance in a task in which subjects learned verbal and pictorial facts about concepts. Anderson and Paulson argued that subjects evaluating test probes focused their memory search (to some extent) on stored verbal facts when the probe involved verbal information and on stored pictorial facts when the probe involved pictorial information. Finally, focused memory search processes have been proposed to explain performance in a number of studies involving scanning of short memorized lists of words, letters, or digits (e.g., Crain & DeRosa, 1974; Naus, Glucksberg, & Ornstein, 1972; see Wescourt & Atkinson, 1976, for a review).

In Experiment 1 of the present study we demonstrate that people attempting to retrieve facts from long-term memory are able to focus on relevant information and to ignore, at least partially, irrelevant information. Experiments 2 and 3 seek to establish how the focusing of the memory search is accomplished.

EXPERIMENT 1

Subjects in this experiment first learned facts about several imaginary people who were referred to by occupation (e.g., "the tailor"). All of the facts were of the form "The (occupation) likes (object)" (e.g., "The tailor likes horses"). The object terms were exemplars taken from two different semantic categories, such as animals and countries (e.g., "The tailor likes horses," "The tailor likes Spain"). These two categories will be referred to as Category A and Category B.

For some of the occupation names (e.g., "architect"), subjects learned one fact involving an exemplar from Category A—for example, "The architect likes elephants." These occupations will be referred to as 1A occupations.

For other occupations, the 6A occupations, subjects learned six facts concerning Category A exemplars—for example, "The lawyer likes wolves," "The lawyer likes rabbits," "The lawyer likes bears," "The lawyer likes tigers," "The lawyer likes pigs," and "The lawyer likes dogs."

Finally, for the 1A-5B occupations, subjects learned one fact about a Category A exemplar and five facts concerning Category B exemplars—for example, "The editor likes lions" (A), "The editor likes Portugal" (B), "The editor likes Italy" (B), "The editor likes Canada" (B), "The editor likes England" (B), and "The editor likes Brazil" (B).

After learning the facts, subjects made true-false decisions for test probes (e.g., "The editor likes bears"). The probes were designed to answer the following question: Do subjects evaluating a probe simply search through all of the facts stored with the occupation in the probe, or do they instead focus on relevant stored facts and ignore irrelevant facts? For example, do subjects evaluating the probe "The editor likes bears" simply search through the six facts learned about the editor, or do they instead focus on the one relevant fact (i.e., "The editor likes lions") and ignore the five irrelevant facts (i.e., "The editor likes Portugal," "The editor likes Italy," etc.)?²

Four types of probes, which are described in Table 1, were employed to answer this question. Each probe type may be designated by two digits (e.g., 1-5) that represent, respectively, the number of relevant facts and the number of irrelevant facts stored with the occupation in the probe. For example, a 1-5 probe is one for which one fact stored with the probed occupation is relevant and five are irrelevant. The four probe types were the following: (1) *1-0 probes*, consisting of 1A occupations paired with Category A exemplars (e.g., "The architect likes wolves"). For this type of probe, the one Category A fact learned for the occupation (i.e., "The architect likes elephants") is relevant for purposes of deciding whether the probe is true or false. (2) *1-5 probes*, involving 1A-5B occupations and Category A exemplars (e.g., "The editor likes bears"). For these probes, the one Category A fact learned for the occupation (i.e., "The editor likes lions") is relevant, and the five Category B facts (i.e., "The editor likes Portugal," etc.) are irrelevant. (3) *5-1 probes*, made up of 1A-5B occupations and Category B exemplars (e.g., "The editor likes Mexico"). For these probes, the five Category B facts stored with the occupation are relevant, and the one Category A fact is irrelevant. (4) *6-0 probes*, consisting of 6A occupations and Category A exemplars (e.g., "The lawyer likes elephants"). For these probes, the

Table 1
Test Probes for Experiment 1

Probe Type	Occupation Type	Exemplar's Category	Example
1-0	1A	A	The architect likes wolves
1-5	1A-5B	A	The editor likes bears
5-1	1A-5B	B	The editor likes Mexico
6-0	6A	A	The lawyer likes elephants

Note—The two numbers designating the probe type refer to the number of relevant and irrelevant facts, respectively, stored with the probed occupation.

six Category A facts learned for the occupation are relevant. Within each of the four types, half of the probes were true and half were false.

At least two patterns of probe RTs may be obtained. First, if evaluation of a probe involves an unfocused search through the facts stored with the occupation in the probe, RT should be a function of the total number of facts learned about this occupation. Hence, RT should be the same for 1-5, 5-1, and 6-0 probes, because in all three cases the total number of facts learned about the probed occupation is six. In addition, RT should be faster for 1-0 probes than for the other three types, because only one fact was learned about the occupations in 1-0 probes.

In contrast, if subjects can focus on relevant facts and ignore irrelevant facts, RT should be a function of the number of relevant facts learned about the probed occupation. Thus, RT should be faster for 1-5 probes (i.e., those with one relevant and five irrelevant facts) than for 5-1 probes (five relevant facts and one irrelevant fact) or 6-0 probes (six relevant facts). If irrelevant facts are ignored completely, RT for 1-5 probes should be as fast as that for 1-0 probes. If, however, irrelevant facts are ignored only partially, RT for 1-5 probes should fall somewhere between RT for 1-0 probes and RT for 6-0 probes.

Method

Subjects. Subjects were eight undergraduate students at Princeton University.

Materials. Twenty occupation names and 80 category exemplars, 10 from each of eight categories, were used to generate a different set of stimulus materials for each of the eight subjects. Each set of materials utilized (1) 10 occupation names chosen randomly from the pool of 20 and (2) 20 category exemplars, 10 from each of two semantically unrelated categories (e.g., animals, countries). The two categories were designated Category A and Category B. Across the eight sets of stimulus materials, each of the eight categories was used twice, once as Category A and once as Category B.

The 10 occupation names and 20 category exemplars were used to construct 40 facts to be learned by the subject. The 40 facts (e.g., "The editor likes lions") were formed by pairing occupations and category exemplars in the following way: Four occupations were designated 1A occupations, and each was paired with one Category A exemplar (e.g., "The architect likes elephants," "The plumber likes monkeys"). Four additional occupations were designated 1A-5B occupations, and each was paired with one Category A exemplar and five Category B exemplars. Finally, the remaining two occupations were designated 6A occupations, and each was paired with six Category A exemplars. Each Category A and Category B exemplar was used in two different facts.

Sixty-four test probes, 32 true and 32 false, were also generated (see Table 1 for examples). The 64 probes were made up of: (1) 8 1-0 probes, 1 true and 1 false for each 1A occupation; (2) 8 1-5 probes, 1 true and 1 false for each 1A-5B occupation; (3) 32 5-1 probes, 4 true and 4 false for each 1A-5B occupation; (4) 16 6-0 probes, 4 true and 4 false for each 6A occupation.

True probes (e.g., "The architect likes elephants") were selected from among the facts to be learned by the subject. False probes (e.g., "The architect likes wolves") were formed by pairing each occupation with exemplars other than those with which it was paired in the set of to-be-learned facts. All 20 category exemplars were used in both true and false probes.

Lists of test probes were constructed by block randomizing the 64 probes. Each list was made up of two blocks of 32 probes, with each block containing half of the true probes and half of the false probes of each type (i.e., 1-0, 1-5, 5-1, 6-0). The ordering of probes within blocks was random. Three different block randomizations of the 64 probes were generated.

Apparatus. A Digital Equipment Corporation PDP-12 computer controlled presentation of probes, recorded responses, and measured RTs in milliseconds. Probes were displayed on a Tektronix Type 602 cathode-ray tube (CRT) unit.

Procedure. One or 2 days prior to the experimental session, each subject was given a list of the 40 facts to be learned. In this list, all facts for a given occupation were presented consecutively (e.g., the six facts for a 6A occupation were shown together). The subject was instructed to study the facts so that he or she could answer questions like "What countries does the editor like?" or "What animals does the lawyer like?"

At the beginning of the experimental session, the experimenter tested the subject's knowledge of the facts by asking questions of the form "What (category name) does the (occupation name) like?" (e.g., "What countries does the editor like?"). All 40 facts were tested in this way. If the subject made any errors, he or she studied the facts further and was then retested. This process continued until the subject could recall all 40 facts without hesitation.

The subject was then seated in a darkened, sound-attenuated booth facing a CRT screen. In front of the subject were two response keys, labeled "true" and "false," which were operated by the index fingers of the left and right hands, respectively. The subject was instructed that probes consisting of an occupation name and a category exemplar (e.g., "editor bears") would be shown on the screen and that these probes should be interpreted as sentences of the form "The (occupation) likes (category exemplar)" (e.g., "The editor likes bears"). The subject was told that his or her task was to decide whether each probe was true or false. A probe was to be considered true if it matched one of the facts the subject had learned and false if it did not. Instructions stressed that true-false decisions should be made as quickly and accurately as possible.

Each subject first performed a 96-trial true-false practice task in which he or she pressed the appropriate response key when the word "true" or the word "false" appeared on the CRT. Following this practice task, the test probes were presented. Each subject received all three randomizations of the 64 probes. Five-minute rest breaks occurred after each set of 64 trials.

The sequence of events on each trial was as follows: First, a fixation point was displayed at the center of the screen for 500 msec. The screen was then blank for 50 msec, after which the occupation and category exemplar were presented. The occupation was centered directly above the location at which the fixation point was displayed, and the category exemplar was centered directly below this location. The subject's "true" or "false" response terminated stimulus presentation and initiated a 1,650-msec intertrial interval.

Results and Discussion

Mean RTs and error rates are presented in Table 2.

Table 2
Mean RTs (in Milliseconds) and Error Rates (in Percentages)
for Test Probes in Experiment 1

Probe Type	True		False	
	RT	Error	RT	Error
1-0	1202	5	1382	4
1-5	1312	9	1620	3
5-1	1570	11	1801	12
6-0	1617	9	1876	13

Mean RTs were calculated from correct responses to the 192 test probes. The RT data were submitted to an analysis of variance with truth value (true, false) and probe type (1-0, 1-5, 5-1, 6-0) as repeated measures. Because each subject received a different set of stimulus materials, this analysis takes into account both subject and item variance (Clark, 1973).

A main effect of truth value was obtained [$F(1,7) = 28.56, p < .05$], indicating that RTs were faster for true probes than for false probes. The main effect of probe type was also reliable [$F(3,21) = 16.10, p < .05$]. The Truth Value by Probe Type interaction was not significant ($F < 1$).

T tests were used to evaluate differences between individual probe types. Because probe truth value did not interact with probe type, tests were made on data collapsed across true and false probes. Four tests were performed, with the significance level for each test set at .01 in order to maintain an overall level of approximately .05.

The first three tests involved the 1-5 (one relevant, five irrelevant facts), 5-1 (five relevant, one irrelevant facts), and 6-0 (six relevant, no irrelevant facts) probes. The purpose of these tests was to determine whether RT varied as a function of the number of relevant facts learned about a probed occupation, even when the total number of facts learned about the occupation was held constant at six. The tests revealed that RT was faster for 1-5 probes than for 5-1 probes [$t(7) = 3.81, p < .01$] or 6-0 probes [$t(7) = 3.60, p < .01$]. The error rate was also somewhat lower for 1-5 than for 5-1 or 6-0 probes. Although mean RT was slightly faster for 5-1 probes (1,686 msec) than for 6-0 probes (1,746 msec), this difference was not significant ($t < 1$).

The finding of faster RT for 1-5 than for 5-1 and 6-0 probes suggests that subjects evaluating probes did not search through all facts stored with the probed occupation. Instead, they focused on relevant stored facts and ignored, at least to some extent, irrelevant facts. The failure to find a significant difference between 5-1 and 6-0 probes may well have occurred because the two probe types differed only by one in the number of relevant facts.

The fourth t test compared 1-0 and 1-5 probes, in order to determine whether irrelevant facts had any effect on RT. While the number of relevant facts was the same for 1-0 and 1-5 probes, the number of irrelevant facts differed for the two probe types (none for the 1-0 probes, five for the 1-5 probes). The difference in RT between 1-0 probes (1,292 msec) and 1-5 probes (1,466 msec) was not reliable [$t(7) = 2.27, p > .05$]. Taken at face value, this result suggests that probe RT was not influenced by the number of irrelevant facts learned about the probed occupation. However, the 174-msec difference between 1-0 and 1-5 probes is large enough to suggest that irrelevant facts may have had an effect on RT that the experiment was not

powerful enough to detect. Thus, while the present findings indicate that people searching memory can focus on relevant stored facts, the data do not provide a clear answer to the question of whether irrelevant facts are ignored completely or only partially.

An important question raised by the results of Experiment 1 is this: How does the retrieval process focus on relevant facts and ignore irrelevant information? In the next section, we consider several models of focused memory search. Predictions derived from these models are then tested in Experiments 2 and 3.

Models of Focused Memory Search

In the following discussion, we will describe models of focused memory search in general terms, rather than developing these models within the framework of specific theories of memory, such as HAM (Anderson & Bower, 1973) or ACT (Anderson, 1976). We adopt this approach because many of the detailed assumptions of the individual memory theories are irrelevant for the models to be considered here. For example, for most of the focused search models, it is irrelevant whether retrieval of facts stored with a concept involves a serial self-terminating search (e.g., Anderson & Bower, 1973) or a parallel spread of activation (e.g., Anderson, 1976). Thus, in discussing focused search models, we will in general assume only (1) that facts stored with a concept may be accessed from that concept (e.g., facts known about Richard Nixon may be accessed by means of a search proceeding from the "Nixon" concept in memory) and (2) that a search of stored facts takes longer the greater the number of facts that must be searched.

We will discuss two classes of focused search models, which may be labeled "selective access" and "nonselective access" models. The selective access models assume that the facts stored with a concept are organized into subsets and that each subset of facts can be accessed selectively (i.e., without accessing other subsets). For example, a person's knowledge of Richard Nixon might be organized into subsets concerning Nixon's family, his trips to China, his resignation, and so forth, such that facts about, say, Nixon's resignation could be accessed independent of facts in other subsets.

A variety of models of the selective access class may be devised. For example, a simple selective access model might assume that retrieval of a fact about a concept (e.g., the name of Nixon's wife) is accomplished by a three-stage process that (1) identifies the relevant subset of information (e.g., the subset concerning Nixon's family), (2) accesses this relevant subset, and (3) searches the subset for the desired information (e.g., Nixon's wife's name).

In terms of the procedure of Experiment 1, this model would assume that the facts a subject learned about a concept (e.g., the editor) were organized into two subsets, one concerning animals (e.g., "The editor

likes lions") and one concerning countries (e.g., "The editor likes Portugal"). When a probe like "The editor likes wolves" was presented, the subject would (1) decide that information concerning the editor's preferences in animals is relevant, (2) access the subset of animal facts stored with the editor, and (3) search this subset to determine whether or not the editor likes wolves.

In this selective access model, then, only the facts in the relevant subset are searched; irrelevant facts (i.e., those in other subsets) are completely ignored. Hence, this model may be referred to as the Relevant Subset Only model.

The Relevant Subset Only model is similar to the Directed Entry model developed by Naus et al. (1972) in the context of the Sternberg (1966, 1969) short-term memory scan paradigm. The model is also similar to the subnode model proposed as part of Anderson's (1976) ACT theory (see also Anderson & Paulson, 1978).

In contrast to the selective access models, non-selective access models do not assume that facts stored with a concept are organized into subsets or that relevant facts may be accessed selectively. Instead, these models argue that both relevant and irrelevant information is accessed during an attempt to retrieve a fact. According to the nonselective access models, memory search focuses on relevant information in the sense that irrelevant facts undergo less complete processing than relevant facts.

We will describe here one simple nonselective access model. This model, which will be referred to as the Relevance Filter model, assumes that retrieval of a fact about a concept is a two-stage process. In the first, or relevance filter, stage, all facts stored with the concept are accessed and checked for potential relevance to the sought-for information. Facts found to be irrelevant are not considered further (i.e., are filtered out). The facts identified as potentially relevant, however, are passed to a second stage, in which they are searched for the desired information. For example, a person attempting to retrieve the name of Nixon's wife would initially access all facts stored with the "Nixon" concept. In the relevance filter stage, these facts would undergo a relevance check that might identify facts involving Nixon's family as potentially relevant and facts concerning other matters as irrelevant. In the second stage, the facts concerning Nixon's family would be searched for the name of Nixon's wife.

Similarly, a subject evaluating a probe like "The editor likes wolves" would first access and check for relevance all facts stored with the editor. Facts found to be irrelevant (e.g., "The editor likes Spain") would not be processed further, whereas facts identified as relevant (e.g., "The editor likes lions") would be searched to determine whether or not the editor likes wolves.

The Relevance Filter model would seem to be a

reasonable one, if we assume that checking for relevance is fast relative to searching for a particular fact. If this were the case, the time saved by not having to search irrelevant facts for a desired piece of information would usually more than compensate for the time needed to filter out the irrelevant facts.

In terms of the procedures of Experiment 1, both the Relevant Subset Only model and the Relevance Filter model (as well as other selective and nonselective access models) argue that when the total number of facts learned about a probed occupation is held constant, RT should be faster the smaller the number of *relevant* facts. Thus, the finding of faster RT for 1-5 probes than for 5-1 or 6-0 probes does not discriminate among the various focused search models. Fortunately, however, the models we have considered differ in their predictions concerning the effect on RT of the number of *irrelevant* facts stored with a probed occupation. Thus, in Experiments 2 and 3, we sought to assess the effects of irrelevant facts. (As we have seen, the results of Experiment 1 concerning irrelevant facts were equivocal.)

Nonselective access models assume that irrelevant as well as relevant facts are processed during fact retrieval. Thus, these models predict that probe RT will be influenced by the number of irrelevant facts stored with a probed occupation. For example, in the Relevance Filter model's first stage, all facts stored with the probed occupation are checked for relevance. Hence, the duration of this stage should be influenced both by the number of relevant facts and by the number of irrelevant facts. However, the duration of the second stage, in which relevant facts are searched, should be a function only of the number of relevant facts. Thus, probe RT should be influenced both by the number of relevant facts and the number of irrelevant facts, although the effect of the irrelevant facts should be smaller than that of the relevant facts.

The predictions of the Relevant Subset Only model are somewhat different. This selective access model assumes that when a probe is presented, the relevant subset of facts is identified, accessed, and searched. Irrelevant facts are never accessed; thus, the number of irrelevant facts stored with a probed occupation should not affect RT.

It should be noted that this prediction does not follow from all models of the selective access class; selective access models in which irrelevant facts are sometimes searched may readily be devised (see, for example, the Random Entry model of Naus et al., 1972). It should also be pointed out that even for the Relevant Subset Only model, we might expect RT to be faster for probes with no irrelevant facts than for probes with at least one irrelevant fact (other things being equal). For probes with at least one irrelevant fact, there are two or more subsets of facts stored with the probed occupation, whereas for probes with no irrelevant facts there is only a single set of stored facts. For example,

for occupations in 1-5 probes, there were two subsets of stored facts, one concerning animals and one concerning countries. For occupations in 1-0 probes, however, there was only a single set consisting of one fact about animals. According to the Relevant Subset Only model, before the relevant subset can be searched, this subset must be selected from among those stored with the probed occupation. This process of selecting the relevant subset may be simpler, and therefore faster, the smaller the number of subsets stored with the probed occupation. Thus, the relevant subset selection process (and, consequently, RT for the true-false decision) may be faster for probes with no irrelevant facts (i.e., those with only one set of stored facts) than for probes with at least one irrelevant fact (i.e., those with two or more subsets of facts).

However, as long as the number of subsets is held constant, the Relevant Subset Only model predicts that the number of facts in an irrelevant subset should not affect RT. Thus, in Experiment 2, we varied the number of irrelevant facts while holding constant the number of relevant facts and the number of subsets into which the facts would, according to the Relevant Subset Only model, be organized. In this situation, the Relevant Subset Only model predicts no effect of irrelevant facts, whereas the Relevance Filter model and other nonselective access models predict slower RT the greater the number of irrelevant facts.

EXPERIMENT 2

As in Experiment 1, subjects learned facts about imaginary people who were referred to by occupation (e.g., "the physicist"). Once again, the facts fell into two categories, A and B.

For some occupations, subjects learned two facts. For 2A and 2B occupations, the two facts were both from the same category (e.g., "The astronomer likes lions," "The astronomer likes giraffes"); for 1A-1B occupations, one fact was from Category A and the other was from Category B (e.g., "The lawyer likes goats," "The lawyer likes France").

For the remaining occupations, subjects learned six facts. These facts consisted either of six facts from a single category (6A and 6B occupations) or of one fact from one category and five from the other (1A-5B and 1B-5A occupations).

Five types of test probes—2-0, 1-1, 1-5, 5-1, and 6-0—were presented. As before, the two digits used to designate each probe type represent, respectively, the number of relevant and irrelevant facts stored with the occupation in the probe. The probe types are described in detail in Table 3.

The 1-5, 5-1, and 6-0 probes provide a replication of the important conditions of Experiment 1. If subjects evaluating probes focus on relevant stored facts, RT should be faster for 1-5 probes than for 5-1 or 6-0 probes.

Table 3
Test Probes for Experiment 2

Probe Type	Occupation Type	Exemplar's Category
2-0	2A	A
	2B	B
1-1	1A-1B	A
	1A-1B	B
1-5	1A-5B	A
	1B-5A	B
5-1	1A-5B	B
	1B-5A	A
6-0	6A	A
	6B	B

The comparison of primary interest, however, is between the 1-1 and 1-5 probes. While the number of relevant stored facts is one for both 1-1 and 1-5 probes, the two probe types differ in the number of irrelevant facts (one for 1-1 probes and five for 1-5 probes). If, as the Relevant Subset Only model assumes, irrelevant facts are never accessed during probe evaluation, RT should be the same for 1-1 and 1-5 probes. If, however, irrelevant facts undergo some processing during fact retrieval (as the Relevance Filter model and other nonselective access models assume), then RT should be slower for 1-5 probes than for 1-1 probes.

Method

Subjects. Subjects were 12 undergraduate students at the Johns Hopkins University.

Materials. A different set of materials was generated for each of the 12 subjects. Each set was constructed from (1) 12 occupation names chosen randomly from a pool of 60 and (2) 24 category exemplars, 12 from each of two semantically unrelated categories (designated Categories A and B). Across the 12 sets of materials, 12 different categories were used. Each category served as Category A in one set of materials and as Category B in another set.

Forty-eight to-be-learned facts were generated by pairing one 2A, one 2B, four 1A-1B, two 1A-5B, two 1B-5A, one 6A, and one 6B occupations with the appropriate number of Category A and Category B exemplars. For example, the two 1A-5B occupations were each paired with one Category A and five Category B exemplars. Each Category A and Category B exemplar was used in two different facts.

The 48 facts were typed on 20 index cards. Each card contained all facts from a single category (i.e., A or B) for a single occupation. For example, for each 1A-5B occupation (e.g., "the pianist"), there were two cards, one containing the one Category A fact (e.g., "The pianist likes wolves") and one containing the five Category B facts (e.g., "The pianist likes Italy," etc.).

Forty-eight test probes, 24 true and 24 false, were also generated (see Table 3 for details). These consisted of 8 2-0 probes, 16 1-1 probes, 8 1-5 probes, 8 5-1 probes, and 8 6-0 probes. Each occupation appeared in two true and two false probes. The occupations that were used in two different types of probes (i.e., the 1A-5B and 1B-5A occupations, which appeared in both 1-5 and 5-1 probes) each appeared in one true and one false probe of each type. Each of the 24 category exemplars appeared in one true and one false probe.

Four different block randomizations of the 48 probes were generated. Each randomization consisted of two blocks of 24

probes each, with half of the true probes and half of the false probes of each type in each block.

Apparatus. The experiment was controlled by a Commodore PET microcomputer system, which displayed probes and recorded responses and RTs.

Procedure. The experiment was conducted in two sessions, which were held on consecutive days. In the first session, subjects learned the 48 facts; in the second session, the subjects evaluated test probes.

At the start of the first session, the subject was told that the experiment involved learning simple imaginary facts. The subject was then given the deck of 20 cards containing the 48 facts and was asked to study the cards one at a time at his or her own pace. After the subject had studied all of the cards, the experimenter tested his or her knowledge of the facts by asking 20 questions of the form "What (category name) does the (occupation name) like?" (e.g., "What animals does the astronomer like?"). Each question corresponded to 1 of the 20 study cards; thus, the 20 questions covered all 48 facts.

Following this test, the subject studied the facts again, after which he or she was tested again, and so forth. On each study trial, the cards were presented in a different random order; on each test, the questions were asked in a different random order. Whenever the subject answered a question correctly on two consecutive tests, the corresponding study card was removed from the set of cards that the subject studied on subsequent trials. In addition, the question was not asked on subsequent tests. Study and test trials alternated until two consecutive correct answers had been given to all of the questions.

The subject then studied all of the cards once more and was tested with all of the questions. If the subject hesitated in answering any of the questions, he or she restudied the corresponding card(s) and was then retested. The first session ended when the subject was able to answer all of the questions (i.e., recall all 48 facts) with little or no hesitation.

In the second session, the subject first studied and was tested on the 48 facts until he or she evidenced no hesitation in recalling the facts. The subject was then seated facing a videoscreen. In front of the subjects were "true" and "false" response keys that were operated by the index and middle fingers, respectively, of the subject's right hand.

The subject first performed a 50-trial true-false practice task in which he or she pressed the appropriate key when the word "true" or the word "false" was displayed on the videoscreen. Test probes were then presented. Each subject received all four randomizations of the 48 probes. Five-minute rest breaks occurred after each set of 48 trials. Procedures for presentation of probes were otherwise the same as in Experiment 1.

Results and Discussion

Mean RTs and error rates are presented in Table 4. Mean RTs were calculated from correct responses to the 192 test probes. An analysis of variance revealed a main effect of truth value [$F(1,11) = 66.4, p < .05$],

Table 4
Mean RTs (in Milliseconds) and Error Rates (in Percentages)
for Test Probes in Experiment 2

Probe Type	True		False	
	RT	Error	RT	Error
2-0	1239	5	1395	6
1-1	1133	4	1424	5
1-5	1070	4	1667	7
5-1	1651	12	1879	15
6-0	1534	3	2083	33

indicating that RT was faster for true than for false probes. The main effect of probe type was also significant [$F(4,44) = 20.8, p < .05$], as was the Truth Value by Probe Type interaction [$F(4,44) = 9.1, p < .05$].

T tests were used to evaluate differences between various probe types. Because of the significant Truth Value by Probe Type interaction, tests were made separately for true and false probes. Six tests were performed, with the significance level for each test set at .01 in order to maintain an overall level close to .05.³

To determine whether subjects evaluating probes focused on relevant facts and ignored irrelevant facts, RT for 1-5 probes (i.e., those with one relevant and five irrelevant facts) was compared with RT for 5-1 (five relevant facts, one irrelevant fact) and 6-0 (six relevant facts) probes combined. For both true and false probes, RT was faster for 1-5 probes than for the other two types [$t(11) = 7.23, p < .01$, for true probes; $t(11) = 3.50, p < .01$, for false probes]. These results replicate the findings of Experiment 1 and indicate that subjects were able to focus on relevant facts.

If memory search focuses on relevant facts, we might also expect RT for 1-1 probes to be faster than that for 2-0 probes. Consistent with this prediction, RT for true 1-1 probes was 106 msec faster than that for true 2-0 probes, and this difference was marginally significant [$t(11) = 2.30, .01 < p < .05$]. For false probes, however, RT did not differ for 2-0 and 1-1 probes ($t < 1$). This latter result will be discussed below.

Finally, and most important, 1-1 and 1-5 probes were compared to determine whether probe RT was influenced by the number of irrelevant facts stored with the probed occupation. For true probes, there was clearly no effect of the number of irrelevant facts [$t(11) = 1.41, p > .10$]. (Note that the 63-msec difference between the 1-1 and 1-5 probes is in the direction of faster RT for the 1-5 items.) For the hypothesis that the 1-1 and 1-5 conditions differ by 100 msec, this test has a power of approximately .82 (assuming $\alpha = .05$). Thus, the probability of detecting any nontrivial effect of irrelevant facts was quite high in this experiment.

The error data support the conclusion that irrelevant facts do not influence RT for true probes: The error rate was the same (4%) for 1-1 and 1-5 probes.

In contrast to the true-probe data, the results for false probes revealed clear effects of the number of irrelevant facts: Mean RT was almost 250 msec slower for 1-5 probes than for 1-1 probes [$t(11) = 3.84, p < .01$]. The error rate was also slightly higher for 1-5 probes (7%) than for 1-1 probes (5%).

These results concerning irrelevant facts are not completely consistent with the predictions of either of the models we have discussed. The finding of no effect of irrelevant facts for true probes cannot readily be accounted for by the Relevance Filter model or by other nonselective access models. According to the nonselective access models, irrelevant facts are examined

during evaluation of a probe and so should influence RT.

The Relevant Subset Only model also has difficulty accounting for the results. This selective access model assumes that when a probe is presented, the relevant subset of facts stored with the probed occupation is accessed and searched. If a fact that matches the probe is found, a "true" response is made; if no match is found, a "false" response is executed. Subsets of irrelevant facts are never searched, and, consequently, the number of irrelevant facts stored with a probed occupation should not affect RT. Thus, the model cannot readily accommodate the effect of irrelevant facts obtained for false probes.

However, a minor modification of the Relevant Subset Only model will account for the present results. Specifically, the findings may be explained by assuming that when a search of the relevant subset failed to locate a fact matching the probe, subjects sometimes searched the other (i.e., irrelevant) subset before responding "false." Given this retrieval process, the data concerning irrelevant facts may be accounted for in the following way: For true probes, a search of the relevant subset will locate a fact that matches the probe. As a result, the irrelevant subset will not be searched, and the number of irrelevant facts will not affect RT. For false probes, however, a search of the relevant subset will not locate a match. Consequently, the irrelevant subset will sometimes be searched, and the number of irrelevant facts stored with the probed occupation will affect RT.⁴

The argument that subjects sometimes searched both the relevant and the irrelevant subsets before responding "false" may also explain our failure to find faster RT for 1-1 than for 2-0 false probes.

The interpretation we have offered raises two major questions. First, why would subjects sometimes search both subsets of facts before making a "false" response? Second, given that subjects do not always search both subsets of facts before responding "false," what determines whether or not both subsets will be searched? We will defer discussion of these questions until after presenting the results of Experiment 3. Although we will not be able to provide definitive answers, we will suggest some possibilities.

One other result from Experiment 2 deserves consideration—namely, the somewhat faster RT and lower error rate for 6-0 true probes than for 5-1 true probes. This result seems to run counter not only to the individual models we have discussed, but to the focused search notion in general. A possible explanation, suggested by the very high (33%) error rate for false 6-0 probes, is that subjects occasionally guessed "true" to 6-0 probes before evaluating them fully. A fast-guessing strategy of this sort would produce fast RTs and a low error rate for true 6-0 probes, along with a very high error rate for false 6-0 probes. Subjects may have adopted this strategy on the basis of a belief that 6-0 probes were very likely to be true. Subjects learned more facts

from a single category for the occupations in 6-0 probes (e.g., "the plumber") than for the occupations in other conditions. Thus, they may have formed impressions such as "The plumber likes a lot of animals." This sort of impression could have led subjects to believe that 6-0 probes like "The plumber likes elephants" were probably true and, consequently, could have caused them to guess "true" to this type of probe.

Although some of the findings of Experiment 2 are difficult to interpret, one result has straightforward implications. Specifically, the finding of no effect of irrelevant facts on RT for true probes (i.e., the finding of no difference in RT between 1-1 and 1-5 true probes) is consistent with selective access models of focused memory search and creates problems for nonselective access models. However, a strong case in favor of selective access cannot be made on the basis of a single null result. Hence, in Experiment 3, we attempted to replicate the results of Experiment 2.

EXPERIMENT 3

As in the previous experiments, subjects learned facts of the form "The (occupation) likes (object)." For 2A and 2B occupations, subjects learned two facts from the same category. For 2A-1B and 2B-1A occupations, subjects learned two facts from one category and one fact from the other category. Finally, for 2A-6B and 2B-6A occupations, subjects learned two facts from one category and six facts from the other category.

Five types of test probes—2-0, 1-2, 2-1, 2-6, and 6-2—were presented. The five probe types are described in detail in Table 5.

These probe types allow us to determine whether subjects evaluating probes focus on relevant facts and ignore irrelevant information. If memory search is focused, RT should be faster for 2-6 than for 6-2 probes and faster for 1-2 than for 2-1 probes.

By comparing RT for 2-0, 2-1, and 2-6 probes, we can assess the effects of the number of irrelevant facts stored with the probed occupation. If, as we have

Table 5
Test Probes for Experiment 3

Probe Type	Occupation Type	Exemplar's Category
2-0	2A	A
	2B	B
1-2	2A-1B	B
	2B-1A	A
2-1	2A-1B	A
	2B-1A	B
2-6	2A-6B	A
	2B-6A	B
6-2	2A-6B	B
	2B-6A	A

argued, irrelevant facts are sometimes searched for false but not for true probes, then we should find slower RT for 2-6 than for 2-1 false probes, but no difference in RT between 2-1 and 2-6 true probes.

Given a finding of no difference in RT between 2-1 and 2-6 true probes, we can, by comparing 2-0 true probes with 2-1 and 2-6 true probes, evaluate the argument that initial selection of the to-be-searched subset is faster when there is only one subset of facts stored with the probed occupation than when there are two or more subsets. This argument would be supported by a finding of faster RT for 2-0 than for 2-1 or 2-6 true probes. In contrast, a finding of no difference among the three probe types would suggest that there was little difference among these probe types in the time required to select the relevant subset.

Method

Subjects. Subjects were 12 undergraduate students at the Johns Hopkins University.

Materials. The occupation names and category exemplars from Experiment 2 were used to generate a different set of materials for each subject. Each set of materials utilized eight occupation names and 16 category exemplars, 8 from each of two categories.

Thirty-two to-be-learned facts were generated by pairing one 2A, one 2B, two 2A-1B, two 2B-1A, one 2A-6B, and one 2B-6A occupations with the appropriate number of Category A and Category B exemplars. Each exemplar was used in two facts. The 32 facts were typed on index cards in the manner described for Experiment 2.

Thirty-two test probes, 16 true and 16 false, were also constructed. These consisted of eight 2-0 probes, eight 2-1 probes, eight 1-2 probes, four 2-6 probes, and four 6-2 probes. Each occupation appeared in two true and two false probes. Each occupation that appeared in two different types of probes was used in one true and one false probe of each type. Each category exemplar appeared in one true and one false probe. Fourteen different block randomizations of the 32 probes were generated.

Apparatus. Apparatus was the same as in Experiment 2.

Procedure. The experiment was conducted in three sessions, which were held on consecutive days. In the first session, subjects were taught the 32 facts in the manner described for Experiment 2. The second and third sessions were test sessions in which subjects made true-false decisions for test probes. The procedure was the same for the two test sessions. Subjects first reviewed the 32 facts learned in the first session and performed the true-false practice task. Seven randomizations of the 32 probes were then presented. The first set of 32 probes was counted as a practice set. Rest breaks occurred after the practice set and after the third of the six test sets. Because of the very high error rate obtained for 6-0 false probes in Experiment 2, instructions to subjects strongly emphasized the importance of accuracy. Procedures were otherwise the same as in Experiment 2.

Results and Discussion

The RT data were submitted to an analysis of variance with session (first vs. second test session), truth value, and probe type as repeated measures. A main effect of session was obtained [$F(1,11) = 36.55, p < .05$], indicating that RTs were faster in the second test session than in the first. The interactions of session with truth value [$F(1,11) = 5.48, p < .05$] and probe type

[$F(4,44) = 2.66, p < .05$] were also reliable. These interactions do not reflect any change across sessions in the basic patterns of results; instead, they indicate merely that differences in RT between true and false probes and among the various probe types were smaller in the second test session than in the first. No other interactions involving the session factor reached significance. Hence, RTs and error rates are presented in Table 6 collapsed across sessions.

The analysis of variance also revealed a main effect of truth value [$F(1,11) = 35.70, p < .05$], reflecting the fact that RTs were faster for true than for false probes. Finally, the main effect of probe type [$F(4,44) = 15.34, p < .05$] and the Truth Value by Probe Type interaction [$F(4,44) = 4.60, p < .05$] were reliable.

Additional tests were conducted to examine differences among probe types in greater detail. Six tests were made, with the significance level for each set at .01.

To determine whether subjects evaluating probes focused on relevant stored facts, *t* tests were used to make comparisons of 2-6 vs. 6-2 probes and 1-2 vs. 2-1 probes. For true probes, RT was faster for 2-6 than for 6-2 probes [$t(11) = 3.47, p < .01$]. In addition, RT was faster for 1-2 than for 2-1 probes [$t(11) = 4.44, p < .01$]. Thus, with the total number of facts learned about the probed occupation held constant, RT for true probes was faster the smaller the number of relevant facts. The error data paralleled the RT results: Error rates were somewhat lower for 1-2 and 2-6 probes than for 2-1 and 6-2 probes, respectively. These results indicate that subjects evaluating true probes focused on relevant stored facts.

In contrast, the results for false probes showed no evidence of focused memory search: RT did not differ for 2-6 and 6-2 probes or for 1-2 and 2-1 probes ($t < 1$ for both comparisons). This finding suggests that subjects evaluating false probes always searched both relevant and irrelevant facts stored with the probed occupation.

In order to evaluate the effects of irrelevant facts in greater detail, RTs for 2-0, 2-1, and 2-6 probes were compared by means of one-way analyses of variance. For true probes, the three conditions did not differ in RT [$F(2,22) = 1.07, p > .25$].⁵ Thus, with the number of relevant facts held constant at two, RT did not vary with the number of irrelevant facts. In addition, there

Table 6
Mean RTs (in Milliseconds) and Error Rates (in Percentages)
for Test Probes in Experiment 3

Probe Type	True		False	
	RT	Error	RT	Error
2-0	1268	4	1484	7
1-2	1051	2	1391	6
2-1	1364	8	1429	6
2-6	1318	3	1880	9
6-2	1614	11	1877	13

was no systematic relationship between error rates and the number of irrelevant facts. For false probes, however, the number of irrelevant facts did influence RT [$F(2,22)=9.58$, $p < .01$]: RT for 2-6 probes was approximately 400 msec slower than that for 2-0 or 2-1 probes.

If we consider the results of Experiment 3 in conjunction with those of the first two experiments, a consistent pattern emerges. In the first place, the findings of all three experiments indicate that subjects evaluating probes focused on relevant stored facts and ignored, at least to some extent, irrelevant information.

In addition, the three experiments produced similar results concerning the effects on RT of irrelevant stored facts. This statement requires some discussion, because it seems at first glance that the results of Experiments 2 and 3 differ from those of Experiment 1. In Experiments 2 and 3, irrelevant facts did not influence RT for true probes, but they clearly did affect RT for false probes. In contrast, in Experiment 1, no significant effect of irrelevant facts was found. However, the test for effects of irrelevant facts in Experiment 1 (i.e., the comparison of 1-0 and 1-5 probes) was conducted using data collapsed over true and false probes, because the Truth Value by Probe Type interaction was not significant. If we reassess the effects of irrelevant facts by making comparisons for true and false probes separately, we find a pattern of results very similar to that obtained in Experiments 2 and 3. For true probes, RT was about 100 msec slower for 1-5 than for 1-0 probes, but this difference did not approach significance [$t(7)=1.38$, $p > .20$]. For false probes, however, RT was almost 240 msec slower for 1-5 than for 1-0 probes, and this difference is significant at the .05 level [$t(7)=2.57$]. It seems reasonable to conclude, then, that in all three experiments, irrelevant facts did not influence RT for true probes, but they did affect RT for false probes.

The effects of irrelevant facts on RT for false probes suggest that when a fact matching the probe was not found in the search of the relevant subset of facts, subjects sometimes searched the irrelevant subset before responding "false." As we mentioned earlier, this conclusion raises two questions. First, why would subjects sometimes search irrelevant facts before making a "false" response? One possibility is that subjects believed that searching all of the facts stored with a probed occupation was the only way to be completely sure that none of the facts stored with the occupation matched the probe. In other words, subjects may have believed that they could minimize the chances of responding "false" to a true probe only by searching both subsets of facts before making a "false" response. This sort of belief may not be as unreasonable as it seems. If, for example, the process of identifying the relevant subset is not entirely error free, then occasional "false" responses to true probes will occur if the subject responds "false" when a match for the probe is not found in the initially selected subset. For instance, if a subset of facts about

animals were identified as relevant for a true probe like "The banker likes Spain," a strategy of responding "false" after an unsuccessful search of the "relevant" subset would result in an error. This sort of error could be avoided by searching both subsets of facts before responding "false."

Note that it is not necessary to assume that a strategy of responding "false" after an unsuccessful search of the initially selected subset would result in frequent errors. Subjects might consider it worthwhile to search both subsets of facts before responding "false" even if this strategy served to avoid only a small number of errors.

The argument we have presented here is clearly based largely on speculation. However, it does seem to provide a plausible rationale for why subjects sometimes search both subsets of facts before making a "false" response.⁶

The second question raised by our interpretation of the false-probe results is that of what determines whether or not a subject will search both subsets of facts before responding "false"? A partial answer to this question is suggested by the preceding discussion. Specifically, if the strategy of searching both subsets of facts before responding "false" represents an attempt to increase accuracy, then we might expect that the greater the emphasis placed on accuracy, the more likely subjects would be to search both subsets before making a "false" response. The present results provide tentative support for this argument. In Experiments 1 and 2, in which no special emphasis was placed on accuracy, the finding of faster RT for 1-5 than for 5-1 and 6-0 false probes suggests that subjects sometimes made "false" responses after searching the relevant subset alone. However, in Experiment 3, in which accuracy was stressed, the finding of no difference in RT between 2-6 and 6-2 false probes seems to indicate that subjects always searched both subsets before responding "false."

The explanation we have offered for the false-probe results is both speculative and incomplete, and additional research will be required to determine whether it is a reasonable one. In particular, the viability of our interpretation will depend upon (1) whether or not evidence can be adduced in favor of the view that subjects search both subsets in order to increase accuracy and (2) whether or not the conditions under which subjects will search both subsets of facts can be specified.

One other aspect of the present results should also be discussed. The finding of no difference among 2-0, 2-1, and 2-6 true probes (as well as the analogous finding for 1-0 and 1-5 probes in Experiment 1) does not provide strong support for the argument that initial selection of the to-be-searched subset may be faster when there is only one set of facts stored with the probed occupation than when there are two subsets of facts. According to this argument, true probes with no irrelevant facts should evidence faster RT than those with at least one irrelevant fact. Although the RT

differences among conditions were in the predicted direction in both Experiment 1 and Experiment 3, these differences did not approach significance. Thus, while we cannot rule out the possibility that the number of subsets stored with a probed occupation affects the time required to select the relevant subset, it would appear that if there is an effect of the number of subsets, it is a rather small one (at least when the number of subsets varies from one to two).

Although it is interesting to consider questions regarding the number of subsets or the searching of irrelevant facts, the major conclusion to be drawn from the present results concerns the effects of irrelevant facts on RT for true probes. Specifically, the finding in all three experiments that the number of irrelevant facts did not influence true-probe RT argues in favor of selective access models of focused memory search and against nonselective access models. In other words, our results suggest that facts stored with a concept can be organized into subsets and that the facts in one subset can be accessed and searched without accessing the facts in other subsets.

GENERAL DISCUSSION

Several fact-retrieval studies have shown that the more facts a person learns about a concept, the longer it takes him or her to retrieve any of these facts (e.g., Anderson, 1974; Thorndyke & Bower, 1974). This result has generally been interpreted to mean that retrieval of a fact about a concept involves a search of all facts stored in memory with that concept. As Smith et al. (1978) have pointed out, however, this interpretation has a rather troublesome implication. Specifically, the interpretation implies that an expert on a particular topic should be very slow in retrieving facts about that topic. However, informal observation suggests to the contrary that experts on a topic can retrieve facts about the topic at least as quickly as nonexperts.

In the present study, we suggested that the expert's lack of difficulty in retrieving facts might be explained by assuming that retrieval usually involves not an unfocused search of all facts stored with a concept, but rather a focused search that concentrates on relevant stored facts and ignores irrelevant information. The results of the present study offer support for this argument. Specifically, the present findings suggest that a person's knowledge about a topic may be organized into small subsets of facts. Given this sort of organization, a person attempting to retrieve a fact about a concept need not search through all facts stored with that concept. Rather, he or she can simply determine which subset of facts is most likely to contain the sought-for fact and search that subset. Thus, regardless of the total amount of information stored with a concept, retrieval of a fact about the concept can be accomplished through a search of only a small number of facts. Therefore, an expert on a topic may, in retrieving a fact about the topic, search through no more facts

than would a nonexpert. In fact, the expert may often search fewer facts than the nonexpert. In comparison with the expert's knowledge on a topic, the nonexpert's information may be organized into fewer subsets with more facts per subset. In some cases the nonexpert's knowledge of a topic may even consist of an unstructured list of facts. Thus, the expert may often be able to retrieve a fact by searching only a small subset of facts, whereas the nonexpert may have to search a larger subset or even the entire set of facts stored about the topic.

The general description of the retrieval process presented here points out several directions for future research. In the first place, although the present results clearly indicate that memory search focuses on relevant facts, the specific assumption that retrieval involves selective access to subsets of stored facts rests primarily on the null finding of no effect of irrelevant facts for true probes. To be sure, this null result was obtained in all three experiments. Nevertheless, additional evidence will be required to make a convincing case in favor of selective access.

Assuming that additional research does in fact support the selective access position, several issues regarding details of the retrieval process deserve further investigation. For example, one question we have not considered is this: How does the process of identifying the relevant subset operate? In the context of the experiments reported here, a very simple process can be envisioned. Because the subsets into which facts were organized (i.e., animal facts, country facts) were the same for all of the occupations (e.g., "the banker"), subjects could identify the relevant subset for any probe by simply deciding whether the probe concerned an animal or a country. In more realistic situations, however, the process of identifying the relevant subset is probably much more complex. For one thing, the subsets into which real-world facts are organized are almost certainly different for different concepts. For example, stored facts about Abraham Lincoln might be organized into subsets concerning his childhood, his years as President, and his assassination, whereas facts concerning Mexico might be organized into climate, tourist attractions, and industries subsets. Thus, it seems that the process of identifying the relevant subset must somehow be keyed to the particular subsets stored with the concept in question. One possibility is that each concept has stored with it an index, or list of subsets, so that during a retrieval attempt the relevant subset may be selected by a process that scans the index in order to determine which subset is most likely to contain the sought-for fact.

Of course, even if we assume that a list of subsets is available for each concept, we leave unanswered the question of how a decision is made regarding which of the subsets is most likely to contain the desired information. It would seem that even for very simple queries (e.g., "Where was Lincoln born?" or "Why are plants green?"), identification of the relevant subset would

require something much more complex than the simple category membership test described for the present experiments.

Another issue for future research concerns how a decision is made to terminate a retrieval attempt. This issue is related to some degree to the questions regarding how information is organized and how the relevant subset is selected. For example, if information organization and subset selection are such that the subset initially identified as relevant is the only subset that could contain the sought-for information, then retrieval should always terminate after the search of this one subset. If, however, the subset initially selected is only one of several that might contain the desired fact, it would probably be reasonable to search more than one subset before concluding that the sought-for information is not to be found.

In conclusion, the results of the present study suggest that retrieval of information from long-term memory is complex yet efficient. However, the present data provide only a vague outline of the processes involved in retrieval. It is to be hoped that future research will enable us to characterize these interesting and important processes in greater detail.

REFERENCES

- ANDERSON, J. R. Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 1974, 4, 451-474.
- ANDERSON, J. R. *Language, memory and thought*. Hillsdale, N.J.: Erlbaum, 1976.
- ANDERSON, J. R., & BOWER, G. H. *Human associative memory*. Washington, D.C.: Winston, 1973.
- ANDERSON, J. R., & PAULSON, R. Interference in memory for pictorial information. *Cognitive Psychology*, 1978, 10, 178-202.
- 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.
- CRAIN, R. D., & DE ROSA, D. V. Retrieval of information from multiple ensembles in short-term memory. *Memory & Cognition*, 1974, 2, 255-260.
- 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.
- NAUS, M. J., GLUCKSBERG, S., & ORNSTEIN, P. D. Taxonomic word categories and memory search. *Cognitive Psychology*, 1972, 3, 643-654.
- OLDFIELD, B. C. Things, words and the brain. *Quarterly Journal of Experimental Psychology*, 1966, 18, 340-353.
- SMITH, E. E., ADAMS, N., & SHORR, D. Fact retrieval and the paradox of interference. *Cognitive Psychology*, 1978, 10, 438-464.
- STERNBERG, S. High-speed scanning in human memory. *Science*, 1966, 153, 652-654.
- STERNBERG, S. Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 1969, 57, 421-457.
- THORNDYKE, P. W., & BOWER, G. H. Storage and retrieval processes in sentence memory. *Cognitive Psychology*, 1974, 5, 515-543.
- WESCOURT, K. T., & ATKINSON, R. C. Fact retrieval processes in human memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 4). Hillsdale, N.J.: Erlbaum, 1976.

NOTES

1. In describing the retrieval process as involving a search through all facts stored with a concept, we do not mean to imply that the search is necessarily exhaustive. Instead, we mean simply that the search process operates on the set of all facts stored with the concept in question. The argument that retrieval should take longer the greater the number of stored facts follows from most self-terminating or exhaustive, serial or parallel formulations of the search process.

2. Although subjects evaluating probes (e.g., "The lawyer likes rabbits") may search facts stored with the category exemplar (e.g., "rabbits") as well as facts stored with the occupation (e.g., "lawyer"), we will consider only the search of facts learned about the occupation. We emphasize facts stored with the occupation because in the experiments reported here we manipulated the number of relevant and irrelevant stored facts only for the occupation terms. In all of the experiments, the same number of facts (i.e., two) was learned for each of the category exemplars. Hence, while subjects may search facts stored with the exemplar in a probe, this type of search will not contribute to differences in RT among the various probe types.

3. Data for 5-1 and 6-0 probe types (both of which involve six total facts and a large number of relevant facts) were combined for purposes of making t tests. Combining these two probe types reduces the number of tests and thus allows us to use a reasonable significance level (i.e., .01) for individual tests. Collapsing the 5-1 and 6-0 data does not eliminate any important comparisons. Although differences between 5-1 and 6-0 probes cannot be evaluated, interpretation of these differences would in any case be difficult: It appears from the data that subjects occasionally made fast-guess "true" responses to 6-0 probes (see discussion at the end of Experiment 2).

4. We argue that subjects sometimes (as opposed to always) searched both subsets of facts before responding "false" on the basis of the finding of faster RT for 1-5 than for 5-1 and 6-0 false probes. If both subsets of facts were always searched before a "false" response was made, we would expect no difference in RT among these probe types. The obtained pattern of results for false probes (i.e., $1-1 < 1-5 < 5-1$ and 6-0) might be taken to suggest that some subjects always responded "false" after searching only the relevant subset, whereas other subjects always searched both relevant and irrelevant subsets before making a "false" response. However, examination of the data for individual subjects indicates that this was not the case; hence, our conclusion was that subjects sometimes but not always searched the irrelevant subset before responding "false."

5. For the hypothesis $2-0 < 2-1 < 2-6$ with 100-msec differences between conditions, the power of this test is approximately .75 when $\alpha = .05$. For differences of 150 msec, the test has a power of approximately .97.

6. It might appear that this argument is contradicted by the finding that irrelevant facts had no effect on RT for true probes. Specifically, it might seem that if subjects evaluating true probes occasionally erred in identifying the relevant subset and consequently searched the irrelevant facts before examining the relevant information, we would expect an effect of the number of irrelevant facts on "true" RT. In actuality, however, as long as the irrelevant subset is searched for only a small proportion of true probes, the effect of irrelevant facts on "true" RT should be extremely small. For example, assume that misidentification of the relevant subset resulted in a search of irrelevant facts for 5% of the true probes. If an irrelevant subset containing one fact can be searched in 100 msec and a five-fact subset requires 500 msec, we would expect a mean difference of $400 \times .05 = 20$ msec between true probes with one irrelevant fact and those with five irrelevant facts. A difference of this magnitude is, of course, far too small to be detected in the experiments reported here.