Journal of Experimental Psychology:
Learning, Memory, and Cognition
1987, Vol. 13, No. 3, 355-367

0278-7393/87/$00.75

Effects of Number of Facts Studied on Recognition Versus Sensibility Judgments

John R. Anderson and Lynne M. Reder
Carnegie-Mellon University

The effect of memory representation on the ease of cognitive judgments can vary depending on the nature of the judgment required. In three experiments, subjects studied sets of sentences and in later phases made recognition judgments or sensibility judgments on these sentences and others constructed from the words in the sentences they had studied. In Experiments 1 and 2, the studied sentences were sensible, whereas in Experiment 3 the studied sentences were nonsensical. Judgment times varied with the *fan* of the concepts in the sentence (i.e., the number of facts known about each concept). Subjects were slowed by fan in retrieving a specific fact but speeded by fan in making a sensibility judgment. In all experiments, subjects were faster at making a judgment in conditions where judgments could be made either by a memory retrieval process or a semantic sensibility process. This implies that subjects can sometimes recognize that they have studied a sentence before they can judge its sensibility. This result calls into question the view that language processing is a faculty that occupies a place separate from memory.

Some of the premiere questions associated with how memory is organized concern the division of faculties within the memory structure. For example, do we have a separate lexicon (Collins & Loftus, 1975; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982)? As another example, do we have distinct semantic and episodic memories (e.g., McKoon & Ratcliff, 1979; Tulving, 1983)? To address such questions, researchers have taken the tack of seeing whether manipulations known to affect one task also affect another (e.g., Seidenberg, Waters, Sanders, & Langer, 1984). If the manipulation affects both, there is evidence (although not conclusive) that it is not necessary to distinguish among faculties (e.g., Anderson & Ross, 1980; Muter 1978). This research is motivated by similar concerns: Will low-level semantic decisions underlying our ability to parse sentences be affected by the variables that affect episodic types of judgments?

In the past decade, there have been numerous demonstrations that the mechanisms used to retrieve information from memory are sensitive to the relations among the contents of memory. Many demonstrations have shown that the time to recognize a fact when other facts share some of the same concepts is slowed in proportion to *fan* (i.e., the number of facts sharing the concepts; see Anderson, 1983b, for a review). More recently, this result has been qualified by the nature of the relation among the facts sharing the concepts or judgments required of the subject (e.g., Keenan & Brown, 1982; Myers, O'Brien, Balota, & Toyofuku, 1984; Reder & Anderson, 1980; Reder & Ross, 1983; Reder & Wible, 1984; Smith, Adams, & Schorr, 1978). Reder and Ross (1983) and Reder and Wible (1984) actually found response times to be faster to decide that a statement was consistent with studied statements, the greater the number of related statements that had been studied using the same concepts.

It is perhaps not surprising that consistency judgments are influenced by the same memory representation variables that influence recognition decisions. Consistency judgments depend in part on knowing what was studied, just like recognition judgments. The surprising aspect of the contrast between the two tasks is that the effect of the memory representation variables is exactly opposite; reaction time (RT) facilitation, instead of interference, with increased fan (Reder & Ross, 1981, 1983; Reder & Wible, 1984). Note that this is quite different from the results of researchers like Myers et al., (1984), who found that fan effects for the same task varied depending on what the subjects committed to memory.

One might wonder, however, whether there would be any effect of the number of facts studied about a concept on judgments that require the subject to retrieve semantic as opposed to episodic information in order to make a decision (Tulving, 1983). Anderson and Ross (1980) found that the time to make a semantic memory judgment of the form *an x is a y* was affected by the amount of information previously studied. Subjects took longer to reject a foil (e.g., *a collie is a cat*) when more sentences had been studied that involved both collie and cat (e.g., *the collie chased the cat*). On the other hand, the time to accept a target tended to decrease with the number of studied sentences.

These studies indicate that fan influences a wide range of memory judgments, although not always in the same way. Whether positive or negative effects are obtained seems to depend on the judgment the subject makes (e.g., recognition vs. consistency, acceptance vs. rejection). It does not seem to depend on the type of sentence representation involved because opposite effects emerge for the same subject on the same studied material. This kind of result seems relevant to the issue of separate faculties for different types of knowledge. The experiments

---

This research was supported by National Science Foundation Grant BNS-82-0889 to John R. Anderson and by National Science Foundation Grant BNS-83-03711 to Lynne M. Reder.

We thank G. Wells for her comments.

Correspondence concerning this article should be addressed to John R. Anderson, Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

reported in the present article attempt to show the same kind of result with a different pair of judgments: recognizing word strings and deciding their sensibility as sentences.

Our basic experimental procedure involved teaching people statements in which the subjects and predicates could be used in more than one sentence. We wanted to see how sensibility and recognition judgments involving these terms were affected by fan. We were particularly interested in the low-level, semantic judgments of deciding whether a string of words is sensible or not. An example of a sensible sentence is *The box was filled with ornaments;* an example of a nonsensical sentence is *The box had mumps when he was eleven.*

One prevailing conception of language comprehension (e.g., in Anderson & Bower, 1973; Clark, 1974; Fodor, 1983; Kintsch, 1974; Norman & Rumelhart, 1975) assumes that a parsing process first interprets the sentence and develops an initial representation of its meaning. This representation is then output to be interfaced with memory or to be matched in some other way, for instance, compared with the parse of a picture in a sentence–picture verification task. This parsing process would bring to bear the selectional restrictions required to make sensibility judgments of the sentences in our experiments. Because these sensibility computations are made before interfacing with memory, the parsing process should not be affected by information in episodic memory. For instance, there should be no difference between judging the sensibility of a sentence which has been studied and one which has not. Furthermore, there should be no effect of the presence of other episodic facts learned about the concepts.[1]

Other theories, such as Anderson's (1976, 1983a) ACT theory and Schank and Birnbaum's (1984) make different assumptions. The ACT theory assumes that all processing, including language processing, proceeds by means of productions operating on a single memory which is not differentiated into semantic and episodic memory. According to this view, it is possible for a subject to judge the sensibility of a sentence by matching it to a stored sentence or by retrieving information about selectional restrictions. As a consequence, it should be possible to more rapidly judge the sensibility of a previously studied sentence. In other words, this research can be seen as a test between a theory that assumes language processing occupies some special place in the information processing separate from a long-term memory data base (see Chomsky, 1980; Wexler & Culicover, 1980) and a model like ACT that sees the computations underlying language processing as mixed in with other processing, all occurring in a common data base.

Predictions

To make some specific predictions about performance in these tasks, we need to specify how we think subjects make recognition judgments and how they make sensibility judgments.

First, our model for recognition judgments is the same as the one we have presented elsewhere (e.g., Anderson, 1983b; Reder, 1982, 1987). When subjects read a sentence like *The box was filled with ornaments,* they form a probe from the content words (box, filled, ornament) and check to see if there is a memory trace involving these concepts. If there is, they recognize the sentence; if not, they reject the sentence.[2]

Second, memory retrieval is involved in making sensibility judgments also. Subjects retrieve the category of the sentence subject (*the box*) and of the sentence predicate (*filled with ornaments*). In this case, both would be containers; in the case of a sentence like *The box had mumps,* one would be container and one would be person. If the two categories match, the sentence is judged as sensible; if they mismatch, they are judged as nonsensical. In principle, sentences could be nonsensical in ways other than a subject–predicate category mismatch, but this was the type of nonsensical sentence we used. Therefore, we assume that this is what subjects checked for.

Having specified the basic judgment processes, we can derive some predictions for the experiments:

1. Subjects should make their judgments more rapidly in circumstances in which they can use either a recognition or a sensibility judgment successfully. (For instance, subjects can decide a studied sentence is sensible either by retrieving it or testing the category of subject and predicate.) In Reder's framework (1982, 1987; Reder & Ross, 1983; Reder & Wible, 1984) subjects are not consistent in the procedures they use to make judgments. Sometimes they adopt a procedure that seems quite appropriate given the task demands, and other times they adopt a different strategy that may or may not work depending on the situation. For example, sometimes subjects try to decide if a statement was studied by deciding whether or not the statement is sensible; other times, subjects try to decide whether or not a statement is sensible by trying to see if that exact statement had been studied. When the statement had been studied and is sensible, either procedure will work for either task; however, in conditions where both procedures will not produce the correct judgment, subjects sometimes have to go on to try the appropriate procedure, having first tried the less appropriate one. When either procedure will work, total judgment time is not inflated by the selection of the wrong procedure. In Anderson's framework (Anderson, 1983a) the procedures can apply in parallel. Under both frameworks, subjects will be better off if either procedure will work.[3]
2. Recognition memory should slow down as more additional facts about a subject or predicate are studied. In both of our theories this is because of a decrease in activation of any particular fact.
3. Ability to make sensibility judgments should improve as a subject or predicate has been studied in more sentences. This is because each sentence provides practice for the connection between the subject and its category and the predicate and its category. This assumes that when subjects study sentences they

---

[1] An interesting question is whether the aforementioned authors really intended to imply that comprehension must be complete before a sentence can be verified. We can report that Anderson and Bower (1973) were aware of this implication of their theory and were uneasy about it; however, at the time, they could not think of a satisfactory way to put this implication to a test.

[2] It is true that the concepts might intersect in a memory trace but have been studied in a different combination (e.g., *the ornament was filled with boxes*). However, we did not use such foils and so we assume subjects did not check for them. Anderson (1976), showed that subjects take longer when there are such foils.

[3] We do not think that the parallel–serial distinction is a substantial difference between our two models. We are convinced by demonstrations of the equivalence between the two frameworks at our level of analysis (Anderson, 1976; Townsend, 1974).

Table 1
Design of Study Materials

No. of facts using subject term	No. of facts using same predicate: 1	2
1	a-A	c-E
1	b-B	d-F
2	e-C	e-E
2	f-D	f-F

Note. Lowercase letters stand for subjects; uppercase letters, for predicates. A repeated letter means the subject or predicate was repeated.

normally check the compatibility between subject and predicate category.

Note that the distinction between Predictions 2 and 3 turns on whether the subject is using knowledge specific to one sentence (Prediction 2) or common to all sentences (Prediction 3). This is the basic distinction Reder and Ross (1983) used to explain positive fan effects for exact recognition judgments and negative fan effects for consistency judgments.

Although Prediction 2 seemed highly likely on the basis of past experiments, neither Prediction 1 nor 3 seemed obvious on the basis of past research and the variety of theoretical positions in the literature. Therefore, these experiments promised to be a significant test of our theoretical positions. Predictions 2 and 3 will be complicated by the fact that subjects can make their decisions by either of two processes in some situations. We have specific expectations about how the patterns will vary as a function of the potential use of different processes. After describing the details of our experimental paradigm and the results of Experiment 1, these more complicated predictions will be derived.

Experiment 1

Method

Subjects. A total of 20 subjects were recruited from the undergraduate population. They either received credit for an introductory course or earned $2.50 per hour. The experiment lasted less than 2 hr.

Design and materials. The subjects and predicates for this experiment are taken from the set given in the Appendix. The subject terms were taken from the categories *people, containers,* and *events,* and the predicates were constructed so that they were appropriate for subjects in only one of the categories. Subjects learned eight person facts, eight container facts, and eight event facts. Within each of these categories, the eight facts could be classified according to a 2 × 2 design in which either one or two sentences were studied about the subject term and one or two sentences were studied about the predicate. Table 1 illustrates the design abstractly by using lowercase letters to denote subjects and uppercase letters to denote predicates. So, for instance, *c-E* from the 1-2 cell is created from a subject *c* that occurred in just that sentence and a predicate *E* that also occurred in the *e-E* sentence from the 2-2 cell. As can be seen, two sentences were assigned to each cell of this design for each category. Table 2 gives examples of the person-fact sets and event-fact sets organized into 2 × 2 matrices. Table 3 gives example test sentences based on the facts in Table 2.

In addition to these target sentences, it was necessary to create test sentences that were sensible but not studied and ones that were nonsensical and not studied. The former were created by combining subjects and predicates from within the cells of the design. So, for instance, *e-E* and *f-F* were studied in the 2-2 cell. Nonstudied sensible sentences were created by making the new combinations *e-F* and *f-E.* Note that these sentences can be categorized according to the same 2 × 2 classification as the target sentences. Nonsensical test sentences were created by pairing subjects with corresponding predicates from another type. So, if we have the *a'-A'* sentence in the 1-1 cell from the *person* material and *a''-A''* in the 1-1 cell from the *event* material, we can create nonsensical items *a'-A''* and *a''-A'*. These items also can be classified according to the 2 × 2 variation in subject–predicate fan.

Procedure. The assignment of materials to conditions was randomly determined for each experimental subject. The experiment was computer controlled; the program generated and displayed the materials on computer terminals and collected subjects' responses and response times. The first phase of the experiment involved presenting subjects with the 24 target sentences to study at the rate of 10 s per sentence. The order of the sentence presentation was also randomly determined for each subject. After studying the material, subjects were given a "drop-out" memory test of the material. This involved presenting predicates to the subjects and having them type the one or two subject terms that occurred with that predicate. There were 18 predicates in all and they were presented in random order. Any predicates that the subject responded to incorrectly were immediately corrected and presented for further study. They were tested again (after other test trials), and this

Table 2
Examples of Study Sentences

No. of facts about person or event	No. of facts using same predicate: 1	2
	Person	
1	The soldier had two daughters and a son.	The banker wrote a short autobiography.
1	The farmer fell and broke an ankle.	The doctor had mumps when he was eleven.
2	The engineer was given a speeding ticket.	The engineer wrote a short autobiography.
2	The fireman learned French in high school.	The fireman had mumps when he was eleven.
	Event	
1	The fair was the event of the year.	The rally was cancelled by the organizer.
1	The rodeo was ended by a thunderstorm.	The dance was postponed until Thursday.
2	The concert was started wih a speech.	The concert was cancelled by the organizer.
2	The festival had been planned for months.	The festival was postponed until Thursday.

Table 3
Examples of Test Sentences

No. of facts studied about subject and predicate	Sentence
Studied	
2-2	The fireman had mumps when he was eleven.
1-2	The rally was cancelled by the organizer.
Nonstudied sensible	
2-2	The fireman wrote a short autobiography.
1-2	The banker had mumps when he was eleven.
Nonstudied nonsensible	
2-2	The concert had mumps when he was eleven.
1-2	The rally wrote a short autobiography.

continued until the subject had given one successful recall to each predicate, and then this drop-out procedure was repeated.

After completing the drop-out procedure twice, subjects went on to the critical reaction time phase of the experiment. There were two types of judgments that subjects had to make. One involved judging whether or not a sentence had been studied and the other involved judging whether or not a sentence was sensible. The judgments were divided into blocks of trials. There were four blocks of trials, each 196 trials long. Different subjects received one of two orders of the four blocks: recognition–sensibility–sensibility–recognition or sensibility–recognition–recognition–sensibility. A recognition block consisted of 96 studied sentences (each sentence presented four times), 48 sensible nonstudied sentences (each possibility two times), and 48 nonsensical sentences (constructed as described earlier and selected so that each subject and predicate occurred in the same relative frequency as in the targets). A sensibility block consisted of 48 studied sentences, 48 sensible nonstudied sentences, and 96 nonsensical sentences. Order of material within a block was random. There were short rest breaks of 30 s after each 64 sentences, with the option for a longer break after each block.

Each RT trial began with the presentation of a sentence on the CRT screen. The subject pressed either the *d* or *k* key to indicate a *no* or *yes* response, respectively. After each trial, the subject's response time and whether the response was accurate were displayed for 1 s. Then there was an additional 1 s of rest before the next trial began. Thus, the disappearance of the feedback served as a warning for the next trial.

Results and Discussion

In computing mean RTs, any incorrect times were excluded and response times were truncated to 5 s to avoid distorted means due to extreme scores. The mean correct RTs and proportion correct are displayed in Table 4. Analyses of variance were performed to obtain an estimate of the standard error of the means. For the recognition judgments the standard error of the reaction times was 33 ms and the standard error of the percent error was 2.3%. For the sensibility judgments the standard error of the reaction times was 20 ms and the standard error of the percentage of error was 1.3%. All of these are based on Subject × Condition interaction. With 20 subjects and 12 conditions there are 209 (19 × 11) degrees of freedom. For recognition judgments, the correct response for studied is *yes;* the correct response for mispairings (sensible) and for mispairings (nonsense) is *no.* For sensibility judgments, *yes* is the correct response for the first two categories and *no* is the correct response for the last category.

Table 4 presents the results of the experiment. The 48 numbers in that table present a fairly complex pattern and, given the standard errors noted above, not all of the apparent differences are real. However, on analysis these data present a systematic and interpretable pattern. First of all, subjects are more accurate and faster on average in making sensibility judgments $t(418) = 16.06, p < .01$, and $t(418) = 7.87, p < .01$, respectively. This presumably reflects the greater lifetime practice people have at judging sensibility. For the recognition data and for the sensibility data, reaction time and error rates are strongly correlated ($r = .98$ and $r = .60$, respectively). So subsequent analyses will focus on reaction time within the recognition data and within the sensibility data.

Within each of these subsets of data there are 11 degrees of freedom that can be partitioned into 2 degrees of freedom reflecting the differences among the three submatrices, 3 degrees of freedom reflecting the effect of subject fan in each matrix, 3 degrees of freedom reflecting predicate fan, and 3 degrees of freedom reflecting the interaction between the two types of fan. The interaction effects are not significant, $F(3, 207) = .07, p > .5$ for the recognition data and $F(3, 209) = 1.40, p > .1$ for the sensibility data.

By comparing the marginals we can measure subject and predicate fan. Generally, the predicate fan appears to be more potent, probably reflecting the fact that the predicate contains multiple concepts from which activation can spread. However, we will just report a combined measure of both fans, which is equivalent to looking at the difference between the 2-2 and 1-1 cells. We are simply interested in fan effects, not their differential size for the two dimensions.

Recognition. Subjects are much slower at rejecting sensible mispairings than at accepting studied sentences, $t(209) = 8.79$, $p < .01$, which reflects the frequently found disadvantage of negative responses. More interestingly, subjects are much faster at rejecting nonsensical mispairings than sensible mispairings, $t(209) = 13.33, p < .01$. We expected this result because nonsensical mispairings can be rejected by either of two processes: the statement is nonsense and it was not studied. In fact, these statements are rejected significantly faster than the studied sensible sentences are recognized, $t(209) = 4.55, p < .01$.

Although nonsensical statements can be correctly rejected by either of two processes, studied sensible and mispaired sensible statements can only be correctly judged by using the recognition process. For these types of sentences, the fan effects (defined as the difference between 2-2 and 1-1) are 60 ms and 129 ms, respectively. The first is not significant, $t(209) = 1.29$ but the second is, $t(209) = 2.88, p < .01$.[4] These fan results are consistent with Prediction 2, although as we noted in the introduction, given past results, this is not surprising.

There is a negative fan effect (−71 ms) for the nonsensical mispairings which is not significant, $t(209) = 1.52$. On the other hand, the difference in the size of the fan effect between the sensible and nonsensical sentences is significant, $t(209) = 2.90$, $p < .01$.

---

[4] We will soon explain why we expect a larger fan effect in the former case.

Table 4
Results from Experiment 1: Mean Reaction Times (in Milliseconds) for Correct Judgments and Proportion Errors

	Predicate fan					
	Recognition judgment			Sensibility judgment		
Subject fan	1	2	*M*	1	2	*M*
	Studied, sensible					
1						
RT	1,085	1,091	1,088	910	910	910
E	.058	.050	.054	.029	.054	.041
2						
RT	1,119	1,145	1,132	943	885	914
E	.090	.085	.088	.013	.050	.031
M						
RT	1,102	1,118	1,110	927	897	912
E	.074	.068	.061	.021	.052	.036
	Mispaired, sensible					
1						
RT	1,255	1,370	1,313	980	941	961
E	.229	.308	.268	.071	.067	.069
2						
RT	1,251	1,384	1,318	1,006	909	958
E	.200	.354	.277	.071	.075	.073
M						
RT	1,253	1,377	1,315	993	925	959
E	.215	.331	.273	.071	.071	.071
	Mispaired, nonsensical					
1						
RT	1,043	998	1,021	1,022	1,012	1,017
E	.054	.017	.035	.073	.081	.077
2						
RT	1,001	972	987	1,011	1,004	1,007
E	.029	.017	.023	.081	.100	.090
M						
RT	1,022	985	1,004	1,017	1,008	1,012
E	.041	.017	.029	.077	.090	.084

Sensibility judgments. Rejection of nonsensical mispairings is significantly slower, $t(209) = 3.67, p < .01$, than acceptance of sensible mispairings for sensibility judgments, which is another instance of the general phenomenon that negative judgments take longer. More important, acceptance of studied sensible sentences is faster than acceptance of sensible mispairings, $t(209) = 3.39$, $p < .01$. This again confirms Prediction 1 because the sensible studied sentences can be accepted either by a sensibility or a recognition judgment.

Although studied sensible statements can be accepted by either of two processes, the sensible mispairings and the nonsensical mispairings can only be correctly judged by using the sensibility process. Here the failure to recognize the sentence provides no evidence as to whether it is sensible. The fan effects for the two conditions are both negative, −71 ms and −18 ms, respectively. The first is significant, $t(209) = 2.51, p < .05$, and the second is not, $t(209) = .63$. The average effect is significant, $t(209) = 2.25$, $p < .05$. This confirms Prediction 3, that increased fan should facilitate sensibility judgments. There is also a nonsignificant negative fan effect (−25 ms) in the case of studied sensible. Below we explain why some conditions produce stronger fan effects than others.

Discussion of Fan Effects

This experiment confirms the three predictions given. For judgments that require direct fact retrieval, there is a conventional fan effect. For judgments that require a sensibility decision, there is a reverse fan effect. The size of these effects varied depending on the condition. Although these size variations were largely nonsignificant, they formed a pattern that was predicted by our theories and supported across all three experiments to be reported. Table 5 presents the relevant data from the three experiments.

In Table 5, each of the six conditions is represented in order of predicted fan effect from largest positive fan effect to largest negative fan effect. We have computed for each experiment the actual signed difference between the 2-2 and 1-1 fan conditions. These predictions are explained below in terms of Reder's (1982, 1987) strategy selection model (see also Reder & Ross, 1983; and Reder & Wible, 1984), which posits differential mixtures of strategy use, depending on the task.

1. Recognition of sensible mispairings. When the statement to be recognized is sensible but nonstudied the subject can come to a correct rejection only by using a direct retrieval strategy. It is still possible that the subject might first try a sensibility strategy. In that case the sentence is judged to be sensible, and the subject might venture a guess of "yes" even though the appropriate task is to make a recognition judgment; however, for sen-

Table 5
Fan Effects (in Milliseconds) for the Conditions of the Three Experiments

	Experiment			
Condition	1	2	3	*M*
	Recognition			
++ 1. Mispaired, sensible				
2-2 vs. 1-1	129	40	211	127
1-1 vs. 0-0	—	280	117	199
+ 2. Studied, sensible				
2-2 vs. 1-1	60	38	148	82
+− 3. Mispaired, sensible				
2-2 vs. 1-1	−71	2	83	5
1-1 vs. 0-0	—	94	215[a]	155
	Sensibility			
+− 4. Studied, sensible				
2-2 vs. 1-1	−25	−41	22	−15
− 5. Mispaired, sensible				
2-2 vs. 1-1	−18[a]	−57	−45	−40
1-1 vs. 0-0	—	50	31	40
−− 6. Mispaired, sensible				
2-2 vs. 1-1	−71	−69	−44[a]	−61
1-1 vs. 0-0	—	−104	−28	−66

Note. For Experiment 3, the sensible and nonsensical categories are reversed (e.g., the mispaired nonsensical cell is in fact the mispaired sensible cell). Plusses and minuses refer to magnitude of effect: see text for further explanation.
[a] Out of predicted order.

sible mispairings these guesses would cause error responses and would be excluded from the RT means. It is not safe to respond on the basis of an affirmative sensibility judgment because sensible sentences can either be studied or not, so typically the subject goes on to try direct retrieval when sensibility yields an affirmative judgment in the recognition task. Thus, in this condition, all correct RTs involve use of the direct retrieval strategy (either exclusively or after using sensibility). Direct retrieval is the strategy that produces positive fan effects.[5] We will denote this condition with two plusses in Table 5 indicating the strongest positive fan tendency.

2. *Recognition of sensible studied sentences.* In contrast to the sensible mispaired condition, if the subject tried the sensibility strategy first and ventured a guess of *yes* because it was sensible, the response would be correct and the reaction time would be included in the data. This means that correct RTs in this condition reflect a mixture of subjects trying direct retrieval first, trying sensibility first but going onto direct retrieval because the sentence was sensible, and trying sensibility first and then guessing that the sensible sentence was studied without going on. We denote this condition with a single plus (+) because the positive fan effects of direct retrieval are mixed with a minority tendency to use sensibility (which produces a negative fan effect), yielding overall a smaller positive fan effect.

3. *Recognition of mispaired, nonsensical sentences.* Because no nonsensical sentences were studied, using either the direct retrieval or sensibility strategy will yield a correct response. Subjects knew they could quit after using the sensibility process because if a statement is nonsensical, it was not studied. Therefore, we assume a relatively equal mixture of the recognition and sensibility processes. The mixture of positive and negative fan effects should cancel each other out. Thus, we denote this condition as "+ −."

4. *Sensibility judgments of studied sensible sentences.* In the case of sensibility judgments, a sentence can be accepted either if it is recognized (because only sensible statements were studied) or if it is judged sensible. Therefore, sentences in this condition, as in the previous one, should involve a relatively equal mixture of recognition and sensibility strategy use. Therefore, we denote this condition as a "+ −" also.

5. *Sensibility judgments of nonsensical mispairings.* Because some mispairings (nonstudied sentences) are sensible and some are not, it is not possible to use a negative recognition judgment to reliably decide sensibility. However, subjects might still guess "no" on the basis of the outcome of a negative recognition judgment to nonsensical pairings, and these would be counted as correct in this condition. Thus, this condition consists of a mixture of subjects sometimes trying sensibility first and deciding on that basis, other times trying direct retrieval and going on to sensibility because the sentence had not been studied, and still other times trying direct retrieval and guessing "no" because the statement had not been studied. Only the last possibility will produce a positive fan effect, and we assume it is a minority course of action because it is not guaranteed to be correct. Therefore, we expect a shallow negative fan effect and represent the predictions of this condition as "−."

6. *Sensibility judgments of sensible mispairings.* In contrast to the previous condition, if the subject responds on the basis of a negative recognition judgment in this condition, such a response would be an error, and the reaction time would be excluded. All correct judgments would be based on a sensibility process which produces a negative fan effect. Therefore, we represent our prediction for this condition as "− −."

Although these predictions were first derived from Reder's (1982) serial strategy selection model, the same predictions could be derived from the parallel race model of Anderson (1976, 1983a). In fact, that model would have the same basic analysis: That is, in some conditions only the sensibility or the direct retrieval process could produce a correct judgment giving a pure fan effect; in other conditions, both could apply, which would cancel out each other's effects; and in other conditions, subjects may occasionally get credit for a response on the basis of an invalid process. Again, this is just a further instance of the mathematical equivalence of parallel and serial models.

Both theories predict a complete ordering of fan effects for the six conditions with the exception of the nonordering of Conditions 3 and 4 in Table 5. Table 5 reports the fan effects defined by the difference between the 1-1 and 2-2 conditions. In Experiment 1 there is only one reaction time that does not fall within the correct ordering. This is the fact that the fan effect for sensibility judgments of nonsensical mispairings is slightly less negative than the fan effect for studied sensibles. In all, our theory predicts 14 pair-wise orderings among the six conditions in Table 5, and 13 are confirmed. It turns out that of the 720 possible permutations of these conditions, only 10 would preserve 13 out of these 14 predictions. Thus, we are predicting significantly better than chance. The theory is clearly picking up something systematic in the ordering of the fan effects. Subsequent experiments will offer more evidence and more varied evidence on how well we are predicting the ordering of the fan effects.

Experiment 2

Experiment 2 was a replication of Experiment 1 except that we decided to look at a more extreme manipulation of fan. We included subjects and predicates in test sentences that had not been presented as part of the set of study sentences. These unstudied terms were classified as *zero fan.* To control for differences in RT attributable to familiarity effects, the unstudied predicates were given exposure by having subjects rate them for pleasantness.[6]

Method

Subjects. A total of 20 subjects were recruited from the same population as in the first experiment. Again, subjects could either receive credit

---

[5] There have been a number of analyses of why positive fan effects are obtained for foils (Anderson, 1976, 1983a; King & Anderson, 1976). All involve some sort of waiting-time model in which the subject adjusts the time he or she waits to reject a foil to reflect the fan of the elements in the foil. For a mechanism that implements this waiting process, see Anderson (1983, pp. 111–112).

[6] The idea of testing zero-fan has some precedent in prior work (e.g., Anderson & Paulson, 1978; Reder & Anderson, 1980; Reder & Ross, 1983); however, in those contexts, zero-fan did not mean the item had been unstudied. Rather, it meant that the item did not belong to an intermediate category associated with the topic of the query. Anderson (1975), Anderson and Ross (1980), and Lewis and Anderson (1976) can be consulted for zero-fan manipulations of the same character as this experiment.

Table 6
Designs Used to Construct Targets and Foils

No. of facts using subject term	No. of facts studied using same predicate		
	0	1	2
	Targets		
1	—	aA	cE
	—	bB	dF
2	—	eC	eE
	—	fD	fF
	Mispairings		
0	gG	gA	gF
	hH	hD	hE
1	aG	aC	cF
	dH	bD	dE
2	eG	eA	eF
	fH	fB	fE

for an introductory course or earn $2.50 per hour. The experiment lasted between 2 and 3 hr.

Materials. The to-be-learned facts were identical in structure to those used in Experiment 1 and involved eight people facts, eight container facts, and eight event facts. The subjects and predicates were taken from the set given in the Appendix. As in Experiment 1, these facts were created according to an abstract target design. This design is given in Table 6, organized according to subject and predicate fan. The mispaired sentences are also illustrated in the table. For these it is possible to have 0-fan subject terms or predicates which means that those terms were not used in any studied facts. Thus, there were nine (3 × 3) foil conditions. The additional conditions necessitated creating two new 0-fan predicates (*G* and *H*) and two new 0-fan subject-terms (*g* and *h*). Because this design was replicated for people, for containers, and for events, there were altogether six 0-fan predicates and six 0-fan subject terms required.

In addition, nonsensical sentences were created from the mispairing design so that they could also have a 3 × 3 manipulation. This involved swapping subject and predicate terms from the different materials categories, as is illustrated in Table 3 for Experiment 1.

Procedure. The procedure for conducting this experiment was similar to Experiment 1. It was run on the same computer system and the first phase required subjects to study 24 sentences, as in Experiment 1. After this, subjects went into the test-study phase of the experiment which proceeded as in Experiment 1 except that this time, subjects were also exposed to the 0-fan topics and predicates: In each pass through the dropout, they saw all 12 0-fan terms and were asked to rate them on a 6-point pleasantness scale. This was the manner in which we attempted to control for effects attributable to basic exposure, so that we could separate the effects on sensibility judgments which were due to learning the material from the effects due to familiarity (see Anderson & Ross, 1980).

The RT phase involved the same alternations among two blocks of sensibility and two blocks of recognition judgments. A recognition block consisted of 96 studied sentences (each studied sentence presented 4 times), 54 sensible foils (each cell of the 3 × 3 design with 6 observations), and the complementary 54 nonsensical facts. A sensibility block consisted of 48 studied sentences, 54 sensible mispairings, and 108 nonsensical mispairings. Each major test block had two 30-s rest breaks and one longer break whose duration was determined by the subject. The administration of the reaction time testing was identical to Experiment 1.

Results

The results of the experiment are presented in Table 7, which uses the same organization as Table 4. The same procedures for calculating mean RTs were used in this experiment as in the previous. An analysis of variance was performed to obtain an estimate of the standard error of the means from the Subject × Condition interaction. For the recognition judgments the standard error of the RTs was 32 ms and the standard error of percentage of error was 2.2%. For the sensibility judgments the standard error of the RTs was 19 ms and the standard error of the percentage of error was 1.5%. The degrees of freedom in all cases were 399. (There were 20 subjects and 22 conditions, hence 19 × 21 = 399).

There are 24 conditions in common between Experiments 1 and 2. The correlation between RTs in these 24 conditions is quite high ($r = .94$), which suggests that the two experiments come to similar conclusions. Again, subjects were more accurate and faster in the case of making sensibility than recognition judgments: 9.2% errors versus 11.6% errors, $t(798) = 4.23$, $p < .01$; 920 ms versus 1005 ms, $t(798) = 4.23$, $p < .01$. We will focus on the reaction time data which is of principle theoretical interest. The error rates and RTs are positively correlated ($r = .92$ for recognition data and $r = .33$ for sensibility data).

As before, the recognition RTs are fastest for nonsensical mispairings, next fastest for studied sensible statements, and slowest for sensible mispairings. Of principle theoretical interest is the highly significant difference between nonsensical mispairings and sensible mispairings, $t(399) = 10.54$, $p < .01$, which supports Prediction 1 that subjects can reject sentences faster when they have two ways to do so. For sensibility judgments, the studied sensible statements are fastest. The two types of mispairings are essentially equally slow. There are four cells in common between sensible studied and sensible mispairings (those not involving zero fan). The average RT of these four cells is significantly faster in the studied sensible condition, $t(399) = 2.77$, $p < .01$. This again supports Prediction 1.

As before, there are no significant interactions between subject and predicate fan in the matrixes of Table 7, $F(9, 399) = .82$ for recognition data, and $F(9, 399) = 1.43$ for sensibility data. The fan effects can be defined as the differences between the 2-2 and 1-1 cells and, where applicable, the differences between the 1-1 and 0-0 cells. Table 5 summarizes these results. The fan effect for the 0-0 versus 1-1 contrast for recognition of sensible mispairings is highly significant, $t(399) = 10.42$, $p < .01$, which supports Prediction 2 and past results. The 1-1 versus 2-2 contrast is not significant for the studied sensibles or mispaired sensibles, although these effects are in the predicted direction. In total, we have again replicated positive fan effects for item recognition.

Also, we have replicated the negative fan effect for sensibility judgments. Consider the case of mispairings where the hypothesis is that subjects are usually making their sensibility judgments by judging sensibility (as opposed to recognizing statements). For sensible mispairings, both the 0-0 versus 1-1 contrast and the 1-1 versus 2-2 contrasts are significant, $t(399) = 3.87$, $p < .01$, and $t(399) = 2.57$, $p < .01$, respectively. For the nonsensical mispairings, the 0-0 versus 1-1 contrast is marginally significant in the wrong direction, $t(399) = 1.86$, $p < .10$, whereas the 1-1 versus 2-2 contrast is significant in the right direction, $t(399) = 2.12$, $p < .05$.

The fan effects for the different conditions do vary in size. The size variations are perfectly ordered in terms of the predictions generated by the detailed model presented after Experiment 1 (see Table 5). The odds of predicting by chance all 14 pair-wise orderings of the 1-1 versus 2-2 contrast is 2 out of 720. The odds of predicting all six pair-wise orderings of the 0-0 versus 1-1 contrast is 1 in 24.

Analysis of Error Times

Our theoretical explanation of the complex RT pattern led one reviewer to suggest that we analyze our error times as a converging measure of our theory. We decided to look at the error times for Experiment 2, which produced our most reliable data. Analyzing error RT data is more difficult than correct RT data because there are missing cells for many subjects (no errors in a condition means a missing cell). To deal with this problem, we defined a contrast for each subject that would estimate the overall size of the fan effect for a matrix,[7] where fan effect refers to increasing by one the number of subject and predicate associations.

The analysis given earlier for different size fan effects can be extended to make predictions for the error times. There are three sets of conditions: one in which we predict a positive fan effect for errors RTs, four in which we predict no clear fan effect, and one in which we predict negative fan effects. The logic used for making these predictions is as follows.

1. Positive fan effect. There should be a positive fan effect for error times for recognition judgments of studied sentences. Either process (sensibility or recognition) should give the correct answer; however, we expect a predominance of recognition judgments in this task (see our discussion of this process with respect to Table 1). Because both processes produce the correct answer, errors must be attributed to some spurious process that races against the correct response and beats it (King & Anderson, 1976; Mohs, Westcourt, & Atkinson, 1975). When a uniform error process races against a correct process that varies across conditions, the mean of the error times across conditions correlates with the mean of the correct times. This is because ~~in~~ conditions of long correct times, there is more room for long error times. Thus, we predict a positive fan effect for error times to mirror the positive fan for correct times. The average fan effect for this condition was 182 ms, which was significant, $t(19) = 2.11$, $p < .05$.

2. No clear prediction. Two of the conditions in which we predict no fan effect for correct RTs are sensibility judgments of studied sentences and recognition judgments of nonsensical mispaired sentences. Here we expect an equal mixture of recognition and sensibility judgments, both of which yield correct responses. Because the sensibility judgments produce a negative fan effect and the recognition judgments produce a positive fan effect, we expect a flat fan effect for correct judgments. Using the race analysis given earlier for positive fan effects, we expect no clear fan effect for errors in these conditions. In fact, the two fan effects were 65 ms for sensibility judgments of studied sentences and 60 ms for recognition judgments of nonsense sentences. Both are positive fan effects; the first is not significant, $t(19) = .97$, but the second is significant, $t(19) = 2.42$; $p < .05$. We will return to a discussion of this second effect at the end of the section.

The third condition for which we have no clear prediction is sensibility judgments of mispaired sensible sentences. Here, errors can occur two ways: either because an error process races against the correct sensibility process producing a negative fan effect, or because subjects make their judgments on the basis of a spurious recognition judgment that produces a positive fan effect. The fan effect was a positive 33 ms, which was not significant, $t(19) = .86$.

There is a fourth condition for which we predict no clear fan effect: Recognition judgments of sensible sentences should produce no fan effect if we use an analysis similar to the preceding one. The average fan effect was a negative 2 ms, which is clearly not significant.

3. Negative fan effect. Finally, we predict negative fan effects for sensibility judgments of nonsensical mispaired sentences. This is like the prediction for the positive fan effect: The only way an error can occur is for an error process to overcome a correct process, which has a negative fan effect. In fact, this condition produced the largest negative fan effect for errors, −10 ms, but it fell far short of significance, $t(19) = .72$.

Given the varied character of the error data, it is not surprising that so few of the fan effects are significant. Our analysis predicts nine pair-wise orderings of the six conditions, and all of these predictions are confirmed. Only 24 of the possible 720 permutations of these six conditions would preserve all nine pair-wise orderings, which is support for our analysis. It is also worth noting that the predictions of fan effects for the error RTs are not the same as those for correct RTs. This is because recognition mispairings switch from producing the strongest positive fan effect in the correct RTs to a slightly negative fan effect in the errors, and sensibility judgments of sensible mispairings switch from producing the strongest negative fan effects in the corrects to mildly positive fan effects for the errors. These reversals were predictions of the analyses.

---

[7] For the 2 × 2 tables or matrices, the fan effect can be estimated by a contrast with the following values:

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

For the 3 × 3 matrices, the fan effect can be estimated by the contrast with the following values:

$$\begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

To deal with unequal cell frequencies and, most significantly, empty cell frequencies, the actual weights w'_i we used were derived from the weights w_i above and cell frequencies n_i by the following formula:

$$w'_i = \frac{n_i(w_i - K)}{\sqrt{\Sigma n_i^2 (w_i - K)^2}}, \quad \text{where} \quad K = \frac{\Sigma n_i w_i}{\Sigma n_i}. \tag{1}$$

This gives a set of weights w'_i such that $\Sigma w'_i = 0$ and $\Sigma w'^2_i = 1$. For a particular matrix of error times T_i we then calculated the contrast $\Sigma w'_i T_i$. This gives us an estimate of $1/\sqrt{2}$ times the fan effect in the case of the 2 × 2 matrix and an estimate of $\sqrt{3}$ times the fan effect in the case of the 3 × 3 matrix. Therefore, in the numbers reported in the article we multiply the contrast by $\sqrt{2}$ for the case of the 2 × 2 matrix and divide it by $\sqrt{3}$ for the case of the 3 × 3 matrix.

Table 7
Results from Experiment 2: Mean Reaction Times (in Milliseconds) for Correct Judgments and Proportion Errors

	Predicate fan							
	Recognition judgment				Sensibility judgment			
Subject fan	0	1	2	*M*	0	1	2	*M*
				Studied, sensible				
1								
RT	—	944	973	959	—	896	844	870
E	—	.084	.091	.088	—	.048	.052	.050
2								
RT	—	977	982	980	—	881	855	868
E	—	.096	.101	.099	—	.065	.079	.072
M								
RT	—	961	978	970	—	889	850	869
E	—	.090	.096	.093	—	.057	.066	.061
				Mispaired, sensible				
0								
RT	912	1,054	1,098	1,021	1,031	925	936	964
E	.017	.104	.188	.103	.204	.096	.108	.136
1								
RT	1,000	1,192	1,208	1,133	958	927	890	925
E	.079	.254	.442	.258	.142	.096	.071	.103
2								
RT	1,004	1,162	1,232	1,133	956	950	858	921
E	.092	.242	.437	.257	.142	.096	.087	.108
M								
RT	972	1,136	1,179	1,096	982	934	895	937
E	.063	.200	.356	.206	.163	.096	.089	.116
				Mispaired, nonsensical				
0								
RT	903	895	921	906	914	943	933	930
E	.008	.029	.033	.023	.064	.064	.096	.075
1								
RT	904	997	919	940	947	964	930	947
E	.054	.058	.054	.055	.067	.112	.085	.088
2								
RT	897	996	999	964	932	972	907	937
E	.050	.038	.054	.047	.083	.112	.096	.097
M								
RT	901	963	946	937	931	960	923	938
E	.037	.042	.047	.042	.071	.096	.092	.087

On the other hand, the significant positive fan effect for recognition judgments of nonsense probes was not predicted. The error rate was lower in this condition (.042) than any other condition. One would have expected a very noisy pattern defined on such a small amount of data; however, of the 11 subjects who showed any fan effect, 10 showed a positive fan effect. We are uncertain as to why this reliable pattern of data was obtained on such a small fraction of the data.[8]

Experiment 3

This experiment was similar in design to the preceding experiments, except that subjects were asked to learn nonsensical sentences instead of sensible ones. In this respect, the experiment was unlike any that have been conducted in the past to explore the fan effect. We were interested in seeing whether we would obtain the same pattern of results, specifically, whether fan from nonsensical sentences would interfere with recognition judgments and facilitate sensibility judgments.

This was intended as a further test of our analysis of the negative fan effect for sensibility judgments. We hypothesized that facilitation from fan for sensibility judgments is due to practice at accessing the category of the subject and predicate so that they can be compared. Studying nonsensical sentences should

---

[8] Actually, Reder's (1982, 1987) model explains it by assuming subjects will select either the recognition or the sensibility process on a given trial when asked to make recognition judgments of nonsense. On those trials where subjects use the sensibility process, there is less chance for the error process to complete because the sensibility judgment process is faster than the recognition process. On those trials where subjects use the recognition judgment, responses are slower, which creates more chance for an error process to complete. Error processes racing against the recognition process give a positive fan effect for error.

also cause subjects to access the categories, but now they would mismatch leading to a recognition of the nonsensical character of the sentence. Therefore, if our hypothesis is correct, nonsensical sentences should facilitate sensibility judgments as much as sensible ones. We know of no alternative explanations for the negative fan effect for sensibility judgments, so it is unclear what another explanation would predict for Experiment 3. Nonetheless, an opposing prediction which seems plausible is that the more nonsensical sentences studied about a concept, the more difficult it will be to judge what is sensible concerning those concepts.

Method

Subject. A total of 25 subjects were recruited from the undergraduate population. They were paid $9 each for an experiment that lasted no more than 3 hr.

Design and materials. The structure of the materials for this experiment was identical to that for Experiment 2 except that the roles of sensible and nonsensical sentences were reversed. The study items were designed according to the layout in Table 6, except that sentences were created by selecting subjects and predicates from different categories. So, for instance, an item from the *a* subject category might represent a person, and an item from the *A* predicate category might represent an activity. There are nine possible combinations of subject types (*person, activity, container*) with predicate types. Six of these combinations are nonsensical. A total of 24 sentences were constructed so that there was one instance of all six possible nonsensical mispairings in each of the four cells of the 2 × 2 target design. With these target materials chosen, it was possible to create nonsensical mispairings to fill the 3 × 3 matrix of mispairings in Table 6. Sensible mispairings were created to fill the 3 × 3 matrix in Table 6 also.

Procedure. The procedure for running the experiment was identical to that for Experiment 2. Subjects were simply instructed at study to learn the nonsense sentences. The reaction time phase alternated between recognition blocks and sensibility blocks as in Experiment 2. The recognition blocks consisted of 96 studied nonsense sentences, 54 nonsensical re-pairings, and 54 sensible re-pairings. The sensibility blocks consisted of 48 studied nonsensical sentences, 54 nonsensical re-pairings, and 108 sensible re-pairings.

Results

The results of the experiment are presented in Table 8, which is organized in a fashion analogous to Table 7. Analyses of variance were performed to obtain an estimate of the standard error of these times from the subject-by-condition interaction. The standard error of estimate of the mean of the correct RT judgments in the recognition task was 29 ms. The corresponding value for error rates was 2.2%. In the sensibility task, the values were 26 ms and 1.6%, respectively. The degrees of freedom in all cases were 504.

Again subjects were faster and more accurate in the sensibility condition: 1,228 ms for recognition versus 1,194 ms for sensibility, $t(1008) = 3.85$, $p < .01$; 12.4% versus 9.1% errors, $t(1008) = 5.59$, $p < .01$. Again, the RT and error rate data are positively correlated ($r = .90$ for recognition data and $r = .73$ for sensibility data). We will continue to focus on the RT data. Again the interaction components (Subject Fan × Predicate Fan) in these matrixes are not significant, $F(9, 504) = 1.76$ for recognition data, $F(9, 504) = .72$ for sensibility data, and again we will focus on the fan effects as reflected by differences down the main diagonal and on differences among the matrixes.

Prediction 1, that subjects make judgments faster when there are two processes that can correctly apply, is confirmed in the two relevant comparisons. Sensible mispairings can be rejected in the recognition task either by failing to recognize the sentence or by judging it sensible. In contrast, nonsensical mispairings can be rejected only by the recognition process. Sensible mispairings are, in fact, rejected faster, $t(504) = 4.46$, $p < .01$. Similarly, studied nonsensical should be judged nonsensical more rapidly than mispaired nonsensical. This prediction is also confirmed when comparing the four cells they have in common, but the difference is not significant, $t(504) = 1.09$.

Prediction 2 is that there should be positive fan effects for recognition judgments for the studied nonsensical and the mispaired nonsensical. This prediction is confirmed by comparing the 2-2 versus 1-1 cells for the nonsensical studied, $t(504) = 3.56$, $p < .01$; by comparing the 2-2 versus 1-1 cells for nonsensical mispairings, $t(504) = 5.14$, $p < .01$; and by comparing the 1-1 versus 0-0 nonsensical mispairings, $t(504) = 2.85$, $p < .01$.

Prediction 3 is that there should be negative fan effects for sensibility judgments for the mispaired nonsensical and the mispaired sensible. The 2-2 condition is significantly faster for the nonsensical mispairings, $t(504) = 1.99$, $p < .05$. All the other comparisons except the 0-0 versus 1-1 for the mispaired sensibles are in the predicted direction, but none are significant.

Table 5 presents a complete analysis of the predictions for fan effects in Experiment 3. In the Experiment 3 column of Table 5, note that the sensible and nonsensical labels are reversed (e.g., studied sensible is studied nonsense). All but two of the fan effects are ordered as predicted. Of the 14 pair-wise orderings for the 1-1 versus 2-2 fan effect, just one is reversed. The probability of getting a confirmation this good by chance is only 10 out of 720. Again, there is just one misordering of the pair-wise orderings of the 0-0 versus 1-1 comparison. The probability of getting this by chance is 4 out of 24.

Conclusions

The three experiments have consistently found subjects faster when they can make their judgments either on the basis of sensibility or recognition processes. This confirms Prediction 1. Particularly important is the result that sensibility judgments are faster for studied sentences. This implies that sentences can be recognized from long-term memory before being judged sensible. As discussed in the introduction, this result rules out certain conceptions about the relation between memory and language processes.

Positive fan effects for recognition judgments were consistently found throughout the experiments when predicted, although they were not always significant. The fan effects (1-1 versus 2-2) in Table 5 for mispaired sensibles and studied sensibles, averaged over the three experiments, are quite significant, $t(1112) = 7.01$, $p < .01$, and $t(1112) = 4.53$, $p < .01$, respectively. This confirms Prediction 2.

Negative fan effects were usually found for sensibility judgments when predicted and usually were significant. Collapsing over experiments, the average fan effect for the 1-1 versus 2-2 comparison is also significant (see Table 5) for the two cases where we would predict it: for nonsensical mispairings, $t(1112) = 3.40$, $p < .01$, and for sensible mispairings, $t(1112) = 5.18$, $p < .01$. This confirms Prediction 3.

Table 8
Results from Experiment 3: Mean Reaction Times (in Milliseconds) for Correct Judgments and Proportion Errors

	Predicate fan							
	Recognition judgment				Sensibility judgment			
Subject fan	0	1	2	*M*	0	1	2	*M*
				Studied, nonsensical				
1								
RT	—	1,118	1,177	1,148	—	1,200	1,195	1,198
E	—	.084	.130	.107	—	.110	.088	.099
2								
RT	—	1,201	1,264	1,233	—	1,212	1,222	1,217
E	—	.123	.159	.141	—	.140	.102	.121
M								
RT	—	1,160	1,221	1,190	—	1,206	1,209	1,207
E	—	.104	.145	.124	—	.125	.095	.110
				Mispaired, nonsensical				
0								
RT	1,111	1,218	1,231	1187	1,295	1,244	1,280	1,273
E	.040	.047	.103	.063	.094	.120	.127	.114
1								
RT	1,184	1,228	1,419	1,277	1,233	1,267	1,215	1,238
E	.090	.130	.323	.181	.097	.097	.113	.102
2								
RT	1,189	1,363	1,439	1,330	1,212	1,210	1,223	1,215
E	.063	.223	.347	.211	.090	.120	.110	.107
M								
RT	1,161	1,270	1,363	1,265	1,247	1,240	1,239	1,242
E	.064	.133	.258	.152	.094	.112	.117	.108
				Mispaired, sensical				
0								
RT	1,042	1,130	1,162	1,111	1,141	1,124	1,130	1,132
E	.047	.073	.063	.061	.055	.060	.053	.056
1								
RT	1,150	1,257	1,322	1,243	1,153	1,172	1,151	1,159
E	.040	.130	.190	.120	.068	.082	.045	.065
2								
RT	1,193	1,243	1,340	1,259	1,099	1,156	1,127	1,127
E	.053	.133	.150	.105	.072	.085	.070	.076
M								
RT	1,128	1,210	1,275	1,204	1,131	1,151	1,136	1,139
E	.047	.105	.134	.095	.065	.076	.056	.066

Table 5 presents the detailed analysis of the ordering of the fan effects according to our elaborations of Predictions 2 and 3. The orderings were not always perfect for individual experiments, but the consistency across experiments is impressive. Averaged over the three experiments, the ordering is perfect. The joint probability of getting all 14 of the pair-wise predictions for the 1-1 versus 2-2 contrast and all six of the pair-wise predictions 0-0 versus 1-1 contrast by chance is less than 1 in 8,000.

The results of these experiments can be seen as having been produced by two processes, a sentence recognition process and a sensibility judgment process. These two processes generate opposite reaction time and accuracy functions when plotted against fan (i.e., the number of other studied sentences sharing the subject term or predicate term). The recognition process, which involves retrieval from long-term memory, is interfered with by the study of additional sentences. The sensibility judgment process is facilitated by the study of additional sentences.

The result that recognition processes are slowed by knowing more facts about the concepts queried is hardly a new result, and we were not surprised to find this effect in conditions where only the recognition process could apply correctly. On the other hand, the finding that sensibility judgments are faster under high fan is a novel result. We explained it in terms of practice during study in retrieving the category of the subject and the predicate in order to compare them.

The effect of fan has once again been shown to be a function of the cognitive process being performed. The most striking cases of this in Table 5 are the sensible mispairings, which either showed large positive fan effects or large negative fan effects, depending on the judgment being performed. In general, tasks that require discrimination of a specific sentence show positive fan effects, and those that do not, show negative fan effects. In addition, this experiment has once again demonstrated again that people do not always execute the strategy that corresponds to the official task at hand (Reder, 1982, 1987; Reder & Ross, 1983; Reder and Wible, 1984), but sometimes use recognition for a sensibility task and vice-versa.

References

Anderson, J. R. (1975). Item-specific and relation-specific interference in sentence memory. *Journal of Experimental Psychology: Human Learning and Memory, 104,* 249–260.

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1983a). *The architecture of cognition.* Cambridge: Harvard University Press.

Anderson, J. R. (1983b). Retrieval of information from long-term memory. *Science, 220,* 25–30.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory.* Washington, DC: Hemisphere Press.

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

Anderson, J. R., & Ross, B. H. (1980). Evidence against a semantic-episodic distinction. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 441–478.

Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences, 3,* 1–61.

Clark, H. H. (1974). Semantics and comprehension. In R. A. Sebeok (Ed.), *Current trends in linguisitics* (Vol. 12). The Hague, The Netherlands: Mouton.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82,* 407–428.

Fodor, J. A. (1983). *The moduality of mind.* Cambridge: MIT Press.

Keenan, J. M., & Brown, P. (1982, November). *Fanning: When does it hurt and when does it help?* Paper presented at the meeting of the Psychonomic Society.

King, D. R. W., & Anderson, J. R. (1976). Long-term memory search: An intersecting activation process. *Journal of Verbal Learning and Verbal Behavior, 15,* 587–606.

Kintsch, W. (1974). *The representation of meaning in memory.* Hillsdale, NJ: Erlbaum.

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

McKoon, G., & Ratcliff, R. (1979). Priming in episodic and semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18,* 463–480.

Mohs. R. C., Westcort, K. T., & Atkinson, R. C. (1975). Search processes for associative structures in long-term memory. *Journal of Experimental Psychology: General, 128,* 103–121.

Muter, P. (1978). Recognition failure of recallable words in semantic memory. *Memory & Cognition, 6,* 9–12.

Myers, J. L., O'Brien, E. J., Balota, D. A., & Toyofuku, M. I. (1984). Memory search without interference: The rate of integration. *Cognitive Psychology, 16,* 217–242.

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

Reder, L. M. (1982). Plausibility judgment versus fact retrieval: Alternate strategies for sentence verification. *Psychological Review, 89,* 250–280.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19,* 90–138.

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

Reder, L. M., & Ross, B. H. (1981, August). Integrated knowledge in differential tasks: Positive and negative fan effects. *Proceedings of the third annual conference of the Cognitive Science Society.* Berkeley, CA.

Reder, L. M., & Ross, B. H. (1983). Integrated knowledge in different tasks: The role of retrieval strategy on fan effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 55–72.

Reder, L. R., & Wible, C. (1984). Strategy use in question-answering: Memory strength and task constraints on fan effects. *Memory & Cognition, 12,* 411–419.

Schank, R. C., & Birnbaum, L. (1984). Memory, meaning, and syntax. In T. Bever, J. Carroll, & L. Miller (Eds.), *Talking minds: The study of language in cognitive science* (pp. 209–251). Cambridge: MIT Press.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14,* 489–537.

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre- and post-lexical loci of contextual effects on word recognition. *Memory & Cognition, 12,* 315–328.

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

Townsend, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* Hillsdale, NJ: Erlbaum.

Tulving, E. (1983). *Elements of episodic memory.* London: Oxford University Press.

Wexler, K., & Culicover, P. (1980). *Formal principles of language acquisition.* Cambridge: MIT Press.

Appendix

Examples of Subjects and Predicates

Subjects	Predicates
	People
The engineer	had mumps when he was eleven.
The farmer	wrote a short autobiography.
The doctor	learned French in high school.
The soldier	took up cross-country skiing.
The plumber	was given a speeding ticket.
The banker	married while in Puerto Rico.
The fireman	had two daughters and a son.
The lawyer	fell and broke his ankle.
	Events
The concert	was started with a speech.
The festival	was held in August every year.
The dance	was organized by older men.
The fair	was ended by a thunderstorm.
The rally	had been planned for months.
The rodeo	was postponed until Thursday.
The picnic	was the event of the year.
The game	was cancelled by the sponsor.
	Containers
The crate	was sent via Railway Express.
The carton	burst open when it was dropped.
The trunk	was damaged when sent by sea.
The package	had only a pineapple in it.
The basket	was filled with ornaments.
The box	was lighter than it looked.
The suitcase	contained a large white shirt.
The barrel	had the owner's address on it.

Received May 3, 1985
Revision received July 18, 1986 ■