COGNITIVE PSYCHOLOGY 16, 217–242 (1984)

Memory Search without Interference: The Role of Integration

JEROME L. MYERS, EDWARD J. O'BRIEN, DAVID A. BALOTA, AND MARIA L. TOYOFUKU

University of Massachusetts, Amherst

Several researchers have shown that the time required to retrieve a sentence increases with the number of unrelated facts learned about concepts in that sentence. L. M. Reder and J. R. Anderson (1980, *Cognitive Psychology*, **12**, 447–472) have argued that such *fan effects* also occur when the facts are integrated, provided that subjects must carry out a search of memory. In the present set of three experiments, we followed Reder and Anderson's procedure but, in a high-integration condition, used facts that were causally linked. In the first experiment, recall and recognition accuracy were better when fan was six than when it was three, and this effect was more pronounced in a high- than in a low-integration condition. In the second experiment, the overall fan effect was negative for recognition time in the high-integration condition, whereas in the low-integration condition, we obtained the usual positive fan effect. In the third experiment, subjects learned the materials on their own to provide a better opportunity for them to integrate facts. All fan effects became smaller or more negative relative to those observed in the preceding study. We consider a class of models for the findings in which subjects use confirming and disconfirming evidence as a basis for early termination of search.

In several current theories (Anderson, 1976; Collins & Loftus, 1975), memory structure is represented as a network of concepts connected by labeled relations. Each relation is assumed to have an associated value of strength which increases with the use of its corresponding path. When one retrieves information from this memory structure, activation spreads in parallel along these associative pathways. The rate at which this activation spreads along a given path is a function of the strength of that path relative to the sum of strengths of all paths emanating from the same concept. Therefore, the more information associated with a concept, the more time required to retrieve (for activation to spread to) a particular association.

This research was supported by NSF Grant BNS80-01180 to J. L. Myers. Support for D. A. Balota was provided by Training Grant NIMH (5-25007). We thank James Chumbley and Charles E. Clifton, Jr. for their comments on an earlier version of this manuscript. Requests for reprints should be addressed to Jerome L. Myers, Department of Psychology, University of Massachusetts, Amherst, MA 01003.

0010-0285/84 $7.50

A number of studies (e.g., Anderson, 1974; Reder & Anderson, 1980) support the above prediction, which is commonly referred to as the "fan effect." Nevertheless, this finding is quite counterintuitive. It seems more reasonable that the more people have stored about some topic, the better they are able to retrieve any fact related to it. Further, some theorists (e.g., Smith, 1981) have proposed memory structures and processes which would be consistent with the absence of a fan effect. Thus, we must establish the boundary conditions that exist for the phenomenon that increasing the number of facts learned about a concept interferes with the speed of retrieval. We designed the present research to contribute to that goal.

PREVIOUS STUDIES OF THE FAN EFFECT

In the typical fan experiment, subjects are presented to-be-remembered lists of sentences. Later, subjects receive a yes–no recognition test in which response latency is the major dependent variable. The principal finding is that subjects are slower to recognize a sentence that contains a given concept as the number of sentences that share the concept is increased. For example, if the probe sentence was "A hippie is in the park," recognition time increases as a function of the number of other sentences in the learned set that also contained "hippie" or "park" (Anderson, 1976).

Several investigators suggest that such interference results hold only when the sentences sharing a concept are thematically unrelated. To address the issue of thematic relatedness, Smith, Adams, and Schorr (1978) constructed integrated texts using sets of facts about an individual that were related by general knowledge; e.g., (1) *The banker was invited to christen the ship,* (2) *The banker broke the bottle,* (3) *The banker did not delay the trip*. Interestingly, there was no effect of fan upon the time to recognize sentences from such integrated passages. Moeser (1979) has reported similar results.

Reder and Anderson (1980), however, have pointed out that the absence of a fan effect does not necessarily indicate that *memory search* with integrated materials is unaffected by fan. Rather, subjects in the Moeser (1979) and Smith et al. (1978) studies may have responded correctly without searching memory. For example, in two of the Smith et al. experiments, a subject who had learned that the banker participated in a ship christening could have decided on the truth of a probe about the banker by noticing whether that probe had anything to do with ship christening; the subject could have conducted what Reder and Anderson referred to as a "thematic consistency check." There was no need to search the set of facts learned about the banker and, therefore, no fan effect. In a further experiment, Smith et al. (1978) attempted to control this possibility by using related foils, false probes which were related to the theme learned for the propositional subject. For example, if an old statement was "The banker broke the bottle," a foil might be "The banker broke the champagne bottle." However, as Reder and Anderson have noted, memory search could still be avoided; the subject could reject the probe because the word "champagne" had not been presented within the experimental session.

To test their explanation of the Moeser (1979) and Smith et al. (1978) results, Reder and Anderson required subjects to learn materials in which two different individuals were associated with the same theme, e.g., several sentences about Bill at the circus and several other sentences about Steve at the circus. In half of the test blocks unrelated foils were used, thereby allowing subjects to base their decisions on a thematic consistency check. In these unrelated foil blocks, no fan effects were observed. In the remaining blocks, related foils were used to increase the likelihood of memory search. For example, a predicate associated with Steve would be presented with Bill as the subject. In this condition, the usual fan effects were found.

Reder and Anderson's results appear to indicate that, whenever subjects must search memory, the duration of the search is a function of the number of associations learned about the concepts in the probe. Some counterevidence, however, may be found in a recent study by Whitlow, Smith, and Medin (1982). They employed propositions which were episodically correlated; for example, all propositional subjects who "moved the bucket" also "pushed the wagon" and "mowed the lawn." In a series of experiments, such correlated materials produced either no fan effect or a negative fan effect; by negative, we mean that recognition was actually faster when fan was greater. Since it is unlikely that a "theme" could be formed for such sentences, the thematic consistency check is not applicable. Thus, Whitlow et al. provided one condition in which increasing the number of associations does not increase recognition time. They also reopen the possibility that materials that are highly correlated on the basis of world knowledge may not exhibit a fan effect. With this in mind, we should take another look at Reder and Anderson's materials.

Three points are important about Reder and Anderson's (1980) sentences. First, they used several passages that they described as "nonscripts" in which real-world knowledge does not constrain the order of action. For example, an individual at a circus could watch various events in any order. Second, some of the passages they characterized as scripts might not follow scripts familiar to college subjects; for example, it is not clear that their subjects new all actions involved in purchasing a car or baking a pie. Third, and most important, the sets of three sentences (fan was one or three) were sampled randomly for each subject from a base set of six. Thus, even if their scripts were well integrated, it is unlikely that the three actions sampled from each script would occur together in close temporal sequence in the real world.

Two further points should be noted about the relations among sentences used in fan studies. First, Black and Bern (1981) have shown that recall is more accurate when one sentence describes an action that "causes" the action of the next sentence than when the two sentences merely describe actions that are temporally sequenced. This result implies that integration is not an all-or-none phenomenon. Reder and Anderson might have observed different results if they had causally integrated all of their passages. The second point concerns the impact integration has upon memory structure. Smith (1981) and Whitlow et al. (1982) have proposed an interconnections model to account for their findings of reduced fan effects. The basic idea is that well-integrated materials are represented by associative networks in which predicates are linked by relational associations ("interconnections") which provide shortcuts in the memory search; activation spreads not only from the terms in the probe to predicates in memory, but from one predicate in memory directly to others that are connected to it by world knowledge. It is unlikely that Reder and Anderson's materials produced such representations. In sum, although thematically coherent, Reder and Anderson's sentences may not have been interrelated in a way necessary to provide an adequate test of the effects of fan with integrated materials.

In the present experiment, we met Reder and Anderson's (1980) objections to the Smith et al. (1978) and Moeser (1979) studies by associating each theme with two propositional subjects and using related foils in the test phase. A high-integration condition employed target sentences that described actions which were enabled or caused by the actions of the preceding sentence. Our control condition was quite conservative; rather than the usual unintegrated set of sentences, we used a low-integration condition in which the predicates were thematically coherent, must as in Reder and Anderson's integration condition. A further difference from previous studies was that we used more predicates than has been typical in previous fan experiments; we pitted fan of three against fan of six. Larger sets should provide more interconnections among propositions because they can produce a more story-like quality. One other difference between our study and that of Reder and Anderson was that our subjects learned only 36 sentences, whereas Reder and Anderson's subject's learned 64 sentences. Reder and Anderson's subjects apparently had great difficulty learning the materials; they report that the sessions in their first experiment lasted 4 to 7 hr. It seems possible that such marathon sessions may attenuate comprehension and integration.

EXPERIMENT 1

In Experiment 1, we try to establish whether our high- and low-integrated passages differed with respect to free recall and recognition memory. Because Black and Bern's (1981) results show causally related materials improve free recall, we expect a similar effect with our sentence. If this did not occur, a subsequent investigation of retrieval times would be questionable.

A second reason for carrying out a study of memory accuracy is to extend the work of Black and Bern. Our study differs from theirs in two major ways. First, we vary the amount of material learned about each propositional subject, thereby permitting us to assess the effect of fan. Second, we include a measure of recognition accuracy; Black and Bower (1979) found that organization of prose affected recall but not recognition, and we were interested in whether this would hold for varying levels of integration.

Method

Subjects. The subjects were 48 University of Massachusetts undergraduates who received course credit for participating in psychological experiments.

Materials. High- and low-integrated lists are presented in Table 1. We designed them with the following characteristics. First, each list consisted of eight sets of six sentences. Each set of sentences described activities of a professional (e.g., the banker, the plumber) in a particular setting (e.g., a ball park or restaurant). We assumed that the setting provided the *theme* of the set.

Second, two sets of sentences described activities in each setting; for example, there were two professionals who were at a restaurant. The two sets of six sentences describing the activities at the restaurant were different. Third, the fourth and sixth predicates in each set of sentences appeared twice in the entire list of 36 base sentences. For example, the predicate *wait in line* occurs in both a ball park set and a restaurant set. Anderson (1976) has advocated pairing every predicate with two persons so that the effect of subject fan might be observed. Because it is impossible to tie up all our predicates in this way, we have limited analyses in all experiments to the fourth and sixth predicates of each set (the *target* sentence).

The high- and low-integrated base lists differed only in the third and fifth sentences of each set.For the high-integrated condition, these antecedent sentences caused or enabled the action of the following target sentence; for the low-integrated group, the antecedent sentences were thematically coherent but not as closely related to the following target. Both versions of Sentences 3 and 5 appear in Table 1.

We created two versions of the material for each integration condition. In Version 1, the first, third, and fifth sentences were dropped from the four sets of sentences describing actions at the restaurant and the ball game. In Version 2, those three sentences were dropped from the four sets describing actions at the race track and the bar. Thus, each target sentence appeared in the fan-6 condition for half of the subjects and in the fan-3 condition for the remaining subjects.

In summary, four lists of materials were derived from the base lists presented in Table 1. Each list consisted of 36 sentences, four sets of fan-6 and four sets of fan-3. Within each

TABLE 1
The Base Set of Stimulus Materials[a]

The actor

had a ticket for the Red Sox game.
went to the ball game.
sat down as the umpire yelled play ball./bought a hot dog from a vendor.
saw the start of the ball game.
found the first few innings boring./looked at his program.
went home early.

The banker

decided to see a baseball game.
arrived at the ball park.
found a crowd buying tickets./bought a souvenir pennant.
waited in line.
entered in time to see his team score./sat near the first base dugout.
cheered loudly.

The lawyer

wanted a good meal.
went to a new restaurant.
found all the tables filled./saw there was a salad bar.
waited in line.
asked about the specialties./spoke to the hostess.
made a selection.

The teacher

took a table near a window.
read the menu.
found that the prices were high./asked for a glass of water.
checked his wallet.
decided he could afford a steak./munched on a breadstick.
placed an order.

The artist

was in the corner saloon.
chatted with the bartender.
asked him to turn on the TV./saw his neighbor enter.
saw the start of the ballgame.
decided he'd better buy something./said hello to a friend.
bought a beer.

The plumber

entered the barroom.
sat down at the bar.
thought he'd like a martini./listened to a folk singer.
placed his order.
decided to have only one drink./watched some people playing darts.
went home early.

The doctor

liked to bet on horses.
went to the racetrack.
studied the odds./stood at the rail.

TABLE 1—*Continued*

made a selection.
debated betting ten dollars./watched the horses at the gate.
checked his wallet.

The fireman

entered the stands at the racetrack.
watched the horses race.
saw his choice win./looked for a seat.
cheered loudly.
felt his throat become hoarse./tore up his ticket.
bought a beer.

[a] For the third and fifth lines of each passage, the first predicate presented belongs to the high-integrated and the second to the low-integrated condition.

level of fan, two individuals appeared in each theme or setting to discourage the theme checking strategy described by Reder and Anderson (1980; see also Reder, 1982). The four lists differed with respect to assignment of themes to fan conditions (the version variable) and with respect to the degree of correlation of target sentences to their predecessors in fan-6 conditions (the integration variable). Note that the sets of threes were identical in the two integration conditions.

Procedure. We tested subjects in groups of two to eight. Each subject was presented with a booklet containing the eight sets of sentences, one set per page; blank pages separated the eight study pages. There were four counterbalanced orders of sets. Subjects were told when to start and stop studying each set; following Reder and Anderson's (1980) procedure of allowing $10 + 10n$ sec (n is the fan), we allowed 40 sec for each set of three and 70 sec for each set of six. After the study phase, subjects solved mathematics problems from the Scholastic Aptitude Test for 15 min. A recall phase followed in which subjects wrote whatever they could recall on eight sheets labeled by profession and setting (e.g., "banker at the ballgame"). They had 15 min to complete this phase. Following recall, we presented 72 sentences in random order. Thirty-six of these were the original acquisition list and the remaining thirty-six were constructed by pairing each professional with a predicate that had been true of the other professional tied to the same theme. Subjects had to indicate whether or not the sentence had been presented before and to indicate their confidence in their response using a 3-point scale. Subjects were allowed as much time as they required to complete the recognition test, and were instructed to proceed through the recognition test in the order the items appeared.

TABLE 2
Experiment 1: Proportion of Correct Recalls [P(R)], Mean Number of Target Intrusions (I), and Area under the ROC Curve Based on Ratings (Area)

Integration	Fan	P(R)	I	Area
High	6	.631	.750	.761
	3	.448	1.333	.659
Low	6	.479	1.167	.730
	3	.401	1.583	.649

Results

Recall accuracy. Table 2 presents the proportion of targets (Sentences 4 and 6 of each set in Table 1) correctly recalled and the number of target intrusions along with a measure of recognition accuracy (area under the ROC curve). First, we consider the proportion of correct target recalls. Table 2 reveals that the addition of three sentences resulted in an average increase in proportion recalled of .13, $F(1,32) = 16.13$, $p < .001$. This negative fan effect appears to be larger for the high- than for the low-integrated text. Although the interaction between text and fan is not significant ($p > .10$), supplementary analyses demonstrate a reliable relationship between text and fan. First, the negative fan effect is highly significant in the high-integrated condition ($F(1,16) = 14.76$, $p < .01$) but is not in the low-integrated condition ($F(1,16) = 3.13$, $p > .10$). Second, the targets in the set of sixes are far better recalled in the high- than in the low-integrated case $F(1,32) = 6.23$, $p < .02$; however, for the threes, in which there is no true integration distinction, the F ratio was less than 1.

The only other significant effect is a fan × version interaction, $F(1,32) = 5.81$, $p < .025$. For reasons that are unclear to us the advantage of fan-6 over fan-3 was greater in Version 2 than in Version 1. This occurred primarily in the low-integrated condition; the Fan × Version interaction was significant for those texts, $F(1,16) = 7.35$, $p < .02$; whereas for the high-integrated texts the corresponding F ratio was less than 1.

Recall intrusions. It is possible that the superior recall scores for the high-integrated sixes represent a form of sophisticated guessing. Scores in this condition may not reflect better memory for the subject–predicate associations but rather stronger preexperimental associations among the predicates. The retrieval of any one predicate triggers the retrieval of other predicates associated with that theme. If this is the case, those conditions showing the best target recall should also show the most target intrusions. This is, target predicates for one propositional subject should often be misrecalled with the other subject sharing the same theme.

These target intrusions are also presented in Table 2. Contrary to the model presented above, intrusions did not correlate positively with correct recall; in fact, the integrated sixes actually had the fewest intrusions. The fan-6 condition had significantly fewer intrusions than the fan-3 conditions; $F(1,32) = 8.11$, $p < .01$. This is particularly surprising because one might expect subjects to guess more to a condition from which more sentences were presented to them. The only other significant effect in this analysis was the interaction of integration, version, and fan, $F(1,32) = 6.82$, $p < .02$. This three-factor interaction reflects a lack of Version × Fan interaction for the high-integrated text ($F < 1$) but a significant Version × Fan interaction for the low-integrated text, $F(1,16) = 12.00$ $p < .01$. Again, the advantage of fan-6 was larger in Version 2 than in Version 1.

Recognition data. Although subjects responded to all 36 sentences in Table 1 and to 36 other sentences involving the same predicates, the data for analysis are the responses to the 16 targets (fourth and sixth sentences of each set in Table 1) and to 16 foils constructed from the target predicates. For example, "The actor went home early" was a target and "The actor cheered loudly" was a foil. Note that all false sentences, including the 16 critical foils, were constructed by pairing a studied predicate with the other propositional subject attached to the same theme.

Subjects had rated each of 72 items as old or new on a 3-point scale. We attached a minus sign to *new* responses and then added 3 to all ratings to convert to a scale between 0 and 6. Each rating was then converted to a hit or false alarm on a unit scale by dividing by 6, the maximum possible rating on the transformed scale. We used each pair of hit and false alarm probabilities to provide an estimate of the area under the ROC curve that would go through that point (see Green & Swets, 1966; Grier, 1971; Pollack & Norman, 1964). This measure does not depend on the choice between threshold and signal detection theories. These areas are presented in Table 2.

In general, the recognition data exhibit a pattern similar to that for recall. The most obvious effect is that of fan; the targets from the threes are less well remembered than those from the sixes, $F(1,32) = 11.54$, $p < .002$. The high-integrated sixes had a higher recognition score than the low-integrated sixes but the effect was not significant.

Discussion

A consistent pattern emerges from Experiment 1. First, subjects recall high-integrated sixes better than low-integrated sixes. This parallels Black and Bern's (1981) finding that causal relations between sentences yield better recall than do temporal relations. This effect holds even when discriminability among stories is difficult, as in our study where two stories are tied to each theme. Furthermore, our data suggest that causal relations also decrease recall intrusions and improve recognition memory. This pattern of data clearly cannot be explained in terms of a guessing strategy.

The second major finding is a significant negative fan effect in all three dependent measures. Target propositions embedded in a set of six highly integrated (causally related) sentences were better remembered than the same target propositions embedded in a set of three thematically coherent sentences. When the sixes were only temporally integrated (the low-integration condition), the fan effect depended upon the version of the materials. Nevertheless, in neither version was there a significant positive fan effect on any measure. This contrasts with results for completely unrelated sets of sentences (Bradshaw & Anderson, 1982; Moeser, 1979). It is consistent, however, with the results of Black and Bower (1979) who found better recall of a base set of four actions when they were embedded

The difference between Moeser's and Bradshaw and Anderson's results on the one hand, and Black and Bower's and our own on the other, is presumably due to the integrated materials in the latter two studies. In our recall experiment, the theme is explicit in the recall cue (e.g., "Banker–Baseball"), and subjects can presumably retrieve it from the sentence probe in recognition. As activation spreads from the theme to related predicates, it will encounter more of the predicates from the set of six than from the set of three. Furthermore, the more facts that are recognized as part of the memorized set, the more bases there are for activating, or reconstructing, other related facts. When the memorized facts are strongly related, as in the high-integration condition, there is an even greater likelihood that one retrieved fact will lead to another member of the memory set. Thus, because of the presence of a coherent theme, a higher proportion of sixes than threes will be remembered, and high-integrated sixes will be more likely to be retrieved than low-integrated sixes. Because the acquisition phase occurred only a short time before the test phase, subjects were able to recognize which persons and predicates were associated with each other; thus, high recall and hit rates were not accompanied by high intrusion and false alarms. If the tests followed acquisition by several days, a positive relation between these measures might be observed.

This model does not require that search time be faster for the fan-6 condition than for the fan-3 condition, nor does it require response time to increase with the degree of integration. Therefore, having established the validity of our integration manipulation, we now turn to an examination of response times.

EXPERIMENT 2

Method

Subjects. Forty-eight University of Massachusetts undergraduates served as subjects for course credit. They were assigned at random to one of eight experimental groups differing with respect to integration (high or low), version of the text (which themes were sixes and which threes; see the *Method* section of Experiment 1 for further details), and experimenter (two of the authors each ran half of the subjects in each group).

Materials. We used the same lists of 36 sentences and 72 probes that we had presented in Experiment 1 (see *Materials* section).

Procedure. Subjects were run individually in an experimental session which lasted 1 ½–2 hr. The experiment was controlled by a PDP 8/E computer, following Reder and Anderson's procedure (1980). Specifically, subjects first viewed on a TV monitor each of the eight sets of three or six sentences. Viewing time was either 40 or 70 sec, depending upon the size of the set of sentences. The order of presentation of the eight sets was randomized on each viewing cycle. After the first viewing cycle, subjects had to say the sentences in response to a cue (e.g., "the banker") on the screen. After subjects responded, the set of sentences was presented for further study. A set of sentences was dropped from

TABLE 3
Experiment 2: Mean Response Times in Milliseconds and Error Probabilities (in Parentheses)

Integration	Fan	Targets	Foils
High	6	2032 (.167)	2133 (.180)
	3	1998 (.143)	2253 (.231)
	Fan effect =	34 (.024)	−120 (−.051)
Low	6	2064 (.165)	2180 (.181)
	3	1967 (.141)	2128 (.225)
	Fan effect =	97 (.024)	52 (−.044)

the study list when the subjects correctly produced the sentences on two successive cycles. This continued until all eight sentence sets met the dropout criterion.

After the acquisition phase, subjects wrote the predicates on a sheet of paper under the appropriate propositional subject. Two subjects had more than two errors in this phase and were dropped from the experiment.

The test phase consisted of four cycles through the 72 original sentences and foils. The probes were randomly sequenced within each cycle, and the four cycle were separated by rest intervals of about 2 min. Subjects rested their index fingers on two triggers labeled "new" and "old" and pulled the appropriate trigger when a sentence appeared on the screen. Reaction time was measured from the appearance of the sentence. Subjects had 5 sec to respond. One second intervened between the response and a ready cue ("xxx"), except on trials on which they made an error. On such trials an additional 1-sec delay was added while the word "error" was presented for feedback. One-half second later, a sentence appeared on the screen.

Results

Cycles to acquisition. The high-integrated condition required an average of 3.94 (fan-6) and 3.89 (fan-3) cycles to meet the dropout criterion; the low-integrated condition had comparable means of 4.23 and 3.97. Neither the integration nor the fan effect approached statistical significance, nor did their interaction.

Recognition times. As in Experiment 1, the analyses are based only on the critical 16 targets and 16 foils. As before, these are the fourth and sixth predicates in each sentence set of Table 1. As noted earlier, all 36 foils were constructed by pairing a predicate for one person with the other person tied to the same theme.

Table 3 presents mean correct recognition times and errors. The fan effect, averaged over targets and foils, was actually negative for the high-integrated text and positive for the low-integrated text. This Fan × Integration interaction was significant, $F(1,40) = 8.17$, $p < .01$. Supplementary analyses carried out separately for the two integration conditions showed a significant positive fan effect (sixes slower than threes) for the low-integrated condition, $F(1,20) = 5.81$, $p < .05$. Although the average negative fan effect in the high-integrated condition was not significant ($p > .10$), it was for the foil data considered alone, $F(1,20) = 6.13$, $p < .025$.

Another major effect in Table 3 follows from this last observation. The overall data set exhibited a significant Fan × Probe Type interaction, $F(1,40) = 4.30$, $p < .05$. Analyses carried out separately for the two integration conditions suggests that this is largely due to the strong negative fan effect noted above for foils in the high-integrated condition. The Fan × Probe Type interaction for this group is marginally significant with $F(1,20) = 4.19$, $p < .054$; in the low-integrated condition, the corresponding F is only 1.12.

A third major effect was the faster response to targets than to foils, $F(1,40) = 39.93$, $p < .001$. The only other effects significant in these analyses involved the experimenter and appeared only in the low-integrated text condition; Experimenter × Fan and Experimenter × Fan × Probe interactions were significant at the .025 level.

Error rates. In general, the pattern of error rates paralleled that for response times. There were fewer errors in responses to targets than to foils, $F(1,40) = 11.50$, $p < .01$. There was also a Fan × Probe interaction; as with response times, a positive fan effect was obtained with targets, but a larger negative fan effect was obtained with foils, $F(1,40) = 7.83$, $p < .01$.

Repetition statistics. One explanation of the pattern of fan effects might be based upon the fact that a probe of a fan-6 passage is more likely than one from a fan-3 passage to have been preceded by a probe of the same theme. If one probe primes the next when they share their theme, then the fan-6 condition will include a larger proportion of primed responses than will the fan-3 condition. The implication is that the fan effect should depend on whether a related probe appeared on the preceding trial. More specifically, the usual positive fan effects should be observed on those trials not recently preceded by related probes. Furthermore, one might expect this priming effect to be larger for the high-integrated text than the low-integrated text.

To test this priming hypothesis, we distinguished four categories of response times: (1) *repetition, same subject*—the probe was immediately preceded by another probe about the same propositional subject; (2) *repetition, different subject*—the probe was immediately preceded by a probe whose subject was the other individual tied to the same theme; (3) *nonrepetition, same subject*—the most recent occurrence of a probe about the current theme was more than one trial back but had the same propositional subject; (4) *nonrepetition, different subject*—the most recent occurrence of a probe testing the current theme was more than one trial back, and its subject was the other individual tied to that theme. The results of this conditional analysis are presented in Table 4.

TABLE 4
Experiment 2: Mean Repetition and Nonrepetition Response Times in Milliseconds and Error Probabilities (in Parentheses)[a]

Propositional subject	Integration	Fan	Repetitions: Targets	Repetitions: Foils	Nonrepetitions: Targets	Nonrepetitions: Foils
Same	High	6	1815 (.183)	2062 (.167)	2014 (.167)	2038 (.222)
		3	1794 (.163)	1855 (.182)	1984 (.161)	2192 (.225)
	Fan effect =		21 (.020)	207 (−.015)	30 (.006)	−154 (−.003)
	Low	6	1758 (.100)	1973 (.145)	2074 (.161)	2200 (.208)
		3	1916 (.167)	1741 (.143)	1936 (.155)	2220 (.259)
	Fan effect =		−158 (−.067)	232 (.002)	38 (.006)	−20 (−.051)
Different	High	6	2064 (.167)	2126 (.089)	2035 (.173)	2134 (.178)
		3	2052 (.102)	1915 (.194)	1996 (.119)	2279 (.225)
	Fan effect =		12 (.065)	211 (−.105)	39 (.054)	−145 (−.047)
	Low	6	1989 (.186)	2132 (.192)	2109 (.138)	2200 (.225)
		3	1871 (.098)	1972 (.123)	1996 (.138)	2033 (.157)
	Fan effect =		118 (.088)	60 (.069)	113 (.000)	167 (.068)

[a] See text for explanation of *same*, *different*, *repetition*, and *nonrepetition*.

Since the ordering of trials within a block was randomized, the statistics in Table 4 are based on observations which varied in number across subjects and conditions. Furthermore, some subjects contributed no observations to some conditions. In view of this, no significance test seemed appropriate; nevertheless, some clear trends emerged from the data.

First, there were priming effects. Response times based on probes that repeat the propositional subject were 218 msec faster than those based on nonrepetitions. The effect was weaker, but still present, for probes sharing the theme (but not the subject) of the previous trial; that priming effect was 83 msec. Second, the priming explanation of the fan effects is not supported; the nonrepetitions for the high-integrated groups exhibited a small fan effect for targets (relative to the low-integrated groups) and a strong negative fan effect for foils. Thus, the patterns in the overall mean times were not due to a disproportionate effect of repetitions upon the fan-6 data. Third, with one exception (the low-integrated text, same propositional subject), repetitions exhibited positive fan effects. This last result may reflect the search of a short-term buffer which contains propositions or themes activated by the preceding probe. The important point to note, however, is that the *non*repetition data should be the purest indication we have of long-term memory structure and processes and these data clearly demonstrate the influence of integration on the observed fan effect.

The apparent difference between repetition and nonrepetition statistics suggests a problem in comparing results from different studies. If the effect is reliable, studies with different lags between related probes will contribute unequal proportions of two processes (priming vs search) to the marginal means that are usually reported. Discrepancies among studies might reflect such disparities. For this reason, it seems wise to provide sequential statistics. Of course, they are of interest in their own right as indicators of underlying probabilistic processes.

Discussion

The results of Experiment 2 indicate that causally integrated sentences produce different results from those reported by Reder and Anderson (1980). Because the theme was tied to two subjects in the present study, Reder and Anderson's theme checking strategy is not applicable. We have also demonstrated the inadequacy of a priming explanation of our fan effects. There are, however, a number of other explanations including a mixed strategy model that incorporates both theme checking and memory search (Reder & Anderson, 1980; Reder & Ross, 1983).

Before considering these alternative models, we will present one further experiment. In this experiment, we investigate whether the way subjects learn the materials influences fan effects. It seemed plausible to us that asking subjects to learn large lists in a laboratory for extended periods may discourage them from finding meaningful relations among sentences; under such conditions, subjects may be more likely to attend to the surface structure of the materials. If subjects learn the sentences at their own pace, they may take greater advantage of the relations inherent in the text. This, in turn, should lead to smaller fan effects. In light of this possibility, we decided to conduct a small follow-up to Experiment 2, in which subjects were sent home to learn the materials on their own.

EXPERIMENT 3

Method

Subjects. Twenty-four University of Massachusetts students served as subjects for course credit. They were randomly assigned to the four combinations of integration and text version. The materials were identical to those in the preceding two experiments. The procedure differed from that of Experiment 2 only in that subjects obtained the to-be-learned sentences the day before the experiment. All sentences were typed on a single sheet and were grouped by subject terms. Each subject was instructed to learn the sentences in any order and any manner she/he felt most comfortable. Upon entering the laboratory for the test phase, they first were given a written test for recall of the predicates as in Experiment 2.

Results

Mean recognition times. Table 5 presents mean correct response times and error rates for Experiment 3. The results are in the direction we had expected. Response times in the high-integrated condition exhibited negative fan effects even on target trials; in the low-integrated condition, there was also a negative fan effect for foils, and the positive fan effect for targets was smaller than for the comparable group in the preceding experiment. Although the results suggest that the take-home acquisition procedure produced smaller, or more negative fan effects than had the paced procedure of Experiment 2, some caution should be observed in

TABLE 5
Experiment 3: Mean Response Times in Milliseconds and Error Probabilities (in Parentheses)

Integration	Fan	Targets	Foils
High	6	1680 (.109)	1901 (.198)
	3	1787 (.133)	2025 (.292)
	Fan effect =	−107 (−.024)	−124 (−.094)
Low	6	1835 (.128)	1975 (.232)
	3	1761 (.122)	2067 (.224)
	Fan effect =	74 (.006)	−92 (−.008)

viewing these results: A test of the Fan × Experiment interaction failed to yield significance ($p > .20$).

As shown in Table 5, responses to foils were again considerably slower than to targets, $F(1,20) = 81.99$, $p < .001$. The only other significant effects obtained in these analyses were a Version × Fan and a Version × Probe × Fan interaction; both were less than .02. Interestingly, similar interactions with version occurred in Experiments 1 and 2 as noted earlier. Subsequent analyses indicated that each of these interactions in all three experiments occurred only for the low-integrated materials. Version 2 consistently yielded smaller positive fan and larger negative fan effects than Version 1. We are unable to specify the critical variables that led to these interactions, but it does appear that when relations among propositions are not highly constrained, incidental factors may influence performance. Another case in point is the Version × Experimenter interaction that occurred in Experiment 2 for the low-integrated texts.

Although the Integration × Fan interaction is not significant, the magnitude of the effect is large. Furthermore, when the data are combined with those from Experiment 2, significant effects observed there receive still stronger support. An analysis based on the pooled data results in a highly significant Integration × Fan interaction with $F(1,60) = 12.68$, $p < .001$. The negative fan effect based on these combined data is clearly significant for the high-integrated text, $F(1,30) = 6.09$, $p < .025$; in contrast for the low-integrated text, the average of the significant positive fan effect of Experiment 2 and the negative fan effect of Experiment 3 yields a clearly nonsignificant F ratio of 1.50.

Error rates. As in Experiment 2, the error rates pattern parallels the response times: foils are significantly more difficult than targets ($F(1,20) = 20.45$, $p < .001$), there is a nonsignificant ($p > .15$) but negative fan effect for the high-integrated text, and no fan for the low-integrated text.

Repetition statistics. Table 6 presents repetition and nonrepetition results for Experiment 3. The magnitudes of the average priming effects are quite close to what they were in Experiment 2: 205 msec when conditioning on the same propositional subject and 64 msec when conditioning on the same theme (different subject). Again, the nonrepetitions exhibit negative fan for the high-integration condition; for the low-integration condition, we now observe small negative fan effects for the foils. The major difference from the results for Experiment 2 is that we now observe a mixture of positive and negative fan effects for repetition trials; in Experiment 2, repetitions consistently exhibited positive fan effects.

Discussion

Using a new procedure for acquisition, we have again demonstrated the absence of positive fan effects. In fact, the average fan effect is now

TABLE 6
Experiment 3: Mean Repetition and Nonrepetition Response Times in Milliseconds and Error Probabilities (in Parentheses)

Propositional subject	Integration	Fan	Repetitions: Targets	Repetitions: Foils	Nonrepetitions: Targets	Nonrepetitions: Foils
Same	High	6	1515 (.130)	1599 (.216)	1639 (.136)	1884 (.208)
		3	1532 (.133)	1684 (.269)	1863 (.134)	1967 (.268)
	Fan effect =		−20 (−.003)	−85 (−.053)	−224 (.002)	−83 (−.060)
	Low	6	1641 (.107)	1845 (.211)	1897 (.143)	1967 (.222)
		3	1713 (.107)	1838 (.080)	1780 (.121)	1960 (.180)
	Fan effect =		−72 (.000)	7 (.131)	117 (.022)	7 (.042)
Different	High	6	1640 (.096)	1732 (.148)	1712 (.093)	2008 (.213)
		3	1797 (.119)	2047 (.324)	1793 (.141)	2144 (.286)
	Fan effect =		−157 (−.023)	−315 (−.176)	−81 (−.048)	−136 (−.073)
	Low	6	1921 (.111)	1968 (.214)	1831 (.129)	2027 (.280)
		3	1806 (.133)	1930 (.250)	1778 (.128)	2058 (.279)
	Fan effect =		115 (−.022)	38 (−.036)	53 (.001)	−31 (.001)

negative, although not significantly so ($p > .20$). If we assume that, when given more time to consider the materials to be learned, subjects are able to take greater advantage of the relations in those materials, the results in this experiment make excellent sense. Under all conditions in Table 5, positive fan effects have been reduced (relative to those in Table 3) or reversed, and negative fan effects have become more negative.

GENERAL DISCUSSION

Reder and Anderson (1980; also Reder & Ross, 1983) have concluded that the time required to search memory increases with the number of facts learned about a propositional subject. The results of the present research suggest that this conclusion requires qualification. With low-integrated materials learned by a paced procedure in the laboratory, we also obtained significant fan effects. However, when those same materials were acquired at the subject's own pace at home, or when the sentence preceding the target was causally related to that target (high-integration materials), fan effects were greatly attenuated and even reversed. Reder and Ross (1983) have also obtained negative fan effects, but only in a condition in which foils were unrelated to the theme associated with the propositional subject, and the subject's task was to judge whether the probe was consistent with the theme. In contrast, none of our critical probes are easily judged for plausibility :"taking a seat" is consistent with any of our four themes.

A variation of the theme checking hypothesis would be one in which subjects organized each set of highly integrated propositions under several subthemes and used these, at least on some trials, as the basis for a consistency check. To account for our data, several conditions would have to be met. First, for reliable discrimination between targets and foils, each subnode must subsume the information in the propositions organized beneath it, but still be related to exactly one propositional subject. Given our materials and foils, this seems extremely unlikely. For example, a possible subtheme in the baseball texts might be the start of the game, but this is associated with both the banker and the actor. A second requirement of such a subtheme model is that to have negative fan effects, the subthemes must be accessed more quickly than the three propositions in the fan-3 condition. This is, not only must subjects establish reliably discriminative subthemes but they must activate them more quickly than some memorized propositions. Finally, the specific subnode model must account for greater negative fan effects on foil than on negative trials. In short, while a subtheme model may possibly be correct, the necessary conditions are unlikely to have been met in the present study.

A key to the development of an adequate theory for our data may be the persistent fan $\times$ probe interaction. In the case in which integration is presumably least, laboratory acquisition of low-integrated materials, the positive fan effect is about half as large with foils as with targets. In two conditions that should encourage integration, laboratory acquisition of high-integrated materials or home acquisition of low-integrated materials, targets exhibit a small positive fan effect and foils exhibit a negative fan effect of about 100 msec. Finally, when both the acquisition method and mateials combine to encourage integration, negative fan effects are found for both targets and foils; the effect is slightly more negative for the foils. This pattern is interesting because neither prior research nor simple search models would have led us to expect it. Typically, with unintegrated materials, fan effects have been similar in magnitude for targets and foils (Anderson, 1976). Simple search models which postulate a self-terminating search for targets and an exhaustive search for foils would predict the opposite of our result: fan effects should be smaller with self-terminating searches. In view of this, a model capable of describing our fan × probe interaction deserves serious consideration. Before presenting such a model, we consider several other possible explanations of the data.

Alternative Models

Priming. As we noted when we presented our repetition and nonrepetition statistics, there is a greater likelihood that a target is preceded by a probe from the same passage (or theme) in the fan-6 than in the fan-3 condition. This could account for negative fan if the first sentence primes the second. Further, if such priming effects are greater with high-integrated than low-integrated materials, the integration × fan interaction follows. But it is not clear how this model would account for the fan × probe interaction described above. More important, negative fan effects should be eliminated, or at least attenuated, when priming is minimal, as on nonrepetition trials. In fact, fan effects were more negative on these trials than on repetition trials in Experiment 2, and equally negative in Experiment 3.

Search from the predicate. Since the target predicates only had a fan of two, a search from those predicates might often have succeeded before a search from the subject. Because this is more likely to be the case when the subject fan is six than when it is three, negative fan effects could result. Although predicate search may influence performance, it fails to account for two important results. First, it does not explain why we have positive fan effects in the low-integrated condition of Experiment 2. Second, it is not clear why fan effects are always smaller in the high- than in the low-integrated condition. We would have expected the reverse: more dependence upon the predicates with less integrated material. Third, this predicate search explanation fails to deal with fan × probe interaction.

Strength in storage. Assume that the strength of a path from a propositional subject to its predicate is greater in the fan-6 than in the fan-3 condition. The better recognition of sixes than threes in Experiment 1 supports this assumption. Following Anderson (1976), the rate of activation of the paths would be:

$$r_6 = \frac{S_6}{K + 6S_6} \tag{1}$$

and

$$r_3 = \frac{S_3}{K + 3S_3} \tag{2}$$

where K is the strength of preexperimental associations to the propositional subject. If S_6 is sufficiently greater than S_3 (the exact difference required will decrease as K increases), the sixes would be searched more quickly than the threes.

There are two problems with this approach. First, it is not clear how the difference in strength arises. Our cycles-to-acquisition data (Experiment 2) showed little effect of either fan or integration, indicating that there were no overt differences in practice in that study. Furthermore, it is difficult to conceive of subjects spending more time studying high- than low-integrated materials, but it is the high-integrated text that shows the pronounced negative fan effect. Thus, we must assume that high-integrated sixes profit from implicit rehearsal presumably because of relations that exist among propositions prior to the experiment. Studying one proposition reminds the subject of another which is closely linked. However, such differences in strength stemming from the short laboratory acquisition phase may not be sufficient to account for the results in Experiment 2. Furthermore, if preexperimental links among predicates serve as the basis for implicit rehearsal, these links might well be stored as part of the episodic representation. Anderson (1976) has employed just such an elaborative encoding mechanism to account for several phenomena (see also Reder, 1982). If such links are stored, however, we no longer have a pure strength model; instead, we have the interconnections model proposed by Smith (1981; also Smith et al., 1978; Whitlow et al., 1982), modified by the assumption of differential strengths.

The second problem with the pure strength explanation is that it provides no obvious way to account for positive fan effects on targets accompanied by negative fan effects on foils. We have considered a number of search models which assume self-terminating search on target trials and exhaustive search on foil trials. A simple example will illustrate the problem. Assume a serial search of the propositions with t as the time to process each proposition; t_6 is less than t_3 because of the difference in strength. Let $T^{(+)}$ and $T^{(-)}$ be search times for targets and foils, respectively. Then, the fan effect for targets is

$$T_6^{(+)} - T_3^{(+)} = 3.5t_6 - 2t_3$$

which is greater than or equal to zero only if t_6 is greater than or equal to $(2/3.5)t_3$. For an exhaustive search of the foils, a negative fan effect will occur only if t_6 is less than $.5t_3$. The two conditions are clearly contradictory. We have examined parallel searches of more complex networks, similar to those presented by Anderson (1976), and have arrived at similar contradictions. While we can offer no general proof that a pure strength model is unable to account for the fan × probe interaction, we have had no success in finding an instantiation that can do so.

A mixed strategy. Reder and Anderson (1980) hypothesized that, in blocks in which related foils were used, subjects sometimes searched memory and sometimes merely checked the probe for consistency with the theme. Such a strategy is capable of predicting negative fan effects, as well as the fan × integration interaction. Let

$$T = a + pc + (1 - p)t \tag{3}$$

where a is the intercept, p is the probability of doing a consistency check, c is the checking time, and t is search time. Assume that p is larger for integrated passages with six propositions (these are more "thematic"); p would also be larger in the take-home than in the laboratory acquisition experiment. Note that t will increase with fan. Starting with Eq. (3), it can be shown that negative fan effects will result if c is sufficiently smaller than t_3, and p_6 is sufficiently larger than p_3.

Once again the problem is the fan × probe interaction. Reder and Anderson have pointed out that their model predicts larger fan effects for foils than for targets, the opposite of our results. The reason for this prediction is that correct response times for foil trials presumably represent search; a subject checking the probe for consistency with the theme on foil trials would ordinarily respond positively, and therefore incorrectly, because the foils are related to the theme.

It is of some interest to consider why the mixed strategy model accounts for Reder and Anderson's data and not for our own. We have already noted that even our low-integrated sixes are more likely to be integrated than the randomly chosen subset of threes employed by Reder and Anderson. Reder and Anderson's subjects may also have had more motivation to check consistency. Because unrelated foils were employed on alternating test blocks, a consistency check would yield a correct response on .75 of all trials, a fairly high rate of reinforcement for such a strategy. Furthermore, subjects could check easily because the theme was always obvious from the probe. For example, typical predicates involved circus performers for the circus theme, rolling dough for the pie baking theme, and references boarding a train for the train ride theme. Under these circumstances, the theme checking strategy makes sense. In contrast, our probes are neutral; such phrases as "took a seat" or "made a selection" are consistent with any of our themes. Therefore, theme checking seems less likely in our study; it is not easy to apply and may require more time than a memory search.

A Class of Process Models

The models just considered have difficulty accounting for the fan × probe interaction. This pattern of fan effects could occur, however, if subjects occasionally terminated searches on foil trials without exhaustively searching all paths from the propositional subject. We must also assume that such early exits from the search on foil trials are more likely with highly integrated materials. Our sixes clearly are more integrated than our threes.

We will consider two models of this general class. Their mechanisms and structures are not mutually exclusive. They differ largely in that one model emphasizes declarative knowledge, encoded elaborations connecting predicates, while the other model emphasizes procedures for judging plausibility of the probe. Both models are oversimplifications and undoubtedly wrong in detail;[1] however, they demonstrate that models of the general class can account for the trends in our data.

Elaborative networks. Smith and his colleagues (Smith, 1981; Smith et al., 1978; Whitlow et al., 1982) have suggested that connections among predicates in the representation of integrated material can attenuate the fan effect. Similarly, Anderson (1981) has proposed an elaborative process which creates connections between concepts in the test sentences and those in a paragraph learned earlier. These provide an added indirect path for the spread of activation between the propositional subject and the probe predicate, resulting in a reduced fan effect with integrated materials. As they stand, these explanations can predict negative fan effects but not the fan × probe interactions observed in the present experiment. What is required is early terminating of the search on foil trials. Let us consider such a process built upon the elaborated structure suggested by Anderson and Smith.

Assume that the presentation of a probe starts activation along paths from the propositional subject. There are m predicates connected to the

[1] To note just a few of the simplifications, we have ignored the complexity of the propositional structure, we have assumed that all the predicates within a passage are equally interrelated, and we have not included the possibility of a successful search from the predicate. Furthermore, for the sake of mathematical tractability, we have assumed a serial search model.

subject of the probe remaining to be processed. It takes time t to find the next connected predicate and to check for a match between the subject–predicate connection found in memory and that embodied in the probe; t does not depend upon any of the variables of our experiment. Consider a target trial. If the predicate found does not match the target predicate, there is probability r that activation will flow from it to the target predicate by way of an interconnection. The time for this path to be traversed is c. We assume that r is a function of integration and, in our study, r will therefore be larger for fan-6 than for fan-3 ($r_3 < r_6$). Given these assumptions, we can describe the time remaining in the search when m predicates remain to be checked:

$$T_m^{(+)} = t + [(m - 1)/m]\,[rc + (1 - r)T_{m-1}^{(+)}] \qquad (4)$$

A closed-form solution to Eq. (4) is available but not particularly transparent. For our purposes it is sufficient to rewrite Eq. (4) as

$$T_m^{(+)} = t + \left[\frac{rc + (1 - r)t}{m}\right] \sum_{i=1}^{m-1} (m - i)(1 - r)^{i-1} \qquad (5a)$$

when $r < 1$ and

$$T_m^{(+)} = t + (c/m)(m - 1) \qquad (5b)$$

when $r = 1$.

Let us examine some of the properties of Eq. 5. If $r = 0$, search time is $.5(m + 1)t$, the result for a self-terminating search; clearly, $T_m^{(+)}$ increases as a function of m. As r increases, $T_m^{(+)}$ will decrease if c is less than t; this is because $rc + (1 - r)t$ is a weighted average of c and t. The fan effect for targets may be obtained by subtracting $T_3^{(+)}$ from $T_6^{(+)}$ with the r's appropriately subscripted. We have computed results with various parameter values but a simple illustration of the model's properties is obtained by setting r_6 to 1 and r_3 to 0. Then the fan effect is $(1/6)(5c - 6t)$, which will be positive (negative) whenever c is greater (less) than $1.2t$. Thus, the model can yield positive or negative fan effects. In order to determine whether it can account for the fan × probe interaction, we must consider search times for foil trials.

The critical assumption here is that the search will be terminated either when all m predicates have been examined or when a contradiction is found. For example, if the probe is "The banker went home early" finding the proposition "The actor went home early" might result in a termination of the search.[2] We assume that such contradictory information is activated by way of interconnections between passages connected to the same theme; for the sake of parsimony, we again assume that the

[2] This assumes that the subject recognizes that the actor and the banker are associated with the same theme (baseball), and that predicates are never repeated within a theme.

probability of such activation is r.[3] The appropriate expression is

$$T_m^{(-)} = t + rc + (1 - \mathrm{r})\, T_{m-1}^{(-)} \tag{6}$$

and its solution is

$$T_m^{(-)} = (t + rc) \sum_{i=1}^{m-1} (1 - r)^{i-1} + (1 - r)^{m-1} t \tag{7a}$$

for $r < 1$ and

$$T_m^{(-)} = t + c \tag{7b}$$

for $r = 1$.

Again, taking the extreme cases in which $r_6 = 1$ and $r_3 = 0$, we find that the fan effect is $c = 2t$ which will be positive (negative) whenever c is greater (less) than $2t$. Note that for these values r, if c is greater than $1.2t$ but less than $2t$, we have positive fan effects for targets and negative fan effects for foils, just as we observed in the high-integrated data of Experiment 2 and in the low-integrated data of Experiment 3. Brining r_3 and r_6 closer together reduces the freedom of choice we have for c but still allows the model to account for the pattern of interest.

Judging plausibility. We can generate a second model by replacing activation of paths between propositions by procedures or productions (Newell, 1973; Anderson, 1976). Assume that these procedures determine whether a probe is plausible on the basis of propositions encountered as activation spreads through the memory network. For example, activating the proposition "The actor was bored" may be sufficient to stop the search and respond "old" to the probe "The actor went home early." Similarly, activation of "The banker cheered for his team" may make the probe "The banker went home early" sufficiently implausible to stop the search with a "new" response. Again assume that the search mechanism is a spread of activation. It again takes time to activate a proposition. We assume that as each proposition in memory is activated, the subject computes the plausibility of the probe; this takes time c and is successful with probability r. The value of r again depends upon the degree of integration. This computationally oriented model results in a mathematical structure very similar to that for the elaborated network model: for targets,

$$T_m^{(+)} = t + [(m - 1)/m]\, [c + (1 - r)\, T_{m-1}^{(+)}] \tag{4'}$$

and for foils,

$$T_m^{(-)} = t + c + (1 - r)\, T_{m-1}^{(-)} \tag{6'}$$

[3] A more realistic model would distinguish between connections within and connections between passages. Such a distinction is supported by the fact that repetition effects based on the same subject are greater than those based on the other subject associated with that theme (see Tables 4 and 6).

The model is similar in spirit to a plausibility judgment model proposed by Reder (1976). She suggested that positive and negative evidence cumulates during the memory search and that the search ends with a decision when either of two cutoffs is reached. Presumably, the search would end if the probe itself, or a direct contradiction of the probe, were found. Models of this class are appealing because they can account for effects of many variables, including probe probabilities and payoffs, upon both response times and error rates. One of the problems with network models is that their account of errors either has been missing or has had a decidedly post hoc flavor.

In summary, two models have been presented which can predict either negative or positive fan effects, and which can account for small positive fan effects on target trials accompanied by large negative fan effects on foil trials. The critical feature of models of this class is that they permit disconfirming evidence to terminate search on foil trials; they embody the assumption that such evidence is more readily available with more integrated materials. Note that confirming evidence provides a similar advantage to more integrated materials on target trials but the savings in time on foil trials tends is greater.

CONCLUDING REMARKS

The difficulties in deciding between declarative and procedural models have been extensively discussed in other contexts. Many of the issues have been noted by Smith (1978) in reviewing theories of semantic verification, and by Anderson (1978) in his consideration of analogical and propositional models of imagery effects. The same difficulties will undoubtedly arise in distinguishing elaborated networks and plausibility models. Rather than become trapped in that morass, we should focus our initial efforts on the following goals. First, if our analysis of the differences between Reder and Anderson's (1981) and our results is correct, it should be possible to prescribe materials and conditions of acquisition which will produce either set of results. Such experiments are basic to a test of any explanation of integration effects. It really reduces to being precise about the meaning of integration. Second, experiments should test the general position that both the elaborated network and plausibility models represent. One such experiment would involve systematically varying the direct contradictions between two passages tied to the same theme. If one propositional subject drinks a martini and the other a manhattan, we would expect negative fan effects on foil trials. Third, we require more complex models than those presented to evaluate the adequacy of our assumptions. Such models would take into account propositional complexity and variable relational strength, include the spread of activation from the predicate, and allow activation to spread in parallel. The eventual goal should be stochastic models that provide an account of repetition and nonrepetition statistics. Research along these lines will further our understanding of the role of integrated text in memory of propositions.

REFERENCES

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology,* **6,** 451–474.

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

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review,* **85,** 249–277.

Anderson, J. R. (1981). Effects of prior knowledge on memory for new information. *Memory & Cognition,* **9,** 237–246.

Black, J. B., & Bern, H. (1981). Causal coherence and memory for events in narratives. *Journal of Verbal Learning and Verbal Behavior,* **20,** 267–275.

Black, J. B., & Bower, G. H. (1979). Episodes as chunks in narrative memory. Journal of Verbal Learning and Verbal Behavior, **18,** 309–318.

Bradshaw, G. L., & Anderson, J. R. (1982). Elaborative encoding as an explanation of levels of processing. *Journal of Verbal Learning and Verbal Behavior* **21,** 165–174.

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

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin,* **75,** 424–429.

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

Newell, A. (1973). Production systems: Models of control structures. In W. G. Chase (Ed.), *Visual information processing.* New York: Academic Press.

Pollack, I., & Norman, D. A. (1964). A nonparametric analysis of recognition experiments. *Psychonomic Science,* **1,** 125–126.

Reder, L. M. (1976). *The role of elaborations in the processing of prose.* Unpublished doctoral dissertation, University of Michigan.

Reder, L. M. (1979). The role of elaboration in memory for prose. *Cognitive Psychology,* **11,** 221–234.

Reder, L. M. (1982). Plausibility judgments versus fact retrieval: Efficient strategies for question-answering. *Psychological Review,* **89,** 250–280.

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. (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.

Smith, E. E. (1973). Theories of semantic memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 6). New York: Erlbaum.

Smith, E. E. (1981). Organization of factual knowledge. In H. E. Howe, Jr., & J. H. Flowers (Eds.), *Nebraska symposium on motivation* (Vol. 28). Lincoln: Univ. of Nebraska Press.

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

Whitlow, J. S., Jr., Smith, E. E., & Medin, D. L. (1982). Retrieval of correlated predicates. *Journal of Verbal Learning and Verbal Behavior,* **21,** 383–402.

(Accepted November 30, 1983)