Interference with Real World Knowledge

CLAYTON H. LEWIS AND JOHN R. ANDERSON

The University of Michigan

Two experiments are described in which subjects studied made-up, fantasy facts about well-known persons and then were asked to verify actual facts about these persons. Reaction time to the actual facts was longer the more fantasy propositions studied about a person. Reaction time was also longer when the verification test involved a mixture of actual and fantasy facts rather than just actual facts. A mathematical version of the ACT model (Anderson, 1976) was fit to the data. It provides a satisfactory fit, better than an alternate model. However, some of the parameter values estimated for the ACT model seemed unreasonable.

A number of recent studies have supported the idea that the more propositions a subject knows about a concept the slower he is to verify any one of them. (See Anderson & Bower, 1973; Anderson, 1974; Thorndyke & Bower, 1974.) Anderson (1976) has developed the ACT model to account for these findings. These studies have employed artificial "facts" learned by subjects in an experimental situation. Intuitively, it seems unlikely that these results would generalize to the verification of propositions which are part of a subject's stock of actual facts. If these results did generalize it would seem to imply that the more one knows about a concept, the slower he would be to verify any particular fact about that concept. In the experiments reported below we explored this matter by having subjects learn varying numbers of artificial facts about well-known people. We then examined the effect of the number of artificial facts learned about a person on the time to verify actual facts about him. Since these experiments involve possible interference of propositions learned in an experimental situation with retrieval of actual knowledge they are also of interest in considering whether memory is so structured that information learned under different conditions can be scanned independently if task demands warrant.

These experiments are concerned with whether experimentally learned material will interfere with pre-experimental material. A similar question

has been subject to extensive experimental analysis—the question of extraexperimental sources of interference (e.g., Postman, 1962; Underwood & Ekstrand, 1966; Underwood & Postman, 1960). It had been hypothesized that high frequency words would be subject to more rapid forgetting because of greater interference due to more competing extraexperimental associations. The initial results in this investigation were largely negative. However, these experiments tended to confound factors such as form class and imagery rating with frequency. There is a recent set of experiments which seem more favorable to the hypothesis of extraexperimental sources of interference. These have controlled for form class and imagibility. M. Fehling (personal communication) has found poorer retention for high frequency paired associates. Kieras (1974) has found poorer recall for sentences composed of high frequency words. Perlmutter, Harsip, & Meyers (1976) also report longer recognition latencies to pairs composed of high frequency pairs.

An experiment by Slamecka (1966) is very similar in intent to the experiments to be reported here. He looked at the effect of learning interfering experimental responses on subject's ability to recall free associations to a stimulus item. There was no effect of these experimental associations, whether measured by frequency of recall or by response latency. The experiments to be reported here will rely on a more accurate and sensitive latency measure than in Slamecka's research. However, his results clearly do not favor the prospects of getting an effect of experimentally acquired facts on pre-experimental facts.

TABLE 1

EXAMPLES OF MATERIAL USED IN THE EXPERIMENT

Examples of artificial facts studied
 George Washington wrote *Tom Sawyer*.
 Napoleon Bonaparte was from India.
 Napoleon Bonaparte was a singer.
 Napoleon Bonaparte is a liberal senator.
 Napoleon Bonaparte had a ranch.
Examples of test probes
 Actual facts
  0 Fidel Castro is Cuban.
  1 George Washington crossed the Delaware.
  4 Napoleon Bonaparte was an emperor.
 Fantasy facts
  1 George Washington wrote *Tom Sawyer*.
  4 Napoleon Bonaparte was from India.
 False
  0 Fidel Castro was a Texan politician.
  1 George Washington is a swimmer.
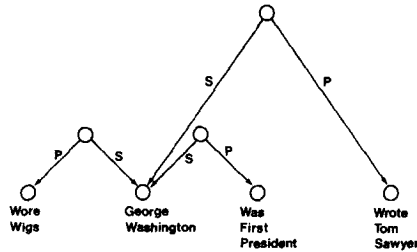 4 Napoleon Bonaparte was a humorist.

FIG. 1. A schematic ACT representation of some of the information attached to George Washington.

Table 1 illustrates the materials used in these experiments. A number of fantasy facts, not actually true, were studied about various well-known public figures. After learning these facts, subjects had to respond true or false to test probes that were either actual facts, fantasy facts, or false sentences (neither true in the real world nor learned during the experiment). Subjects were instructed to respond true to fantasy facts as well as to actual facts. The digits in Table 1, not shown to the subject, indicate how many fantasy facts were studied about the individual. The major question is whether the number of fantasy facts studied about a person would influence the speed with which an actual fact about that person can be verified.

According to the ACT model the subject's knowledge about a person in the experiment can be represented schematically as in Fig. 1. Figure 1 illustrates a very simple network in which the subject, *George Washington*, is connected to a number of predicates. Figure 1 is a simplification of the ACT representation which is similar to the HAM model of Anderson and Bower (1973). We are using this simplification to ease the task of exposition. If a given proposition is to be verified, nodes representing its subject and predicate are identified in memory, and search processes are started from both nodes. These searches spread outward along all arcs from these nodes until they intersect, at which time a check is performed to determine whether the connection between subject and predicate matches the proposition to be verified. If it does, the appropriate response is made. This search mechanism is similar to that proposed by Collins and Loftus (1975). The spread of search processes from a node is assumed to be a limited-capacity parallel process (see Townsend, 1974). This means that the more arcs are encountered at a node the slower will be the progress of the search along each arc. Referring to Fig. 1 it will be seen that the process of verifying an actual fact about George Washington will accordingly be slowed down by the presence of the fantasy fact about him. If additional propositions with the same predicate as an actual fact were studied, verification of the actual fact would be slowed for the same reason, since search also proceeds from the predicate node.

EXPERIMENT I

Experiment I allows a test of these predictions. The number of artificial facts studied about a person was varied from zero to four. So that the effect of extra arcs at the predicate node could be assessed, predicates from actual facts to be tested were used in two fantasy facts or were not used at all in the fantasy facts.

Method

Materials. A list of 40 well-known actual facts about public figures was prepared, two facts for each of 20 individuals (see Appendix). Each fact was of the form NAME PREDICATE, the 40 predicates being all different. From these 40 facts materials were constructed for a learning phase and for a later reaction time test phase. One of the two actual predicates associated with each name was randomly designated a *study* predicate, the other being *reserved*. Reserved predicates were not used in constructing fantasy facts for study. A total of 40 fantasy facts was constructed by repairing names and study predicates, four fantasy facts being produced for each name in treatment 4, three for each in treatment 3, and so on for treatments 2, 1, and 0. These 40 propositions, together with six filler propositions constructed from similar materials, formed the study materials for the subject. To form the test materials the original 40 actual facts were added to the 40 fantasy facts, and 80 new false propositions to be used as foils were constructed by repairing names and predicates in such a way that among the test materials each name and each predicate appeared equally often in real or fantasy facts as in foils. Note that all fantasy facts and all false facts used predicates that came from actually true facts. One motivation for this was to assure that the predicates used for the false facts and the fantasy facts had credibility as the sort of thing that might be said of a public figure. Building fantasy facts in this way has the disadvantage that their predicates were connected to other individuals in a nonsystematic way; for example, in the fantasy fact "George Washington wrote *Tom Sawyer*" the predicate is strongly associated with another individual, while this is not true of "George Washington was a scientist." Randomizing the use of predicates and individuals was intended to deal with this problem.

Another point to note about the fantasy facts is that many were inconsistent with known facts. This gave the fantasy facts a different character from obscure actual facts that subjects had not seen before. It is not clear how results might differ if such true but not known facts were used.

It was intended that each subject have his own randomly constructed set of materials. However, through an oversight only nine different sets of materials were used.

Subjects. Twenty-six female subjects from the Human Performance Center paid subject pool participated in the experiment in groups of one to four. Of these, three failed to perform sufficiently accurately in the test phase of the experiment and were eliminated. Subjects were paid $2.00 per hour for their services.

Procedure. Study and test materials were presented and responses and response times collected by an IBM 1800 computer with four cathode ray tube terminals, each in a semi-isolated booth. The GEPS II programming system (Kieras, 1973) was used, allowing up to four subjects to be run independently at one time. A subject was first given a list of the 20 names used in the materials, plus those appearing in filler items, and asked to mark any which were unfamiliar. The names were well known to the subjects: the mean number marked was .2. No aid was given to subjects' recognition beyond mentioning a common nickname if one existed.

The subject then entered the study phase of the experiment. The study propositions were presented one at a time on the CRT screen. The subject was told to regard these

propositions as describing a fantasy world peopled by the individuals described. She was asked to write a short continuation of each proposition on a numbered sheet, then press a button to see the next proposition. The nature of the continuation was not prescribed beyond the requirement that it relate in some way to the subjects' understanding of the proposition. Examples given in the instructions included adding embellishing detail and providing a reconciliation of the proposition with actually known facts.

The 40 study propositions were presented in three random orders, preceded and followed each time by three filler items. The subject was asked not to consult her previous written continuations when seeing a proposition again but was free to use the same continuation again if she chose. To decrease the likelihood that the subject might try to memorize the study materials verbatim, she was told that the continuations would be used to assess her handling of fantasy materials and was not told that she would be tested for retention of propositions or continuations. The study procedure was self-paced and took about 50 min.

On completing the study phase the subject entered the test phase of the experiment. Test propositions were presented one at a time on the CRT and the subject was told to press a true or false button as quickly as was consistent with reasonable accuracy. In particular, she was to respond *true* to any proposition which was an actual fact or which was one of the fantasy facts seen during study, and *false* to other propositions. To be sure the subject understood this requirement she was asked to write out in her own words the relevant part of the instructions and get the experimenter's approval or clarification before beginning the test phase. The subject was free to choose which hand to use for *true* responses and which for *false*. After each response the subject was told whether she was correct or incorrect and the next item was presented after a delay of 1 sec. The test items were presented in three random orders to each subject, each order preceded and followed by six filler items and separated by rest periods. The test phase lasted about 50 min, including two rest periods.

After the test phase the subject was given a list of the actual facts used in the experiment including fillers, and asked to mark any which had been unfamiliar when first encountered in the test or which the subject felt were untrue. The materials were reasonably familiar: the mean number of facts marked was .5.

Results

For each subject, response times for correct responses for each type of test item were averaged, times longer than 5 sec being treated as 5 sec and times shorter than 200 msec being classed as errors. The subject means were then averaged, and are shown in Fig. 2. The abscissa in Fig. 2 is *fan*: the number of fantasy propositions studied about the name in the test item. False test propositions and actual facts are broken down according to whether the predicate in the item was a *study* predicate or a *reserved* predicate; that is, whether the predicate in the test item was seen during the study phase or not.

The subjects × conditions interaction mean square was used to estimate the standard error of a condition mean, which was 49.6 msec. This estimate may be too low given that a large number of conditions was used in a within-subjects design, but except where noted it is used below to give some idea of the reliability of differences in the data.

The first aspect of the results to note is the increasing trend of time to verify actual facts with reserved predicates as fan increases. Because of
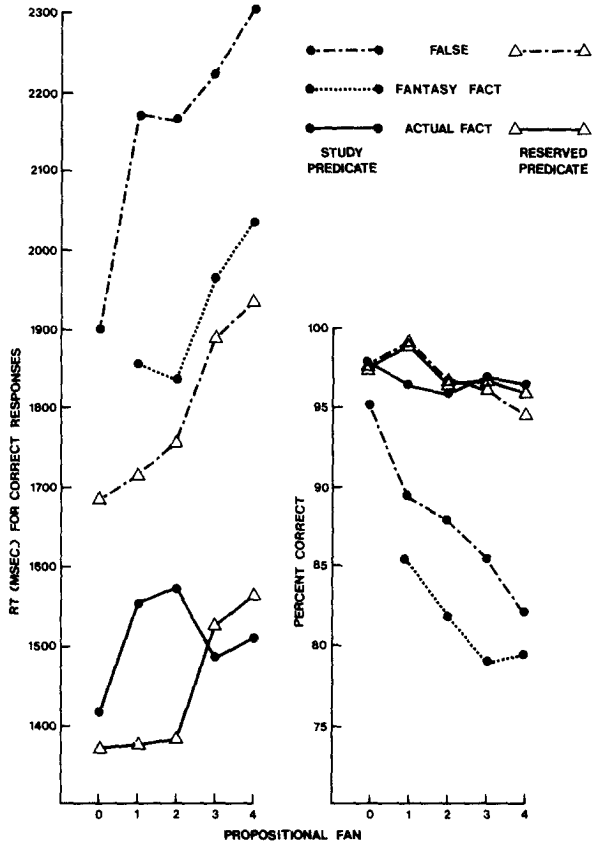
FIG. 2. Reaction times and accuracy in Experiment I.

the importance of this trend its reliability was assessed by calculating the value of a linear trend contrast for each subject and estimating the variability of this contrast directly rather than using the overall estimate of standard error. This computation indicates that the trend is quite significant ($t_{22} = 4.0, p < .001$). Time to verify fantasy facts also increases with fan, ($t_{506} = 2.8, p < .01$) and verification of artificial facts takes longer than verification of actual facts of the same fan ($t_{506} = 14.0, p < .001$).

On the average, actual facts with study predicates take longer to verify than those with reserved predicates ($t_{506} = 2.0, p < .05$), and the interaction with fan is significant ($t_{506} = 2.3, p < .025$). That is, the effect of fan is significantly less for those actual facts with study predicates. Time to reject false propositions increases with fan ($t_{506} = 5.5, p < .01$). Propositions with reserved predicates can be rejected faster than those with study predicates ($t_{506} = 10.1, p < .005$).

Proportions correct for the various types of test items are also shown in Fig. 2. Errors tend to increase with fan.

Discussion

The basic predictions of the ACT model are borne out: Studying fantasy facts about a person slows down the verification of actual or fantasy facts about that person. Studying extra occurrences of a predicate slows down verification of facts using that predicate. The interaction between prior use of predicate and fan was not expected. The higher the fan associated with a name, the greater the probability that the name would occur close to its study predicate during the study phase. Thus, there is a greater probability that the subject may notice and implicitly rehearse the relation between the name and its actual predicate. This might produce decreased reaction time to the name–predicate combination thus causing the interaction between fan and prior use of the predicate.

To account for the overall speed advantage of actual facts over fantasy facts one more feature of the ACT model must be described. Memory connections can have different strengths, and search energy is divided among connections in proportion to strength. It is reasonable to assume that connections making up fantasy facts are weaker than those making up actual facts, so that it takes longer to find a fantasy fact if it is present.

The principle means for rejecting false propositions in the ACT model is by means of a process that waits a certain period of time for an intersection to occur. If an intersection does not occur in that period of time, a false response is generated. This waiting process is attuned to the various variables, such as fan, which can delay the occurrence of an intersection for true probes. That is, the process waits longer for probes consisting of elements with greater fan. This accounts for at least the qualitative features of the rejection data: Rejection is slowed by fan and by extra use of predicate.

The accuracy data in Fig. 2 are not accounted for directly by the ACT model. One may speculate that imperfect original learning is responsible for the misses in verifying fantasy facts, with the proviso that some sort of interference accounts for poorer learning of facts about names with higher fan. Another possibility is that guesses are produced after allowing processing to go on for some interval, so that items taking longer to process will get more guesses and hence more errors. This could account for the trend in errors for fantasy facts and false test items with study predicates. This sort of guessing model has been developed and tested by King and Anderson (in preparation) and Mohs, Westcourt, and Atkinson (1975).

Note that while large differences in accuracy occur among the conditions the principle differences in latency cannot be attributed to differ-

ences in speed accuracy tradeoff: Slower conditions are less accurate. It is possible that the accuracy differences do reflect differences in underlying processing which would make comparison of latencies problematical, but we will argue below that the sort of model proposed by Mohs *et al.* can account for the accuracy data in a way that does not require serious reinterpretation of the latencies.

While the ACT model is fairly well supported by these results, there is another kind of model which can also explain them. In the ACT model the interfering effect of fantasy facts occurs during *search*: Fantasy and actual facts are stored together in memory, and fantasy facts tie up search capacity when actual facts are being sought. In the other kind of model interference occurs because of a *decision* problem: Fantasy and actual facts can be searched independently, but the subject does not know in advance where to search for a test item. If the test item is an actual fact, but the subject searches first among his store of fantasy facts, a penalty is incurred which increases with fan.

The simplest form of decision–interference model assumes that fantasy and actual facts are searched serially. If m fantasy facts can be searched in time $d + me$, d being a constant overhead and e the time to check one fact, and all actual facts can be searched in time A, and if actual facts are searched first with probability p, then the average search time for an actual fact with a name of fan m will take $A/2 + (1-p)(d + me)$. The factor $A/2$ in this expression reflects the assumption that the correct proposition will be uncovered halfway through the test of actual propositions. This formula accounts for the increase in verification time attendant upon fan.

The main effect of predicate type on verification could be captured in this model at the expense of another parameter: The probability of starting search with actual or artificial facts could be allowed to vary according to the number of times the elements of the test item were seen during study. This is plausible, since if a subject could be certain that he had not seen a predicate during study he should never search his artificial facts at all.

EXPERIMENT II

Both search–interference and decision–interference models can account for the Experiment I data reasonably well. Experiment II was carried out in an attempt to separate these two models. Experiment II was very similar to Experiment I, the major change in procedure being that subjects were tested on actual and fantasy facts separately to minimize the decision problem that they may have faced in the Experiment I test in which these items were mixed.

According to the decision–interference model there should be no effect

of fantasy facts on verification of actual facts if subjects know they are being tested only on actual facts. In the ACT model, interference should be just as great whether test items are mixed or not. A mixed test was also given to Experiment II subjects to provide a check on the comparability of the procedures. Another change in Experiment II was that some fantasy facts were presented more often than others during the study phase. According to the ACT model this should have the effect of increasing the strength of these propositions, and hence allowing faster responses to these but also increasing the interference they cause.

Because of the interaction of fan with prior use of predicates found in Experiment I and the possibility it suggested of implicit rehearsal occurring during study, it was decided to reserve in Experiment II all predicates associated with actual facts to be tested. Finally, only fans of 0, 1, and 4 were used, as it was felt these gave an adequate picture of fan effects.

Method

Materials. The set of 40 facts about 20 names was used again, but the makeup of study and test materials was different from Experiment I. For each subject one of two groups of 10 names was selected at random and called the *test* group. Two names from this group were assigned randomly to each of the treatments 0, 1, 1r, 4, and 4r. These treatment designations indicate the number of fantasy facts to be studied about the individual, and whether the fantasy facts will be repeated during study. Thus an individual in treatment 4r will have four fantasy facts each presented twice as often during study as the four fantasy facts about an individual in treatment 4. The predicates associated with the other group of names, the nontest group, were used to construct the fantasy facts. Altogether there were 20 fantasy facts. These 20 facts, together with six filler items, formed the study materials for the subject. In preparing the study materials for presentation each fantasy fact with a name in treatments 1r or 4r was repeated, so that these items were seen twice as often during study as items with names in treatments 1 and 4.

There were three types of reaction time test blocks—pure tests of actual facts, pure tests of fantasy facts, and mixed tests of actual and fantasy facts. In each of these blocks there were an equal number of true and false probes. The false probes were constructed so that each name and each predicate appeared equally often in true and false probes. This was to prevent the names or the predicates from acquiring a response bias. To achieve this balance it was necessary to use some filler actual trues (constructed from the nontest group) and some filler falses.

Subjects. Thirty female subjects from the Human Performance Center paid subject pool participated in the experiment. They were paid $2.00 per hour.

Procedure. The procedure used in Experiment II was very similar to that used in Experiment I. Since a smaller number of fantasy facts was used subjects spent less time on the study phase, about 40 min. On completing the study phase the subject was given instructions describing three different types of tests she would encounter during the next part of the experiment. In each test, propositions were presented, responses collected, and feedback given exactly as in Experiment I, but the subject's rule for responding was different for the three types of tests. In a *pure fantasy* test the subject was to respond *true* only to items she remembered from the study phase, and *false* to all others. No actual facts were included in the materials presented in this kind of test. In a *pure actual* test the subject was to respond *true* only to actual facts. No fantasy facts were presented here. In a *mixed* test the

subject was to respond *true* to actual or fantasy facts as in Experiment I. To ensure that subjects understood the types of tests they were required to paraphrase versions of the relevent parts of the instructions as in Experiment I before starting the test phase.

Each subject received a series of six tests, two of each type, arranged in such a way that each type appeared once among the first three tests, with the last three being simply the first three in reverse order. These test orders were counterbalanced over subjects. Each test was preceded by a summary of the rules for responding, and the subject was told to consult her instruction sheet before proceeding. The test series took about 40 minutes.

The familiarity of the materials was checked exactly as in Experiment I. An average of 0.9 names and 1.3 facts were marked as not known.

Results

Data for Experiment II were treated as described in Experiment I. The mean reaction time for correct responses are shown in Fig. 3 and the accuracy data is shown in Fig. 4. Test results are shown broken down by fan, the number of fantasy facts studied about the individual in a test item. They are also broken down according to which test type they come from—
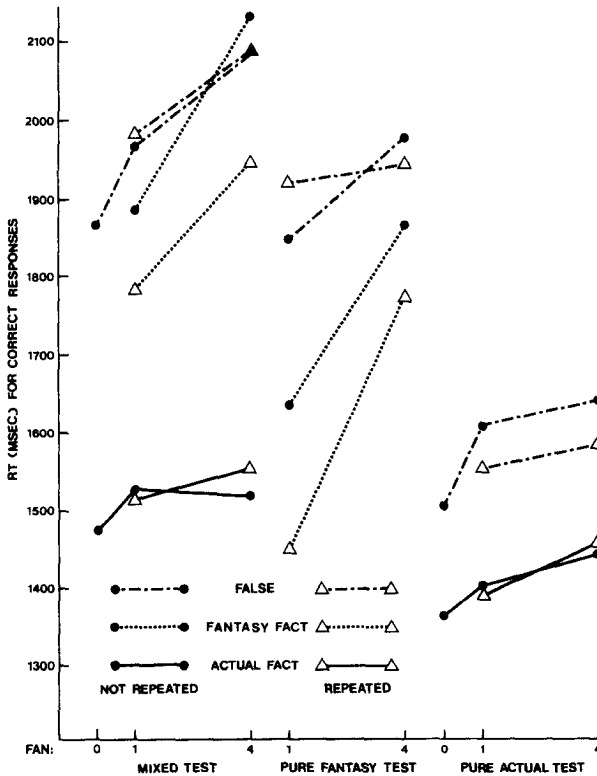

FIG. 3. Reaction times in Experiment II.

mixed test, pure fantasy, or pure actual. Points designated *repeated* show results for those items for which the fantasy facts were shown twice as often during study. Thus items about individuals in treatment 4r appear as fan 4, repeated. The standard error of the 32 reaction time means in Fig. 3, estimated from the interaction mean square in a conditions × subjects analysis of variance, is 56 msec.

Considering first the results for the mixed test, one can note a pattern of times similar to that found in Experiment I, with verification of actual facts being more rapid than verification of fantasy ($t_{899} = 10.2, p < .005$). Repetition makes verification of fantasy facts faster ($t_{899} = 2.52, p < .01$) and reaction time increases with fan ($t_{899} = 3.59, p < .005$). There is no interaction between the effect of fan and repetition on fantasy facts ($t_{899} = .71$). Time to reject a false fact is also increased by fan ($t_{899} = 3.36, p < .005$). Effects of fan and repetition on verification of actual facts are less clear. Their effect was investigated by computing a fan 4–fan 0 contrast for each subject and calculating its variability. The difference is marginally significant ($t_{899} = 1.47, p < .1$). There is no apparent effect of repetition on actual facts ($t_{899} = .18$).

In the pure fantasy test the same effects of fan found in the mixed test appears for verification ($t_{899} = 4.86, p < .005$). The effect of repetition on verification is also significant ($t_{899} = 2.44, p < .025$). The effect of fan on rejection is marginally significant ($t_{899} = 1.33, p < .1$). The fan × repetition interaction for rejection is not significant ($t_{899} = .94, p < .2$).
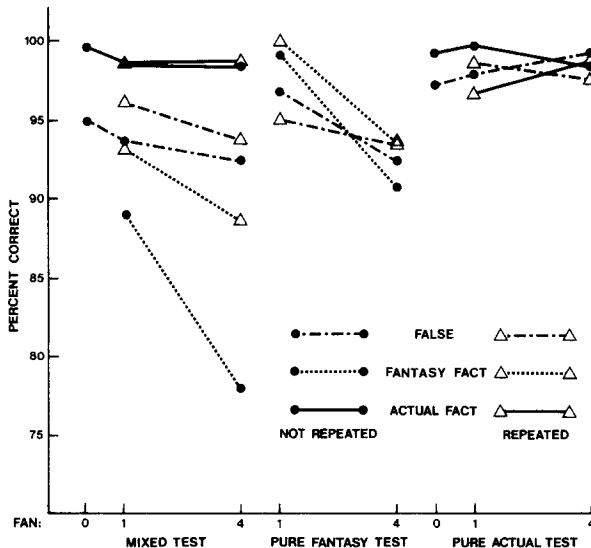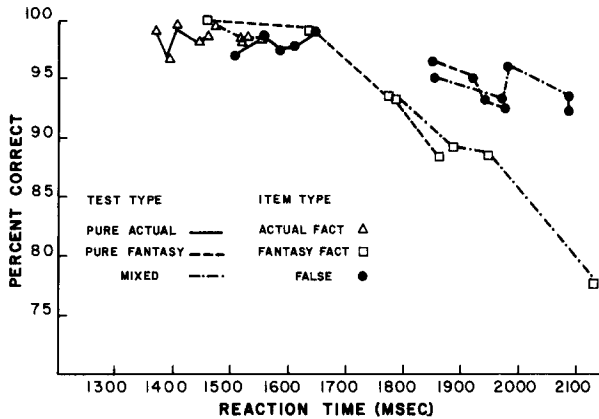

FIG. 4. Accuracy in Experiment II.

FIG. 5. Speed–accuracy plot for Experiment II. The lines are intended simply to indicate test type.

In the pure actual tests the effect of fan and repetition on verification was investigated as for the mixed test by computing a fan 4–fan 0 contrast for each subject. The contrast is positive ($t_{29} = 1.97$, $p < .05$); 21 of 30 subjects showed positive contrast, ($p < .05$ by sign test). The effect of fan on rejection time is marginally significant ($t_{899} = 1.55$, $p < .01$). There is not a significant effect of repetition on either time to accept an actual fact ($t_{899} = .10$) or time to reject a false ($t_{899} = .96$).

Comparison of the times for the three tests show that verification is faster in the pure tests than for corresponding items in the mixed test; actual facts are faster in the pure actual test ($t_{899} = 2.91$, $p < .005$), and fantasy facts are faster in the pure fantasy test ($t_{899} = 6.42$, $p < .005$). Similarly rejections are faster in the pure actual test than in the mixed test ($t_{899} = 11.7$, $p < .005$), faster in the pure fantasy test than in the mixed test ($t_{899} = 2.77$, $p < .005$), and faster in the pure actual test than in the pure fantasy ($t_{899} = 8.08$, $p < .005$).

With respect to the accuracy data, the mixed test results in Fig. 4 are quite comparable to the Experiment I results in Fig. 2; in particular, accuracy for fantasy facts is poor and gets worse as fan increases. In the pure fantasy test, accuracy for fantasy facts is much better, but still decreases with fan. Figure 5 provides a plot of the accuracy in the 32 conditions of Experiment II against reaction time. As can be seen there is a general trend such that longer latencies are associated with lower reaction times. An analysis of Experiment I reveals a similar trend.

Discussion

Returning to the models proposed to account for the results of Experiment I, how do they fare now? The most important new facts to be

TABLE 2

PREDICTIONS OF THE DECISION–INTERFERENCE MODEL FOR EXPERIMENT II

	FAN		
	0	1	4
Mixed test			
False, not repeated	$K_f + A + d$	$K_f + A + d + e$	$K_f + A + d + 4e$
False, repeated	—	$K_f + A + d + e_r$	$K_f + A + d + 4e_r$
Fantasy fact, not repeated	—	$K_t + d + d + P_mA$	$K_t + d + 2.5e + P_mA$
Fantasy fact, repeated	—	$K_t + d + e_r + P_mA$	$K_t + d + 2.5e_r + P_mA$
Actual fact, not repeated	$K_t + A/2 + (1 - P_m)d$	$K_t + A/2 + (1 - P_m)(d + e)$	$K_t + A/2 + (1 - P_m)(d + 4e)$
Actual fact, repeated	—	$K_t + A/2 + (1 - P_m)(d + e_r)$	$K_t + A/2 + (1 - P_m)(d + 4e_r)$
Pure fantasy test			
False, not repeated		$K_f + d + e + P_fA$	$K_f + d + 4e + P_fA$
False, repeated		$K_f + d + e_r + P_fA$	$K_f + d + 4e_r + P_fA$
Fantasy fact, not repeated		$K_t + d + e + P_fA$	$K_t + d + 2.5e + P_fA$
Fantasy fact, repeated		$K_t + d + e_r + P_fA$	$K_t + d + 2.5e_r + P_fA$
Pure actual test			
False, not repeated	$K_f + A + (1 - P_a)d$	$K_f + A + (1 - P_a)(d + e)$	$K_f + A + (1 - P_a)(d + 4e)$
False, repeated	—	$K_f + A + (1 - P_a)(d + e_r)$	$K_f + A + (1 - P_a)(d + 4e_r)$
Actual fact, not repeated	$K_t + A/2 + (1 - P_a)d$	$K_t + A/2 + (1 - P_a)(d + e)$	$K_t + A/2 + (1 - P_a)(d + 4e)$
Actual fact, repeated	—	$K_t + A/2 + (1 - P_a)(d + e_r)$	$K_t + A/2 + (1 - P_a)(d + 4e_r)$

accommodated are, first, that fantasy facts have an effect on verification time even in the pure actual test; second, that verification is faster in the pure tests than in the mixed; and finally, that repetition speeds verification of fantasy propositions.

On first thought the first of these findings might seem to rule out the decision–interference model, since subjects should never examine any fantasy facts in the pure actual test. If one allows some degree of confusion on the subject's part on this point, however, some decision–interference is still expected. This hedge is made attractive by the fact that pure tests are faster. This lends some support to the decision–interference model, since it is a direct prediction from the smaller difficulty in deciding which facts to search in verifying an item. The effect of repetition on verification can be accounted for by assuming that the artificial facts studied more often can be scanned more rapidly.

To provide a rigorous test of the decision–interference model we decided to test a mathematical expression of it. This involved trying to predict the 32 reaction times displayed in Fig. 3. For the time being we are ignoring the large differences in accuracy across conditions; how these may be accounted for is considered below. The model involved the following parameters: There is a time d to access the fantasy facts. It is assumed that the time to search the fantasy facts increases by an amount e, if not repeated, and e_r, if repeated, for every fantasy fact searched. The difference between repeated and nonrepeated facts was to accommodate the fact that subjects are faster to verify repeated fantasy facts. The fourth parameter was A, the time to access and search the actual facts. Then there were three probabilities of accessing the actual facts first: P_e, for the pure fantasy test; P_m for the mixed test; and P_a for the pure actual test. Presumably, $P_e < P_m < P_a$. Finally there are two parameters K_t and K_f which represent the encoding time, response time, and

TABLE 3

PARAMETER ESTIMATES FOR THE DECISION–INTERFERENCE MODEL

| Description	Parameter	Estimate
Time to access the fantasy fact	d	324 msec
Time to search a fantasy fact	e	73.5 msec
Time to search a repeated fantasy fact	e_r	53.5 msec
Time to access and search actual facts	A	181 msec
Probability of searching actuals in fantasy test	P_e	0
Probability of searching actuals first in mixed test	P_m	.78
Probability of searching actuals first in actual test	P_a	.96
Encoding, response, etc. for true response	K_t	1310 msec
Encoding, response, etc. for false response	K_f	1391 msec
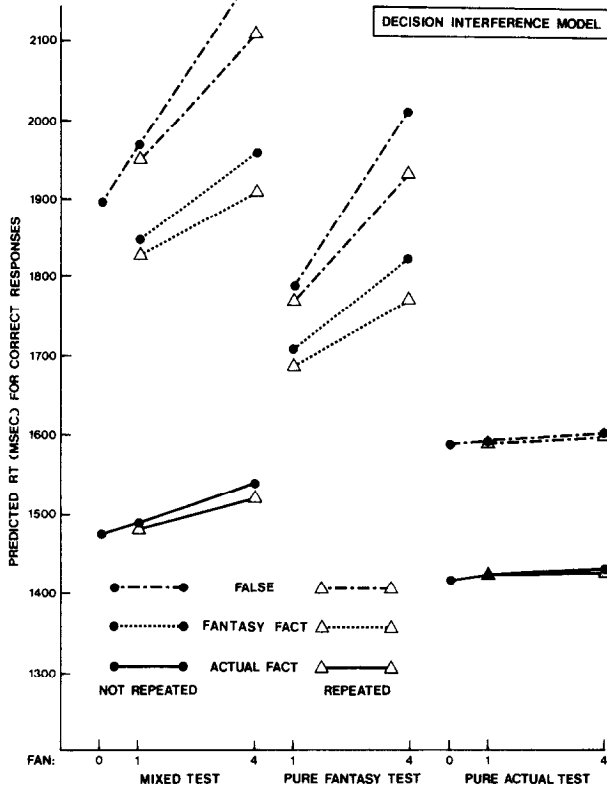
FIG. 6. Predictions of the Decision–Interference Model—to be compared with Fig. 3.

other processes not associated with the search of memory for true and false probes.

The equations giving the predictions of this model for the 32 conditions of this experiment are displayed in Table 2. The parameters of the model were estimated using the STEPIT program (Chandler, 1965), minimizing the squared deviation. These parameters are given in Table 3. Figure 6 displays the predictions of the model. It is to be compared with Fig. 3 which displays the actual data. The model does account for 91% of the variance among conditions, but the variance not accounted for is significant [$F(23,844) = 2.137, p < .005$].

A source of difficulty for the model is the effect of repetition. The model predicts that repeated fantasy facts, since they are searched more rapidly, should interfere less with actual verification. Also, since more facts must be scanned in verifying actual or fantasy facts at fan 4 than at

fan 1 the effect of repetition should be greater at fan 4 than at fan 1, but this does not appear in the data. These problems with repetition appear in extreme form in the predicted rejection times. Another difficulty is that the effect of fan on latency to an actual fact is predicted to be smaller in the pure actual test than in the mixed. Still another difficulty is that the model predicts the effect of fan to be twice as great for rejecting falses as accepting trues—because of the self-terminating nature of the search process. A further problem for the model is its failure to predict the change in relative speed of processing of fantasy facts vs false facts between the pure fantasy test and the mixed test. These defects could of course be alleviated by elaborating the model. One way to do this would be to abandon the serial search assumption of the present decision–interference model in favor of separate ACT-like searches of fantasy and actual facts. But the fit of the model as it stands is good enough to make the limited point that the data in hand do not rule out a decision–interference model of some kind.

Next we attempted to fit the ACT model as described by Anderson (1976) to the data. The actual equations used in fitting the ACT model are given in Table 4 using parameters described in Table 5. The reader may wish to refer to these tables while following the exposition below. The ACT model assumes that a subject responds true by finding an intersection of search processes among the concepts in the probe and then checking the pattern of intersecting search paths to see if a true response is justified. In developing the predictions of the model we will consider separately the processes of finding an intersection and checking it.

In the ACT model, search is carried out by an activation of memory nodes that spreads out from all concepts mentioned in a sentence. If the sentence is true, after a certain time has passed activation from all these sources will intersect. The time to respond to a sentence thus depends in part on how slowly or quickly intersection occurs. As an approximation we will assume that a sentence like *George Washington crossed the Delaware* has just two sources of activation—one for the subject, *George Washington*, and one for the predicate, *crossed the Delaware*. The time needed for an intersection to appear will depend on the rate of spread of activation from these sources along the path leading to the intersection. As an approximation, the time is given by

$$a \left(.5 \, \frac{1}{R_s} + .5 \, \frac{1}{R_p} \right), \quad \text{where}$$

a sets the time scale, and R_s and R_p are the rates of spread from the subject and predicate, respectively. This approximation is used throughout Table 4 to yield predicted times for true probes.

TABLE 4

PREDICTIONS OF THE ACT MODEL FOR EXPERIMENT II

	FAN		
	0	1	4
Mixed test			
False, not repeated	$K_f + \beta_m \dfrac{1}{\dfrac{1}{f} + \dfrac{1}{f+1}}$	$K_f + \beta_m \dfrac{1}{\dfrac{1}{f+1} + \dfrac{1}{f+1}}$	$K_f + \beta_m \dfrac{1}{\dfrac{1}{f+4} + \dfrac{1}{f+1}}$
False, repeated	—	$K_f + \beta_m \dfrac{1}{\dfrac{1}{f+1} + \dfrac{1}{f+1}}$	$K_f + \beta_m \dfrac{1}{\dfrac{1}{f+4} + \dfrac{1}{f+1}}$
Fantasy fact, not repeated	—	$K_m + a(1+p)$	$K_m + a(2.5+p)$
Fantasy fact, repeated	—	$K_m + a(1+p/r)$	$K_m + a(2.5+p/r)$
Actual fact, not repeated	$K_m + \dfrac{1}{x} ap$	$K_m + \dfrac{1}{x} a(p+.5)$	$K_m + \dfrac{1}{x} a(p+2)$
Actual fact, repeated	—	$K_m + \dfrac{1}{x} a(p+.5r)$	$K_m + \dfrac{1}{x} a(p+2r)$
Pure fantasy test			
False, not repeated		$K_f + \beta_e \dfrac{1}{\dfrac{1}{f+1} + \dfrac{1}{f+1}}$	$K_f + \beta_e \dfrac{1}{\dfrac{1}{f+4} + \dfrac{1}{f+1}}$
False, repeated		$K_f + \beta_e \dfrac{1}{\dfrac{1}{f+1} + \dfrac{1}{f+1}}$	$K_f + \beta_e \dfrac{1}{\dfrac{1}{f+4} + \dfrac{1}{f+1}}$
Fantasy fact, not repeated		$K_e + a(1+p)$	$K_e + a(2.5+p)$
Fantasy fact, repeated		$K_e + a(1+p/r)$	$K_e + a(2.5+p/r)$
Pure actual test			
False, not repeated	$K_f + \beta_a \dfrac{1}{\dfrac{1}{f} + \dfrac{1}{f+1}}$	$K_f + \beta_a \dfrac{1}{\dfrac{1}{f+1} + \dfrac{1}{f+1}}$	$K_f + \beta_a \dfrac{1}{\dfrac{1}{f+4} + \dfrac{1}{f+1}}$
False, repeated	—	$K_f + \beta_a \dfrac{1}{\dfrac{1}{f+1} + \dfrac{1}{f+1}}$	$K_f + \beta_a \dfrac{1}{\dfrac{1}{f+4} + \dfrac{1}{f+1}}$
Actual fact, not repeated	$K_a + \dfrac{1}{x} ap$	$K_a + \dfrac{1}{x} a(p+.5)$	$K_a + \dfrac{1}{x} a(p+2)$
Actual fact, repeated	—	$K_a + \dfrac{1}{x} a(p+.5r)$	$K_a + \dfrac{1}{x} a(p+2r)$

TABLE 5

PARAMETER ESTIMATES FOR THE ACT MODEL

Description	Parameter	Estimate
Strength of repeated artificial fact	r	1.3
Strength of actual fact tested	x	2.4
Total strength of actual facts	p	4.8
Search rate	a	120 msec
Average number of actual facts	f	3.9
Falsification rate, mixed test	β_m	235 msec
Falsification rate, fantasy test	β_e	197 msec
Falsification rate, actual test	β_a	76 msec
True constant, mixed test	K_m	1212 msec
True constant, fantasy test	K_e	951 msec
True constant, actual test	K_a	1112 msec
False constant	K_f	1383 msec

The rates R_s and R_p depend on the strength of the links leading toward intersection, (that is, the links making up the proposition being sought) relative to the total strength of all links attached to the source subject and predicate. For example, if the strength of links making up the target proposition is s, and the subject has links connected to it whose strengths total D_s, then $R_s = s/D_s$. We calculate the relevant strengths as follows. The strength of links making up a nonrepeated fantasy proposition is arbitrarily set to 1 to establish scale. A repeated fantasy proposition has links of strength r, and any actual fact has links of strength x. A parameter p was estimated as the total strength of all links, other than those representing fantasy facts, attached to a predicate or individual.

We can now calculate times for intersection to occur for various kinds of probes in terms of these parameters. For example, consider a fan-four, repeated fantasy fact. The strength of links connected to the subject is $4r$ for the links representing the four repeated fantasy facts plus p for all other links, for a total of $p + 4r$. The link that leads toward intersection is part of one of the repeated fantasy facts, and so has strength r. So $R_s = r/(4r + p)$. There is only one fantasy fact attached to the predicate, so similarly $R_p = r/(r + p)$. Plugging these rates into our approximation of time to intersection gives

$$a \left(.5 \, \frac{4r + p}{r} + .5 \, \frac{r + p}{r} \right),$$

which simplifies to $a(2.5 + p/r)$. Times for other kinds of probes are computed in the same manner.

We now must consider the time taken to check an intersection to see if the response true is appropriate. One thing this checking process must

make sure of is that any modifier connected to a candidate path in memory does not conflict with the proposition being sought. For example, in verifying the proposition "George Washington was a woman" one would not wish to respond *true* if a structure representing "It is false that George Washington is a woman" were found. We will assume that the time required to perform checking varies according to the conditions of test; this is a plausible assumption since there are presumably fewer possibilities that need to be checked in a pure test than in a mixed. In this way we will account for the speed differences between tests, given that the time for an intersection to be found, as computed above, depends only on the kind of item being verified and not on test type. Three separate "intercept" parameters were used, K_m, K_e, and K_a, representing the time required for encoding, checking, and responding true in the mixed, pure fantasy, and pure actual tests, respectively.

We can now calculate the total response time for a true sentence by adding time to find an intersection to the lumped estimate for encoding, checking, and responding. For example, if the fan-four, repeated fantasy fact considered above appears in the pure fantasy test, the total response time is calculated to be $K_e + a(2.5 + p/r)$.

The ACT model for falsification is not as well developed as for verification. It is assumed that a waiting process is set up and if there is not an intersection of activation after a certain period, a false response is initiated. It is further postulated that this waiting process is sensitive to variables like fan which affect the time for activation to intersect. So for probes of higher fan, the process will wait longer before initiating the false response. Let f_1 be the fan associated with the subject, and f_2 the fan associated with the predicate. Then, in Table 4 we assumed that the waiting time would be: $(1/(1/f_1 + 1/f_2))\beta$, where β establishes the time scale. This is the waiting time one would get if the waiting time were determined by the fastest of two exponential waiting processes — one associated with the subject and with mean time $f_1\beta$ and the other associated with the predicate with mean time $f_2\beta$. The HAM model (Anderson & Bower, 1973) proposed that the falsification times were determined by a race between processes with independent exponential time distributions. However, we do not mean to endorse an exponential distribution of waiting times. We regard this equation as somewhat arbitrary and chose it because it had been used in fitting earlier fact-retrieval experiments (e.g., Anderson, 1974; Thorndyke & Bower, 1974). Many expressions relating reaction time to the fan of subject and predicate would have done as well. For instance, we might have used $\beta(f_1 + f_2)$.

To derive predictions for the false probes required that four additional parameters be estimated. We needed to know f, the mean number of pre-experimental links or number of actual facts. Also we estimated three

values of β for the three test types; β_m for the mixed test, β_e for the pure fantasy test, and β_a for the pure actual test. It seemed reasonable to estimate a different parameter for each test because the mean processing time for true probes would be different in each test. The waiting process should adjust the amount of time it waited to reflect this fact.

To see how predictions for rejection time were computed, consider a fan-four, repeated item occurring in the mixed test. The fan from the subject is $f + 4$ and from the predicate $f + 1$. Note that only the *number* of links enters here, not their strength, so the fact that the subject occurred in repeated facts does not matter. The waiting time is now given by

$$\beta_m \left(\frac{1}{\dfrac{1}{f+4} + \dfrac{1}{f+1}} \right) ,$$

and the total response time is given by adding K_f, the time to respond *false*.

The ACT model requires the estimation of 12 parameters, in contrast to the nine parameters for the decision interference model. The parameters estimated by the STEPIT program are given in Table 5. The predictions of the ACT model are shown in Fig. 7. As can be seen by comparison with Fig. 3, ACT does a good job of fitting the data. The predictions account for 98% of the variance. The variance not accounted for is not significant [$F(20,899) = .500$].

Although the fit of the model is quite good some of the parameters estimated are a bit bizarre. Note that the total strength of links other than those in fantasy facts p, is 4.8 which is just twice the strength of one of the actual facts tested ($x = 2.4$). Given that there were two actual facts tested about each individual this means that these facts took up all the pre-experimental strength. Actually, the STEPIT program wanted to estimate a value of p less than 4.8, but we constrained the value of p to be at least $2x$. Second, the value of f, the number of actual facts, is 3.9. To avoid an out-and-out contradiction one could imagine that the actual facts tested were so salient, particularly over the course of repeated testing, that they accounted for nearly all the pre-experimental strength. It is also possible that more sensible values of x, p, and f could have been estimated without much changing the goodness of the fit. That is, with 12 parameters it is difficult to know how sensitive the estimation procedure is to each parameter value. Another possibility is that the adequacy of the simplified representation in Fig. 1 as an approximation is being strained, and that more complex network structures must be considered. For instance, Anderson (1976, Section 8.3) proposed an hierarchical
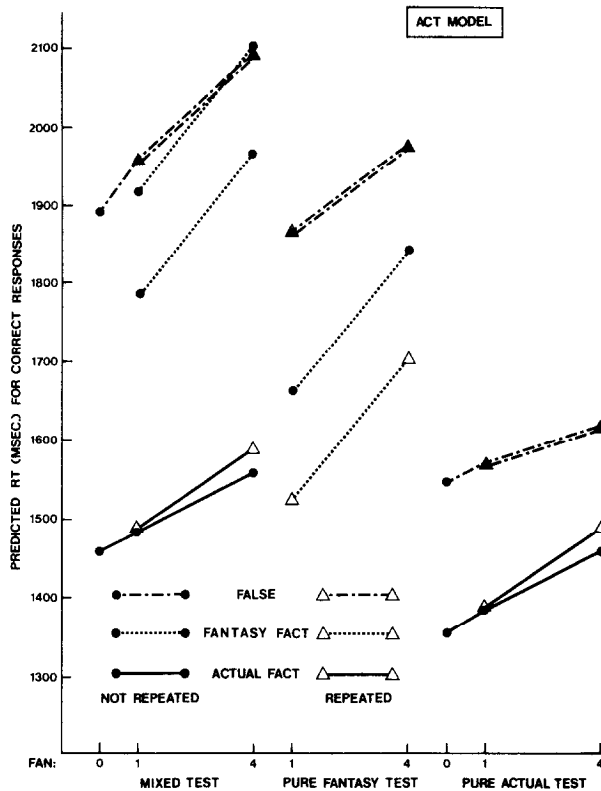
FIG. 7. Predictions of the ACT Model—to be compared with Figure 3.

structure in which a separate node would be set up to collect experimental facts about *George Washington*. Such a model begins to look like a mixture of the ACT model described here and the decision inference model.

As noted above we have so far ignored accuracy data: We have concentrated on establishing that these models are able to capture some of the principle latency effects. We will now consider how the models might be elaborated to account for accuracy effects as well.

In discussing the Experiment I results the suggestion was made that failures in learning might account for some of the misses in verifying fantasy facts. This is apparently not an important factor, in the light of the Experiment II results, since accuracy for fantasy facts in the pure fantasy test is quite high. So what does account for the low accuracy in the mixed condition? A look at Fig. 5 reveals an orderly relation between accuracy and processing time: longer latencies go with lower accuracy. As with the Experiment I results this suggests a race model for errors, such as that proposed by Mohs *et al.*, in which errors occur when a

guessing process finishes before normal handling of an item is completed. Conditions in which normal processing takes longer will therefore show more errors.

The effect of the race with a guess process is to systematically change the observed distribution of reaction times from the "true" distribution of reaction times that would have been obtained were it not for the guess process. Under reasonable assumptions about the distributions of processing and guessing times the racing guess process tends to truncate the distribution of latencies since trials for which normal processing takes a long time tend to be turned into errors by the prior completion of the guess process. Thus the observed mean reaction times are less than the "true" mean reaction times. However, the ordinal relations among the mean reaction times for the various conditions are preserved. The differences among conditions are also reduced because of this guess factor. The ordinal relation among the differences are not always preserved. The differences among longer reaction times tend to be more greatly reduced. Thus, the effect of this guess process is to produce a "shrinkage" of the true graph of reaction times into an observed graph like Fig. 3. This shrinkage is greater for longer reaction times.

Despite this shrinkage, it seems likely that either the ACT or decision inference models could have been fit to the true times about as well as the observed times. This is because the shrinkage preserves all ordinal relations among means and does not wildly alter ordinal relations among differences. The fit of the models to the observed times gives a good indication of the fit to the true times, could that have been obtained. The decision interference model might have given a somewhat improved fit to the true data since it did overpredict the size of some of the observed differences at long reaction times. More thorough investigation (such as actual model fitting) of the effects of extending the ACT and decision interference models to include a racing guess process would require consideration of a number of further points including biasing of guesses, possible changes in the guess processes under different test conditions, and the portion of normal processing with which the guess process is to race. We believe more evidence would be required to adequately motivate choices about these matters.

SUMMARY AND CONCLUSIONS

These two experiments have uncovered two facts which we feel have important implications for the retrieval of information from memory. First, it was shown the more new facts learned about an individual the slower subjects were to retrieve facts already well learned about the individual. Second, it was shown that subjects were slower when the range

of possible facts they must verify was larger. That is, they took longer in the mixed than the pure fantasy test. Thus, we have found evidence for two distinct types of interference to the retrieval process.

Two models were presented to account for this pattern of data. The decision interference model proposed that the well established preexperimental facts were stored in one location and the fantasy facts were stored in another. The effect of number of fantasy facts on actual facts was due to subjects sometimes searching the fantasy facts before the pre-experimental. The advantage of the pure testing context arose because subjects would more often search the correct location first.

The ACT model proposed that the information was stored together. The effect of fantasy facts on pre-experimental was due to a slowing down of the activation rate. The advantage of a pure testing context was attributed to the greater efficiency of the process that checked the path of intersection.

Neither of the two models gave totally satisfactory quantitative fits. The decision–interference model mispredicted some important quantitative relations while some of the parameters of the ACT model took on somewhat bizarre values. We suspect the deficiencies of both models could be remedied with further complications, but such complications did not seem justified. Each model accounts naturally for one of the main two interference effects, but had to be modified to accommodate the other. This serves to indicate the importance of these two effects to a theory about the retrieval of real-world knowledge from long-term memory.

APPENDIX

Facts Used as Materials

GEORGE WASHINGTON CROSSED THE DELAWARE.
GEORGE WASHINGTON CHOPPED DOWN A CHERRY TREE.
NAPOLEON BONAPARTE WAS FRENCH.
NAPOLEON BONAPARTE WAS AN EMPEROR.
JULIUS CAESAR WAS A ROMAN LEADER.
JULIUS CAESAR WAS ASSASSINATED.
FIDEL CASTRO IS CUBAN.
FIDEL CASTRO IS A COMMUNIST.
CHRISTOPHER COLUMBUS SAILED TO THE NEW WORLD.
CHRISTOPHER COLUMBUS WAS AN EXPLORER.
ABRAHAM LINCOLN FREED THE SLAVES.
ABRAHAM LINCOLN GAVE THE GETTYSBURG ADDRESS.
JOSEF STALIN WAS A RUSSIAN.
JOSEF STALIN WAS A DICTATOR.

HENRY AARON PLAYS BASEBALL.
HENRY AARON BROKE A SPORTS RECORD.
GALILEO GALILEI WAS A SCIENTIST.
GALILEO GALILEI WAS ITALIAN.
BEBE REBOZO LIVES IN FLORIDA.
BEBE REBOZO IS NIXON'S RICH FRIEND.
WALT DISNEY MADE MOVIES.
WALT DISNEY CREATED MICKEY MOUSE.
GOLDA MEIR IS ISRAELI.
GOLDA MEIR IS A WOMAN.
SPIRO AGNEW WAS VICE PRESIDENT.
SPIRO AGNEW RESIGNED HIS OFFICE.
ELIZABETH TAYLOR IS AN ACTRESS.
ELIZABETH TAYLOR MARRIED RICHARD BURTON.
LYNDON JOHNSON HAD A RANCH.
LYNDON JOHNSON WAS A TEXAN POLITICIAN.
JOHN LENNON IS ENGLISH.
JOHN LENNON IS A SINGER.
MARK SPITZ IS A SWIMMER.
MARK SPITZ WON OLYMPIC MEDALS.
MARK TWAIN WROTE *TOM SAWYER*.
MARK TWAIN WAS A HUMORIST.
TEDDY KENNEDY IS A LIBERAL SENATOR.
TEDDY KENNEDY IS FROM MASSACHUSETTS.
MAHATMA GANDHI WAS FROM INDIA.
MAHATMA GANDHI WAS A NONVIOLENT LEADER.

REFERENCES

Anderson, J. R. *Language, memory, and thought.* Hillsdale, NJ: Erlbaum Associates, 1976, in press.
Anderson, J. R. Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 1974, **5**, 451–474.
Anderson, J. R., & Bower, G. H. *Human associative memory.* Washington: Winston, 1973.
Chandler, J. P. STEPIT. Program 90PE66, Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN, 1965.
Collins, A. M., & Loftus, E. F. A spreading-activation theory of semantic processing. *Psychological Review*, 1975, **82**, 407–428.
Kieras, D. E. *Analysis of the effects of word properties and limited reading time in a sentence comprehension and verification task.* Unpublished doctoral dissertation, University of Michigan, 1974.
King, D. R. W., & Anderson, J. R. Long-term memory search: An intersecting activation model. In preparation.
Mohs, R. C., Westcourt, K. R. & Atkinson, R. C. Search processes for associative structures in long-term memory. *Journal of Experimental Psychology: General*, 1975, 103–121.
Perlmutter, J., Harsip, J., & Myers, J. L. The role of semantic knowledge in retrieval

from episodic long-term memories: Implications for a model of retrieval. *Memory & Cognition*, 1976, in press.

Postman, L. Extra-experimental interference and the retention of words. *Journal of Experimental Psychology*, 1961, **61**, 97–110.

Slamecka, N. J. Differentiation versus unlearning of verbal associations. *Journal of Experimental Psychology*, 1966, **71**, 822–828.

Thorndyke, P., & Bower, G. Storage and retrieval processes in sentence memory. *Cognitive Psychology*, 1974, **5**, 515–543.

Townsend, J. T. 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: Lawrence Erlbaum Associates, 1974.

Underwood, B. J., & Ekstrand, B. R. An analysis of some shortcomings in the interference theory of forgetting. *Psychological Review*, 1966, **73**, 540–549.

Underwood, B. J., & Postman, L. Extraexperimental sources of interference in forgetting. *Psychological Review*, 1960, **67**, 73–95.