Journal of Experimental Psychology:
Learning, Memory, and Cognition
1985, Vol. 11, No. 1, 136–153.

0278-7393/85/$00.75

The Role of Practice in Fact Retrieval

Peter L. Pirolli and John R. Anderson
Carnegie-Mellon University

Four experiments examined the relationship among massive practice, speed-up of memory retrieval, and the reduction of long-term memory interference. All experiments employed a multiday fact recognition procedure. Interference was varied by a fan manipulation: Concepts could appear in more than one fact (fan) or only one fact (no fan). We found that fact retrieval speeds up as a power function of days of practice but that the number of daily repetitions beyond four produced little or no impact on reaction time. In addition, interference decreased in proportion to the degree of practice but did not disappear even with 25 days of practice. Practice on specific facts and practice on the general task had multiplicative effects in reducing recognition time. Further, general task practice was found to decrease interference, suggesting that general central processes were speeding up. Most but not all of these practice effects and their interactions with interference are predicted by the ACT* model of fact retrieval.

There are many experimental demonstrations in the psychological literature of the old adage "practice makes perfect." A major effect of practice is a continual speed-up in performance. An apparently ubiquitous feature of this speed-up is the power-law relationship between performance time (T) and practice (P), $T = P^{-b}$, where b is in the range of 0 to 1 (Lewis, 1979; Newell & Rosenbloom, 1981). Such a power function implies a continuous improvement in performance with practice.

Typically, such power-function reaction-time improvements have been demonstrated with perceptual-motor skills such as pointing (Card, Moran, & Newell, 1983) and cognitive skills such as geometry problem solving (Neves & Anderson, 1981). Research in the area of fact memory has been less concerned with what happens to our access to a fact as it is repeated many times. Frequently when practice effects have been acknowledged in fact learning they are regarded as peripheral and as not affecting factors of fundamental theoretical importance. This attitude is manifest in a large number of studies looking at memory retrieval (e.g., Anderson, 1976; Shoben, Wescourt, & Smith, 1978). Indeed, the typical procedure is to report results averaged over repeated tests.

In certain situations, repeated testing is thought to result in qualitative reorganization of the knowledge. For instance, questions formerly answered by inference may be answered by direct retrieval upon repetition. Mohs, Wescourt, and Atkinson (1975) suggested that subjects transform a list organization to a paired-associate organization with extensive practice. Schneider and Shiffrin (1977) argued that short-term memory search changes from a controlled serial process to an automatic parallel process. Such two-state qualitative changes are not easily made compatible with the continuous change implied by the power function associated with practice, unless one assumes that there is a gradually changing mixture of the two states.

Hayes-Roth (1977) proposed a two-state, *knowledge-assembly theory* in which practice reduces both interference and retrieval time in fact memory. The theory is formulated within a spreading-activation framework (e.g., Collins & Loftus, 1975). It assumes that facts are represented as propositional traces, each trace consisting of concept nodes associated by relational links. A proposition is retrieved

This research was supported by Grants BNS 78-17463 and BNS 82-08189 from the National Science Foundation to John Anderson. Preparation of this manuscript was in part supported by an IBM Research Fellowship to Peter Pirolli. We would like to thank Al Corbett and Lynne Reder for comments on earlier drafts of this article. We also thank N. J. Slamecka for pointing out the relevance of Postman's (1969, 1982) work to our own.

Requests for reprints should be sent to Peter L. Pirolli, Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213.

when all of its constituent nodes and links are activated. Knowledge-assembly theory (and other spreading-activation theories, e.g., Anderson, 1976) addresses interference by proposing that activation spreading from a concept node is divided up among its associates. Increases in the number of associates of a concept decrease the amount of activation spread to any particular associate. The number of associative links emanating from a concept node is often referred to as the *fan* of a concept.

According to knowledge-assembly theory, a fact trace may initially be subject to interference because its constituent concepts are associated with many other facts. However, traces, concept nodes, and links vary in strength as a function of their frequency and recency of activation. With repetition, the strength of a trace reaches some threshold and the trace becomes *unitized.* A unitized trace behaves as a discrete concept in and of itself. A stimulus pattern (e.g., a sentence) corresponding to a unitized trace can activate the trace without activating the trace's former constituent concepts. Thus, a unitized trace is not affected by the interference associated with its former constituent concepts.

Hayes-Roth (1977) performed an experiment supporting the knowledge-assembly theory prediction that interference disappears with repetition. Hayes-Roth's subjects learned fact sentences of the form *The doctor hated the lawyer,* and interference was varied using a *fan* manipulation. The fan of the sentence subjects (e.g., doctor) was varied from one to six. That is, subject concepts could occur in from one to six different sentences. This manipulation typically produces a *fan effect* (e.g., Anderson, 1976): A fact containing concepts related to many facts produces longer reaction times than a fact containing concepts associated to fewer facts. For example, recognition time for the above "doctor" fact would generally increase as one added other "doctor" facts to memory. After learning sentences, Hayes-Roth's subjects went through 10 daily recognition reaction time sessions in which each item was repeated 10 times per day. Day 1 recognition times showed a fan effect. Reaction times to targets increased as a function of the number of other facts related to a target subject. However, on Day 10 there was no effect of fact fan on reaction times. Thus, the fan effect seems to disappear with massive practice.

Hayes-Roth's work has been quite influential in current conceptions of interference and memory retrieval (e.g., see Smith, 1981). Her results indicate that repetition can produce qualitative changes in the nature of memory. However, it is not clear how to relate her results to the anticipated power-function speed-up. Unfortunately, her data are no longer available in a form that would allow us to trace out the improvement over time. Experiment 1 tracks recognition time and the fan effect over the course of 25 days. Certain aspects of this experiment have been briefly summarized elsewhere (Anderson, 1983a, 1983b), but here we present the procedures, results, and a mathematical model in greater detail.

Experiment 1

To our knowledge, very few studies of long-term memory have been performed that track memory performance over an extended practice period. The Hayes-Roth (1977) study is one exception to this scarcity. However, Hayes-Roth only compared data from the first and final sessions of a 10-day experiment. We sought to determine the character of the speed-up in memory retrieval and the reduction of the fan effect with practice by tracking memory performance over an entire 25-day period.

Subjects in the current experiment initially learned fan and no-fan sentences. Fan sentences contained subjects, verbs, and objects that appeared in two sentences, whereas no-fan sentences contained subjects, verbs, and objects that appeared uniquely in one sentence. The current experiment was also an attempt to determine whether decreases in recognition time were strictly a function of number of item repetitions. Previous studies of the effects of massed versus spaced practice (see Crowder, 1976, chap. 9) have suggested that the facilitative effects of item repetition on memory performance decrease as one increases the rate of repetition. There were two groups of subjects in the current experiment. One group received 12 presentations of items daily, and the other group received

24 presentations per day. If retrieval speed-up is strictly a function of number of retrieval attempts, then we should expect that 24-repetitions/day subjects should be speeding up faster than 12-repetitions/day subjects in recognition reaction time.

Method

Subjects. Eight Carnegie-Mellon work-study students participated in the experiment for pay ($3/hr with bonus for accuracy and speed performance). Four were randomly assigned to the condition of 12 repetitions per day, and 4 to the condition of 24 repetitions per day. However, 1 subject in the 24-repetition condition had to quit after the 4th day, and his data were not analyzed.

Materials. For each subject, 16 target sentences of the form *The doctor hated the lawyer* were constructed by placing random selections from a pool of nouns and verbs into a sentence frame. Nouns were profession names (e.g., doctor, lawyer) ranging from four to eight letters in length, and past tense verbs (e.g., hired, flattered) ranged from five to nine letters in length. Eight targets were no-fan sentences in which each subject, verb, or object occurred uniquely within that sentence. The other eight targets were fan sentences in which each subject, verb, or object occurred in two sentences. The fan targets were constructed with the constraint that no pairwise combination of subject, verb, or object occurring in one sentence could occur in another. A pool of foils was created by recombining subjects, verbs, and objects from targets. There were three types of foils: SV foils, which combined the subject and verb from one target with the object of another; SO foils, combining subject and object from a target with the verb of another; and VO probes containing the verb and object of one target and the subject of another. The target words selected to form a foil could come from fan or no-fan items. Thus, each set of SV, SO, and VO foils could be divided into an equal number of 1-1, 1-2, 2-1, and 2-2 foils, where the first digit indicates the fan of the target from which the single foil word was selected and the second digit indicates the fan of the target from which the pair of foil words was selected.

Apparatus. All phases of this and the following experiments were presented via computer terminals controlled by a PDP 11/34 computer running under the YEPS (Proudfoot, 1978) extended UNIX system.

Procedure. Subjects in this and subsequent experiments participated in three tasks: (a) *study,* in which each target was displayed for study, (b) *test,* in which subjects were tested on targets with a cued-recall procedure and given feedback until they reached a criterion memory performance, and (c) *recognition reaction time,* which consisted of a yes–no recognition task. The goal of the study and test procedures was to ensure that subjects had committed target facts to long-term memory, whereas the recognition phase tested retrieval of those items. In this experiment, subjects studied and were tested on target materials in the first day's session. They participated in a reaction time phase in all 25 sessions. The 25 sessions took place only on weekdays, so subjects had four, 2-day weekend breaks over the course of the experiment. The study phase during Session 1 consisted of presenting subjects with each of the 16 target sentences in random order for 15 s each. Following the study phase, subjects were tested on targets with a dropout procedure. Each target appeared with a subject, verb, or object missing, and subjects had to type in the missing word. Each target and missing word defined a *query* for a total of 16 × 3 or 48 queries. This pool of queries was repeatedly presented, and on each iteration queries were presented in random order. A query was dropped from the query pool when it had been answered correctly twice.

In the reaction time phase, subjects were instructed to indicate whether a presented item was a target or foil. Blocks of 96 trials were presented to subjects with brief breaks between blocks. Subjects in the 12-repetitions group received four blocks of trials, whereas the 24-repetitions group received eight blocks. In each block, each target was presented three times. There were also 16 presentations each of SV, SO, and VO, foils. On each foil presentation a foil of the appropriate type was randomly selected from the foil pool and presented to the subject. Thus, foils were not explicitly repeated. On each trial, subjects were presented with a single sentence in the center of their terminal screen and were to respond by depressing "k" on their keyboards to targets and "d" to foils. Speed and accuracy feedback was provided on each trial, and subjects were motivated by a points system to be as rapid as possible without sacrificing accuracy.

Results and Discussion

For each subject on each day, a mean reaction time was computed for targets and foils. Recall that there were two types of target (fan and no fan) and 3 (SV, VO, SO) × 4 (1-1, 1-2, 2-1, 2-2), or 12, types of foil. Preliminary analyses revealed that reaction times to SV, VO, and SO foils (mean reaction times were 888, 869, and 872 ms, respectively) did not reliably differ nor enter into any significant interactions. Consequently we collapsed subject data over this factor in subsequent analyses. In addition, there were no reliable reaction time differences between 1-1, 1-2, and 2-1 foils (mean reaction times were 798, 807, and 828 ms, respectively). Subject data for these foils were pooled and considered as no-fan foil, and the 2-2 foils (M = 1,071 ms) were renamed as fan foils. The mean error rate (false alarms and misses) was 3.7%.

Mean reaction times to fan and no-fan targets are presented in Figure 1. The presented data are collapsed over the 12- and 24-repetition/day groups. The decrease in reaction times evident in Figure 1 resembles power-function decreases in performance time

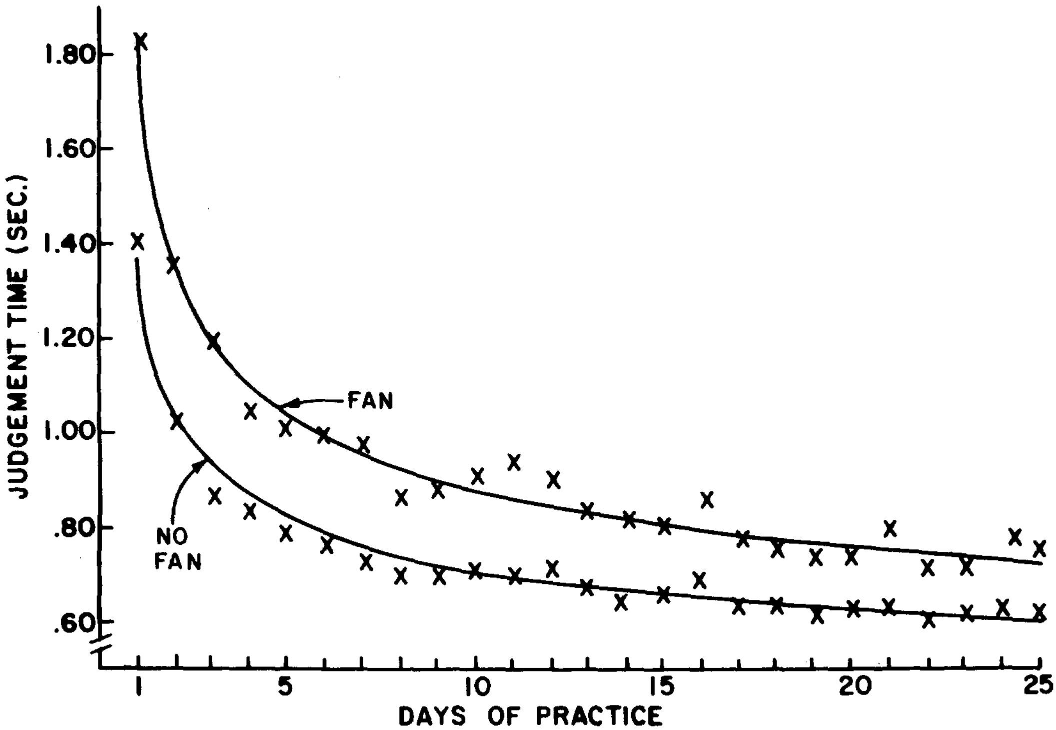

Figure 1. Practice effects on reaction time and interference in Experiment 1.

found in other studies of practice. An analysis of variance consisting of the between-subjects factor of practice (24 repetitions/day, 12 repetitions/day) and the within-subjects factors of fan (fan, no-fan), by days (1 to 25), and item (target, foil) was conducted on the reaction time data. Reaction times in the 12-repetitions group (M = 860 ms) did not reliably differ from those in the 24-repetitions group (M = 902 ms), $F(1, 5) = .10$, MS_e = 2,920. This curious lack of a practice effect is generally consistent with findings from the massed versus spaced practice literature (e.g., Hintzman, 1969). However, it should be noted that the majority of such studies (cf. Crowder, 1976; Hintzman, 1974) varied the frequency of item repetition within a *fixed* period of time to achieve spacing effects. In the current experiment, subjects getting 24 repetitions per day received their first 12 item repetitions at the same rate over the same period, on average, as the 12-repetitions/day subjects. The additional 12 fact repetitions provided no reliable facilitative effect on memory.

The above analysis also indicated that the decrease in reaction times from Day 1 (M = 1,758 ms) to Day 25 (M = 723 ms) was significant, $F(24, 120) = 32.44$, $MS_e = 440$, $p < .01$. As has been found in other similar studies (e.g., Anderson, 1976), targets were responded to faster (M = 774 ms) than foils (M = 983 ms), $F(1, 5) = 10.29$, $MS_e = 800$, $p < .05$, and this main effect did not enter into any reliable interactions.

We also found a fan effect: Reaction times to fan items were significantly slower (M = 939 ms) than those to no-fan items (M = 817 ms), $F(1, 5) = 39.92$, $MS_e = 60$, $p < .01$. This effect indicates that retrieval times increase with the number of facts related to a concept. Of special interest to us was determining if this fan effect was attenuated by massive practice. In Figure 1, reaction times to fan and no-fan targets reduce and converge with days of practice. The fan effect (fan minus no-fan reaction time collapsed across targets and foils and groups) was 250 ms on Day 1 and 80 ms on Day 25. This Fan × Days interaction was reliable, $F(24, 120) = 10.47$, $MS_e = 1$, $p < .01$. Unlike the Hayes-

Roth (1977) finding of a complete reduction of the fan effect over 10 days (100 target repetitions), it is apparent that a significant fan effect remained in our experiment after 25 days (300 or 600 target repetitions), $t(120) = 8.75$, $p < .01$.[1] All seven subjects were slower in the fan condition than the no-fan condition on Day 25.

An ACT Model Relating Practice and Interference*

The current fact retrieval results appear to follow a power law relating time (T) to practice (P). This power law is predicted by the ACT* theory (Anderson, 1982, pp. 398–400) strength and decay mechanisms. ACT* is a *production system* theory (Anderson, 1983a). The general model for the fact retrieval task involves representing facts as traces stored in a propositional network with processes represented as condition-action production rules operating on those facts. Responses are generated by an interaction of a spreading-activation mechanism and a pattern-matching mechanism. The spreading-activation mechanism retrieves propositions from long-term memory into working memory. The pattern-matching mechanism can be thought of as a comparison process, and it matches the conditions of production rules to active information in working memory. These production rules determine the sequencing of cognitive events. When a production rule is selected by the pattern-matching mechanism, the actions specified by that rule are executed. These actions include additions to working memory, including calls to response generation routines, and the setting of new goals or subgoals. In the current paradigm, we assume that there are a set of production rules that test whether a given probe corresponds to a long-term memory proposition (see Anderson, 1983a, pp. 107–114, for details regarding a similar paradigm).

The operation times of the spreading-activation and pattern-matching mechanisms are not additive. It is assumed that higher levels of network activation produce more rapid pattern matching. One could think of the pattern matcher as an engine, and network activation as the flow of fuel to that engine; as the flow of activation is increased, the pattern-matching engine works faster.

According to Anderson (1982, pp. 398–400), the ACT* reaction time equation for the fact retrieval task is

$$\mathrm{RT} = I + FD/QP, \tag{1}$$

where I represents a lower bound on performance time due to constraints in the task environment, F is the fan of the fact, Q is the strength of the specific fact being matched, P is the strength of the retrieval production rule, and D is the multiplicative constant representing the time to match a one-fan fact of one unit strength, by a one-unit-strength production.

Suppose the production has been matched G times and the fact S times. Then, assuming power-law increases in strength, we can replace Q by S^{-e} and P by G^{-f}. Thus, the expression in Equation 1 may be written

$$\mathrm{RT} = I + FD/S^eG^f = I + BS^{-e}G^{-f}, \tag{2}$$

where $B = FD$. The component S^{-e} represents the specific practice of the fact and the component G^{-f} represents general practice of the fact retrieval procedure. For the current experiment we can simplify Equation 2 even further, because the rate of general practice is a multiple of the rate of specific practice. That is, for every S repetitions of a specific fact there will be $G = nS$ trials of the task. Therefore the term $BS^{-e}G^{-f}$ can be expressed as $n^{-f}BS^{-(e+f)}$, and letting $B' = n^{-f}B$ and $g = e + f$, Equation 2 becomes

$$\mathrm{RT} = I + B'S^{-g}. \tag{3}$$

We fit Equation 3 (using STEPIT, Chandler, 1965) to the data presented in Figure 1. In doing so we considered S in terms of days and used $S - ½$ to represent the fact that each data point in Figure 1 is an average over a day's session. Assuming separate I, B', and g parameters for the fan and no-fan conditions, there are six potential parameters to be estimated in fitting the data in Figure 1. The I parameter reflects the asymptote that the conditions will reach in the limit, g the rate of speed-up, and B' the amount of time that can be improved with practice (we

[1] All t tests reported in this paper are linear contrasts using the mean square error term of the factor or interaction of interest as an estimate of the variance.

Table 1
Best-Fitting Parameter Estimates of the Power-Function $RT = I + B'(S - ½)^{-g}$ to Experiment 1 Reaction Times

Parameters fixed across conditions	Best fit equations	χ^2	Adjusted r^2
All independent	$T_F = .44 + 1.06(P - ½)^{-.39}$ $T_{NF} = .50 + .63(P - ½)^{-.52}$	78.97	.985
I	$T_F = .48 + 1.01(P - ½)^{-.42}$ $T_{NF} = .48 + .65(P - ½)^{-.50}$	80.29	.985
B'	$T_F = .61 + .85(P - ½)^{-.52}$ $T_{NF} = .32 + .85(P - ½)^{-.35}$	129.92	.975
g	$T_F = .48 + 1.02(P - ½)^{-.42}$ $T_{NF} = .42 + .72(P - ½)^{-.42}$	86.38	.984
I, B'	$T_F = .34 + .99(P - ½)^{-.27}$ $T_{NF} = .34 + .99(P - ½)^{-.43}$	619.92	.885
I, g	$T_F = .36 + 1.14(P - ½)^{-.36}$ $T_{NF} = .36 + .77(P - ½)^{-.36}$	100.94	.981
B', g	$T_F = .57 + .84(P - ½)^{-.44}$ $T_{NF} = .38 + .84(P - ½)^{-.44}$	218.05	.960

Note. T_F = time for fan; T_{NF} = time for no-fan.

will refer to this as the *initial time*). We looked at all possible combinations of setting parameters to be equal in the fan and no-fan conditions. The results are reported in Table 1. There we have reported the equations estimated along with two measures of goodness of fit. The first is a chi-square statistic:

$$\chi^2 = \sum_i (\bar{x}_i - \hat{x}_i)^2/S^2_{\bar{x}}, \tag{4}$$

where $\bar{x}_i$ is the observed mean, $\hat{x}_i$ is the predicted mean, and $S^2_{\bar{x}}$ is the standard error of the means. This chi-square statistic has 50 degrees of freedom less the number of parameters estimated. The second statistic (from Reed, 1976) is an r^2 correlation adjusted to reflect the number of free parameters used in estimation:

$$r^2 = 1 - \frac{\sum_i(\bar{x}_i - \hat{x}_i)^2/(h - k)}{\sum_i(\bar{x}_i - \bar{X}_i)^2/(h - 1)}, \tag{5}$$

in which h is the number of empirical points, k is the number of free parameters, and $\bar{X}_i$ is the grand mean of the $\bar{x}_i$s.

Inspection of Table 1 reveals a rather clear picture. Reasonable fits can be achieved by setting asymptotes and rates to be the same across the two conditions but allowing for two different initial times across the two conditions. Much worse fits are produced under the assumption of same rate and initial times but different asymptotes or under the assumption of same asymptotes and initial times but different rates. Also the absolute quality of the r^2 fit with asymptote and rates fixed across conditions is not much worse over the case of independent estimates of all parameters.

Figure 1 presents the fits obtained by assuming that the asymptote (I) and rate (g) parameters are the same in the fan and no-fan conditions. The major source of deviation from the predicted function derives from the fact that subjects are slower following a weekend break (see Days 6, 11, 16, and 21). This reflects forgetting over the weekend.

Conclusions from Experiment 1

There are three noteworthy results from this experiment. The first is that memory retrieval appears to obey the same power-law speed-up as does skill performance. The second result is that the fan effect decreases but does not disappear as Hayes-Roth claimed. In fact, our power-law equations indicated that the decrease in the fan effect is directly related to the amount of speed-up. That is, the parameter B' reflecting fan is multiplied by the component $(S - ½)^{-g}$ reflecting speed-up. It is hard to reconcile such a gradual process with the qualitative changes in the memory process envisioned by Hayes-Roth (1977).

The third result is the apparent lack of effect of 12 versus 24 trials of practice per day. This result came as a surprise to us. It is based on comparing only 3 versus 4 subjects, but this lack of difference was obtained in an experiment that found large differences of other variables. Also, it is consistent with later studies in the series. It should also be noted that practice over days has a large effect. Thus, 24 trials over 2 days is much better than 24 in 1 day. The total ineffectiveness of large numbers of practice trials per day has the potential of being an important finding about memory. The finding has its analog in the skill acquisition literature: It has been found that students learn Morse code at the same rate whether they get 4 or 7 hours of training per day (reported in Bray, 1948).

The lack of a beneficial effect of extra repetitions appears to be consistent with the *habituation* account of the spacing effect (Hintzman, 1974; Posner & Warren, 1972). This account states that the facilitating effects of repeating an item are attenuated by habituation, which is assumed to be a basic physiological phenomenon that dissipates with time. Increasing the spacing of repetitions decreases the attenuating effects of habituation on repetition effects. Consistent with this view are results (Hintzman & Rogers, 1973) showing that the spacing effect is relatively independent of whether the interval between repetitions is filled with presentations of other items. Our results seem to indicate that habituation can accumulate over repetitions to the point that it washes out the effects of further repetitions.

Experiment 2

The second experiment was devoted to exploring the replicability of the first two results noted above—the power-function speed-up and the systematic decrease in the fan effect. We decided to explore these phenomena with sentences of a different type: person-location sentences of the form *The doctor is in the bank.* We also wanted to explore what seemed to be a plausible explanation of the difference between our results and those of Hayes-Roth (1977). She manipulated only the fan of the subjects of her sentences. The verbs and objects always occurred uniquely in one sentence. It might be that her subjects came to retrieve from the unique verb and object and so avoided the fan effect. Our material in the first experiment had equal fans for all terms. Therefore we were motivated to look at sentences with asymmetric fan to see if different speed-up functions would be obtained for them.

Method

Subjects. Thirteen subjects participated in what was originally two experiments. The first group of 9 subjects participated for a fee of $3/hr in an experiment consisting of 600 trials daily. The second group of 4 participated for the fee plus a speed-accuracy bonus in an experiment consisting of 900 trials daily.

Materials. Fifteen person-location target facts of the form *The hippie is in the park* were created for each subject. There were four types of target facts: 1-1, 1-3, 3-1, and 2-2, where the first digit indicates the fan of the target subject and the second digit indicates the fan of the target location. Each subject's material consisted of three targets of types 1-1, 1-3, and 3-1, and six targets of type 2-2. We can characterize these sentences as having symmetric or asymmetric fan. A symmetric fan sentence (1-1 or 2-2) has a person and location that appear equally often in experimental facts. An asymmetric fact (1-3 or 3-1) has one concept (person or location) that appears in three facts, whereas the other concept is unique to that fact. The asymmetric fan sentences were analogous to those used by Hayes-Roth (1977), whereas the symmetric fan sentences were analogous to those used in Experiment 1. Sentence subjects were profession names ranging from four to eight letters in length. Similarly, locations (e.g., *church, park*) ranged from four to eight letters in length. Foils were created by randomly selecting from the pool of target persons and locations and creating a nontarget sentence. Foils were randomly generated on each foil trial during the experiment.

Procedure. The experiment extended over a 10-day period with a weekend break between Sessions 5 and 6. On the first day, subjects studied and were tested on targets, and they then performed a recognition reaction time task. This recognition task was repeated on the subsequent 9 days.

In the study phase on Day 1, targets were presented individually for 12 s each on subjects' CRT (cathode ray tube) screens. The test procedure followed the study phase, and it consisted of a cued recall task involving the studied targets. Each target was (a) probed with a person noun (and subjects had to type in all the locations studied with that person) and (b) probed with a location noun (and subjects had to type in all the people studied with that location). Thus, there were 10 location queries and 10 person queries. This pool of queries was repeatedly presented to subjects. At the end of each pass, all correctly answered queries were dropped from the pool before starting the next pass.

The recognition phase, which was conducted over all 10 sessions, served to give subjects practice and to gather their reaction time data. The two groups of subjects received different numbers of fact repetitions per session. One group of subjects received 20 fact repetitions per session, and the other group received 30 repetitions per session. All subjects received an equal number of target and foil trials, which means that one group got 600 trials per day while the other got 900 trials. On each trial, subjects indicated whether a presented sentence was a target or foil. Sentences were presented on CRT screens, and subjects indicated their responses by pressing either a "k" key for targets or a "d" key for foils on their computer terminal. Trials within a session were partitioned into blocks of 60, which permitted subjects to take brief breaks between blocks. The order of presentation of items within any session was random.

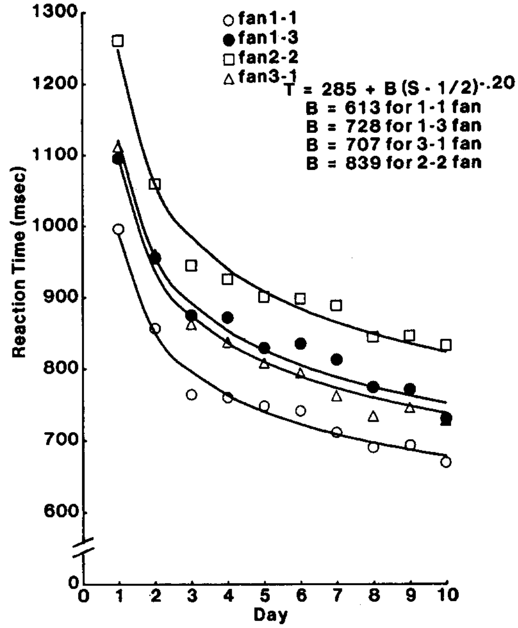

Figure 2. Practice effects on symmetric and asymmetric fan target reaction times in Experiment 2.

Results and Discussion

Figure 2 presents mean reaction times for targets, collapsed over 20- and 30-repetition/day groups. We submitted these target data to an analysis of variance consisting of the between-groups factor of repetitions/day (20, 30) and the within-subjects factors of fan (1-1, 1-3, 3-1, 2-2) and days of practice (1 to 10). First, let us point out how the results of this experiment replicate those of Experiment 1. Mean reaction times for the 20- and 30-repetitions groups were 867 ms and 802 ms, respectively. As in Experiment 1, this repetition contrast was not significant, $F(1, 11) = 0.77$, $MS_e = 630$, $p = .40$, although the marked decrease in reaction times (from M = 1,116 ms to M = 737 ms) over days was highly significant, $F(9, 99) = 40.56$, $MS_e = 20$, $p < .01$. Again we are confronted with this interesting finding which suggests that it is the number of days of practice and not the number of repetitions per se that has the greatest impact on retrieval speed.

It is evident from Figure 2 that the 1-1 fan conditions produced the fastest reaction times on all days, the 1-3 and 3-1 conditions were next, and the 2-2 conditions produced the slowest reaction times on all days. This fan main effect was significant, $F(3, 33) = 13.00$, $MS_e = 40$, $p < .01$. Planned comparisons indicated that the 1-3 and 3-1 fan conditions did not differ significantly, $t(33), = 1.17$, whereas the 1-1 conditions were significantly faster than the pooled 1-3 and 3-1 fan conditions, $t(33) = 5.40$, $p < .01$, and the 2-2 conditions were significantly slower than the pooled 1-3 and 3-1 fan conditions, $t(33) = 6.28$, $p < .01$.

We were especially interested in whether the fan effect would disappear for the asymmetric fan sentences (1-3 and 3-1). The Fan × Days interaction was reliable, $F(27, 297) = 2.42$, $MS_e = 20$, $p < .01$, indicating that the fan effect decreased over the course of 10 days. However, there was still a marked difference among fan conditions on Day 10. The reaction time difference between the 1-1 and pooled 1-3 and 3-1 fan conditions decreased over days, and this interaction of latency differences over 10 days of practice was significant, $t(297) = 4.87$, $p < .01$. Similarly, the reaction time differences between the 1-1 and 2-2 fan conditions decreased over the course of 10 days, $t(297) = 7.07$, $p < .01$, and so did the 2-2 versus pooled 1-3 and 3-1 fan differences, $t(297) = 2.20$, $p < .05$. The mean reaction times on Day 10 were as follows: M = 668 ms for 1-1 fan; M = 729 ms for 1-3 fan; M = 727 ms for 3-1 fan; and M = 831 ms for 2-2 fan. Even on Day 10 the 1-1 fan condition was significantly faster than the pooled 1-3 and 3-1 fan conditions,

$t(297) = 5.67$, $p < .01$, and the combined 1-3 and 3-1 conditions were faster than the 2-2 condition, $t(297) = 9.74$, $p < .01$. More important, we can see that reaction times in the asymmetric and symmetric fan conditions appear to follow the same pattern over days. We did not see the asymmetric 1-3 and 3-1 reaction times approach the levels of those in the 1-1 conditions as might be expected on the basis of Hayes-Roth's (1977) results.

The best-fitting parameter values of Equation 3 (using STEPIT), assuming a constant intercept and rate across conditions, are presented in Figure 2. The chi-square statistic, with 36 degrees of freedom, for the best-fitting parameters was 66.68, which is not significant.

Given the difference in materials, the results of the two experiments are remarkable in their similarity. We still do not have any compelling explanation of the difference between our results and those of Hayes-Roth (1977). This experiment shows that asymmetric fan sentences do not behave differently from fan sentences. Perhaps the actual difference was that Hayes-Roth's material all involved the same asymmetric fan, whereas we had a mixture of different types of asymmetric-fan plus symmetric-fan materials. Her subjects may have implemented a general strategy that initiated search from the unique predicate, whereas such a general strategy would not have worked for our subjects. In any case, our results call into serious question the generality of Hayes-Roth's knowledge-assembly theory and conclusions about interference based on it. There is no evidence in our research for a qualitative change in the nature of the knowledge, although we have used considerably more extensive practice than Hayes-Roth.

Experiment 3

The third experiment was devoted to understanding better the effects of practice per se, rather than its interaction with fan. One of the interesting results of the first experiment was that there was no effect of 12 versus 24 repetitions per day. Similarly, Experiment 2 found no effect associated with 20 versus 30 repetitions. Experiment 3 explored the range from 1 to 8 repetitions to study the limits of this result.

A second goal of this experiment was to tease apart the effects of general practice at the task versus specific practice on the items. These factors were confounded in the first two experiments. It seems more natural to assume that the effects of speed-up were due to improvement in memory for specific items, particularly in light of the interaction with fan. However, it is logically possible that part or all of the speed-up was due to general practice in the task and that the benefit would extend to completely new items.

Previous research by Postman (1982) would, at first glance, lead us to expect that general practice would have *no* effect on recognition memory. Postman (1982, Experiment 1) found no differences in accuracy in a yes–no recognition paradigm among groups of subjects who had or had not previously practiced the task with different materials. However, general practice effects have been found in other recognition memory studies (e.g., Anderson, 1981; Atkinson & Juola, 1973). For example, Atkinson and Juola (1973, see Figure 1) found that recognition latencies for the first test of an item decreased with the number of trial blocks subjects had participated in. Anderson (1981, Experiment 2) had subjects who studied and were tested on two lists of paired associates, one list after the other. Anderson found that List 2 recognition latencies were less than those of List 1, and that List 2 accuracy was greater than that of List 1.

The discrepancy between the results of Postman and those of Anderson and Atkinson and Juola seems to be due to the sensitivity of the dependent measures that were used. The Anderson and Atkinson and Juola studies tested subjects' recognition memory in time-pressured situations, whereas Postman's procedure allowed subjects to take any amount of time to respond. In the current experiment we expected to pick up on any general practice effects by testing subjects' recognition latencies in a time-pressured situation.

In this experiment we sought to manipulate three factors: number of days items were presented, number of days of task practice, and number of item repetitions per day. On each of 5 days, subjects learned new materials in study and test phases. Subjects then entered a recognition reaction time phase in which the materials consisted of just-learned items and materials presented on earlier days. This allowed us to control the item age (number of days presented for recognition) of materials. For example, on the fifth and final day of the experiment, materials having item ages of 0 (new), 1, 2, or 3 days were presented for recognition. This design also permitted us to study the effects of task practice because we could examine recognition times for items having the same item age across different days of task practice. For example, we could look at recognition times to new items on Days 1 through 5 of the experiment and examine how general task practice affected recognition time holding item age constant. Orthogonal to the item age and task practice manipulations, we varied the number of daily repetitions of each item from one to eight.

Method

Subjects. A total of 14 Carnegie-Mellon University undergraduates were paid for participation in this experiment.

Materials. Items for each subject consisted of 128 random combinations of subjects and verb-object predicates, forming sentences identical in structure to those of Experiment 1. Sentence subjects were person nouns (e.g., "punker"). Objects were person nouns or highly concrete nonperson nouns (concreteness ratings of greater than 6.00 on Paivio, Yuille, & Madigan, 1968, norms). Subjects, verbs, and objects varied in length from four to eight letters. Because no interference manipulation was performed in the current experiment, all sentences contained subjects, verbs, and objects that appeared in no other sentences.

Design and procedure. In all five daily sessions, subjects participated in a study, test, and recognition reaction-time phase. On each day, subjects studied and were tested on a new subset of materials. Following study and test, subjects entered a recognition phase in which the just-studied items were probed along with materials presented for recognition on previous days. On each day, one new subset of material was learned in the study and test phases and entered into the pool of recognition items, and an equal sized subset of materials was dropped from the pool. The exception to this procedure was the first day on which all items were new.

An abstract representation of the materials presented for recognition each day is presented in Table 2. Each day's recognition material consisted of four subsets of 16 sentences each. Each subset of materials is denoted by a letter in Table 2. Subsets of materials could be presented for 1 to 4 days. Materials A, presented for 1 day, were replaced by Materials E, which would then appear for 4 subsequent days. Materials B, started on Day 1 and presented for 2 days, were replaced by Materials F, which would be presented for 3 days. Similarly, Materials C, started on Day 1 and presented for 3 days, were replaced by Materials G, which would appear for 2 days. Finally, Materials D, which were presented 4 days beginning on Day 1, were replaced by Materials H, which only appeared for 1 day. Thus on Days 4 and 5, the materials had four levels of item age: 0 (new), 1, 2, or 3 days old.

In addition, recognition materials having the same item age were presented across various sessions (see Table 2). For instance, new items were seen on all days, 1-day-old items occurred on Days 2 through 5, 2-day-old items occurred on Days 3 through 5, and 3-day-old items occurred on Days 4 and 5. This allowed us to examine the effect of days of general task practice holding item age constant. Within each item age condition there were four levels of item repetitions per day. Sentences could be presented 1, 2, 4, or 8 times per day. There were four sentences in each of these repetitions/day conditions.

The study, test, and recognition procedures of this experiment were essentially the same as those in Experiment 1, with the following exceptions. On Day 1, subjects studied 64 sentences, and on subsequent days subjects learned only the new sentences (16) for that day. In the test phase, subjects were queried with a question of the form "WHO predicate?" and were to respond with the matching target subject. Subjects iterated through the query pool until queries had been answered correctly twice for each target. Subjects were tested only on sentences presented in the study phase. Foils were created by randomly re-pairing target subjects and predicates within each item age condition. The fact recognition procedure consisted of 480 trials per day, divided into 20 equal blocks.

Results and Discussion

Subjects' mean reaction times for correct responses to targets and foils in each presentation condition on each day (four repetitions per day by four item age conditions) constituted the raw data. The mean error rate was 6.2%. These reaction time data, averaged over correct responses to targets and foils, are

Table 2
Design Outline for Experiment 3

Day 1	Day 2	Day 3	Day 4	Day 5
A (0)	E (0)	E (1)	E (2)	E (3)
B (0)	B (1)	F (0)	F (1)	F (2)
C (0)	C (1)	C (2)	G (0)	G (1)
D (0)	D (1)	D (2)	D (3)	H (0)

Note. Letters denote subsets of materials presented for recognition. Parenthesized values indicate the number of days of prior practice on materials.

Table 3
Recognition Times (in Milliseconds) for Item Age Conditions Across Practice Days in Experiment 3

	Repetitions per day				
Item age	1	2	4	8	*M*
			Day 1		
0	1222	1177	1091	1057	1137
			Day 2		
0	1288	1142	1110	1036	1144
1	1243	1146	1022	961	1093
			Day 3		
0	1139	1078	947	926	1022
1	1183	1055	1003	906	1037
2	1094	967	906	850	954
			Day 4		
0	1139	1063	955	890	1012
1	1314	1025	925	877	1035
2	1045	971	894	836	935
3	1054	926	845	817	910
			Day 5		
0	1200	1010	912	886	1002
1	1148	1049	893	845	981
2	1095	913	847	815	918
3	999	891	847	798	884

presented in Table 3. The first question of interest concerned the effect of task practice. It can be seen in Table 3 that reaction times decreased substantially as a function of the number of days subjects performed the recognition task, holding item age constant. For example, mean reaction times to new items decreased from 1,137 ms on Day 1 to 1,002 ms on Day 5. This suggests that task practice produces faster retrieval times over and above item-specific practice.

To verify this task practice effect, a Days (2 to 5) × Item Age (new, 1 day old) × Repetitions per Day (1, 2, 4, 8) × Item (target, foil) within-subjects analysis of variance was performed.[2] This analysis yielded a significant effect of days of practice, $F(3, 39) = 6.90$, $MS_e = 90$, $p < .01$; of repetitions per day, $F(3, 39) = 42.53$, $MS_e = 90$, $p < .01$; but not of item, $F(1, 13) = .27$, $MS_e = 50$, $p > .5$. Item age was not significant in this particular analysis, $F(1, 13) < 1$, $MS_e = 40$. This probably reflects the fact that new items have the advantage of being more recently practiced (i.e., they are seen in the study phase prior to the reaction time phase) than 1-day-old items on all days, whereas 1-day-old items have greater frequency of prior practice.

We were also interested in the effects of item age and number of repetitions per day on recognition times. Although the above analysis showed a significant effect of repetitions/day, that analysis confounds the effects of item repetitions received in prior sessions with the within-session effects of repetition. To illustrate, consider that target reaction times for the eight repetition conditions (collapsed across Days 4 and 5 and item age) decrease from 987 ms on first presentation

[2] We should point out that foils were not explicitly repeated. Randomly generated foils from a condition were arbitrarily assigned identifiers corresponding to target identifiers from that condition. It is these identifiers that are tabulated when counting repetitions.

to 821 ms on the eighth presentation within a session. The above analysis, by averaging across all within-session repetitions, might have artificially shown faster reaction times for the 8-repetitions/day conditions than for the 1-repetition/day condition. We thus sought to avoid this confound by examining data for the first recognition presentation of items within each session.

Figure 3 plots reaction times for the first target presentation for all item age and repetitions per day conditions collapsed across Days 4 and 5. These data were submitted to a Day (4, 5) × Item Age (0, 1, 2, 3) × Repetitions per Day (1, 2, 4, 8) × Item (target, foil) within-subjects analysis of variance. In this particular analysis there was no significant effect of days of practice, $F(1, 13) = 1.85$, $p > .05$, $MS_e = 104$, which is consistent with the power-law prediction that day-to-day speed-up increments decrease with days. The item main effect approached significance, $F(1, 13) = 4.06$, $p = .07$, $MS_e = 32$. One can see in Figure 3 that the number of days an item was seen produced decreases in reaction times, $F(3, 39) = 18.92$, $p < .001$, $MS_e = 69$, as did the number of times it was seen per day, $F(3, 39) = 17.08$, $p < .001$, $MS_e = 96$.

It is interesting that there was no reliable Item Age × Repetitions per Day interaction, $F(9, 117) = .86$, $p > .5$, $MS_e = 32$. One curious case to consider is the decrease in reaction times in the new item condition as a function of repetition condition. Because the data plotted are for the first presentation of an item, one might expect this curve to be relatively horizontal. That is, given that new items had the same amount of prior practice at the first presentation, one might expect reaction times to be constant for all repetition conditions. However, one must recall that items were presented in a random order during an experimental session. Therefore, on average, the first presentation of one-repetition-per-day items occurred later in an experimental session than the first presentation of eight-repetition-per-day items. Thus, the decrease in reaction times for new items as a function of repetition conditions probably indicates that new items produced slower reaction times if they were presented for the first time later in an experimental session.

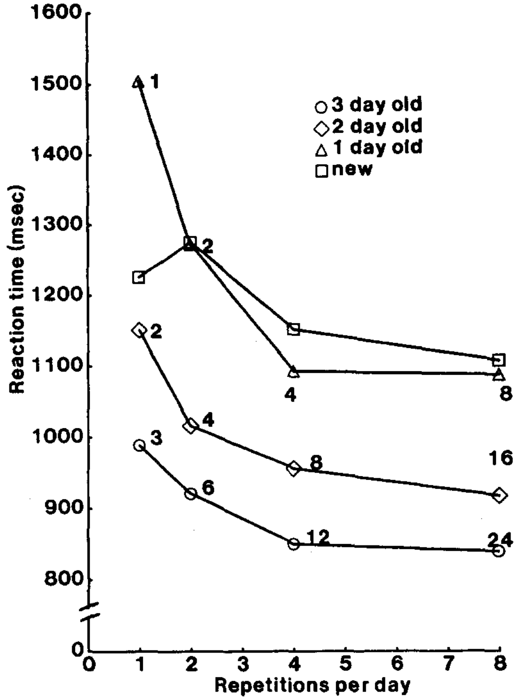

Figure 3. Item age and repetitions/day effects on first target presentation reaction times in Experiment 3. Numbered labels on data points indicate the number of prior repetitions of an item on previous days (no label indicates that the item is new).

We did another analysis to determine if there was an effect of number of repetitions per day independent of serial position effects: We collapsed data for first item presentation across Days 4 and 5. We also collapsed the one- and two-repetitions-per-day conditions together (low repetitions) and collapsed the four- and eight-repetitions-per-day conditions (high repetitions). We selected a subset of the data such that the first presentations of high- and low-repetition items were equated for serial position. This was achieved by the following procedure: We selected the first presentation of a low-repetition item. It was included in the analysis only if we could find a first presentation of a high-repetition item within 10 trials. If so, we yoked these low and high items together. Selecting items in this manner, the low-repetition items had a mean serial position of 77.22 (out of 480 trials/day), whereas the high-repetition items analyzed had a mean serial position of 73.86. The mean difference of low- minus high-

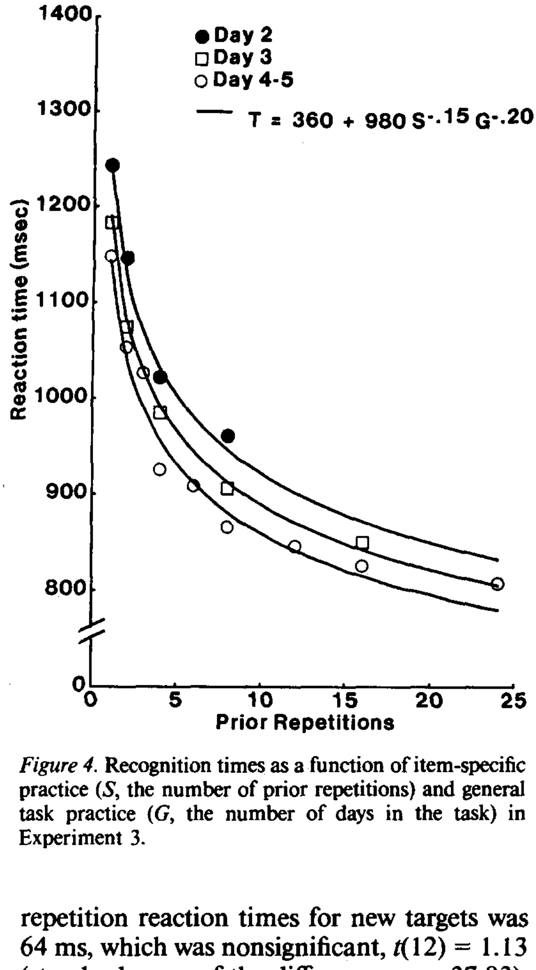

Figure 4. Recognition times as a function of item-specific practice (S, the number of prior repetitions) and general task practice (G, the number of days in the task) in Experiment 3.

repetition reaction times for new targets was 64 ms, which was nonsignificant, $t(12) = 1.13$ (standard error of the differences was 37.83). Collapsing across the 1-, 2- and 3-day-old targets, the mean low- minus high-repetition reaction time difference (116 ms) was significant, $t(13) = 3.07$, $p < .01$ (standard error of the differences was 25.05). Thus, there appears to be some benefit of more than two repetitions per day.

Although there is some effect of number of repetitions per day, the stronger variable appears to be number of days of repetition. This can be seen quite clearly in Figure 3, which gives recognition times on first presentation of a target on a day. The 1-, 2- and 3-day old curves are labeled with total number of previous repetitions over all days (unlabeled points are new items). When points are equated for total number of repetitions, subjects are faster when those points are distributed over more days with fewer presentations per day. For example, items that were repeated eight times over the course of 2 days produced shorter response latencies than items repeated eight times on 1 prior day (contrast the points labeled "8" on the 1-day-old and 2-day-old curves in Figure 3). Note that this occurs despite the serial position effects that tend to favor times in conditions with more repetitions per day.

Figure 4 re-presents the data from Table 3 as a function of day of practice (G) and total number of specific tests (S) of a fact in prior days. According to the theory sketched out in Experiment 1, these data should be fit by the power function in Equation 2. We used a STEPIT program to estimate best fits to this data, and the best-fitting functions are reported in Figure 4. The intercept I was fixed at .36. This value of I was estimated in Experiment 1, which we assumed provided a better estimate because it looked at more extensive practice (values of I and the other parameters are hard to separate in estimation). The best-fitting parameter values are $B = .98$, $e = .15$, and $f = .20$. The exponent e estimates rate of item-specific speed-up, and the exponent f estimates rate of general task speed-up. It is interesting that the the rate of general speed-up appears to be slightly greater than the item-specific speed-up. It is also interesting that the combined value of $e + f$, .35, is very close to the .36 exponent estimated for Experiment 1.

There are two noteworthy results in this experiment. Consistent with the first two experiments, there seems to be little effect of amount of practice per day relative to the effect of number of days of practice. There was no evidence of an effect of four versus eight trials and a minimal effect of one and two versus more trials. The second result was the evidence that general speed-up is at least as important as practice of specific items. This might not seem surprising if we can ascribe such general factors to peripheral processes such as stimulus encoding and response execution. From this view the more interesting result is that we have evidence for an effect of specific practice over and above general practice. However, there is only one difficulty in the peripheral interpretation of the general practice effect. The first two experiments indicated that the practice effect (not separating general and specific components) interacted with the fan effect. The fan effect is surely a central process and would not be expected to interact with peripheral practice. There is the possibility that the general processes speeding up may include central as well as peripheral process. These general, central processes might include comparison of a probe to memory traces, and so forth.

Note that Equation 2 for the ACT* theory allows for general processes involving strengthening of the fact recognition productions. Close inspection of Equation 2 reveals an interesting prediction. This equation predicts a multiplicative relation between general practice (G^{-f}) and fan B). In other words, we should expect the fan effect to decrease in proportion to the amount of general practice on a task. This is a surprising possibility on many theoretical accounts, including the 1976 ACT theory (Anderson, 1976). This is because the fan effect is frequently thought of as affecting only retrieval of the fact and not such general processes as comparing the probe to memory.

Experiment 4

The fourth experiment explored the intriguing possibility of an interaction of general practice and the fan effect. Typically, general practice effects have been ascribed to peripheral or noncognitive processes (e.g., Atkinson & Juola, 1974). Such proposals would predict additive effects of general practice and fan rather than the multiplicative one predicted by ACT* (Anderson, 1982, pp. 398–400). However, results from Postman (1969, pp. 285–293) indicate that interference does decrease with general practice. That study involved a test phase contrasting recall of a list of A-B paired associates among groups learning either lists of A-B and A-C paired associates (interference conditions) or only A-B paired associates (control conditions). Subjects in the interference conditions showed poorer recall than subjects in the control conditions, but these differences were attenuated for subjects who had prior practice with the paired-associates recall task with different materials.

In the current experiment, we varied factors of general practice, item-specific practice, and fan. We did so by presenting fan and no-fan materials on either 4 or 8 days over the course of 8 days of a fact retrieval task. On the 9th and 10th day of the task, subjects were presented with new fan and no-fan material. We could thus hold item age constant and examine the fan effect at different levels of task practice (e.g., contrasting reaction times to new materials on Days 1 and 9). By presenting materials for 4 versus 8 days over eight daily sessions we were also able to look at the interactions of fan with item-specific practice.

Method

Subjects. Twelve Carnegie-Mellon undergraduates completed the study for a combination of pay ($3/hr plus performance bonus) and partial course credit.

Materials. Sentences identical in form to those used in Experiments 1 and 3 were randomly constructed from 36 subjects, verbs, and objects for each experimental subject. A total of 48 target sentences was constructed. Half of these targets were no-fan sentences; the others were fan sentences. Foils for a condition were constructed by a random re-pairing of subjects, verbs, and objects from the target condition into a sentence frame.

Design and procedure. The current experiment took place over a 10-day period with a weekend break between Days 5 and 6. An abstract representation of the current design is presented in Table 4. We manipulated specific practice on items by presenting materials for either 4 or 8 days over the course of the first eight daily sessions. To examine effects of general practice, we presented subjects with new materials on Days 9 and 10. Thus, specific practice with materials was equated for Days 1 and 9 as well as for Days 2 and 10. However, subjects on Days 1 and 2 had less general practice with the task than they did on Days 9 and 10.

Materials were divided in half for each subject, with half the materials presented on Days 1 to 8 and the other half on Days 9 and 10. In the initial eight sessions, fan and no-fan materials were divided equally into materials that were presented on either 4 (low-practice materials) or 8 (high-practice materials) days. High-practice fan and no-fan materials (HIF and HINF, respectively, in Table 4) appeared every day over the initial eight sessions. Low-practice fan and no-fan materials (LOF and LONF, respectively, in Table 4) were presented on Days 1, 4, 7, and 8. Since high- and low-practice materials were presented on Day 7, both were equated on Day 8 for the number of intervening days since last practice. On Days 9 and 10, new fan (NEWF) and new no-fan (NEWNF) sentences were introduced, and all previous materials were dropped from the experiment. These new sentences allowed us to determine if there had been any impact of general practice on the fan effect.

Study, test, and recognition reaction time phases were conducted as in Experiment 3. The study and test phases of the experiment occurred on Days 1 and 9 (when new materials were introduced). All target sentences were repeated 10 times in the recognition phase, and each target presentation was matched with a foil presentation

Table 4
Design Outline for Experiment 4: Presentation Schedule for Materials

Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10
HIF	HIF	HIF	HIF	HIF	HIF	HIF	HIF	NEWF	NEWF
HINF	HINF	HINF	HINF	HINF	HINF	HINF	HINF	NEWNF	NEWNF
LOF			LOF			LOF	LOF		
LONF			LONF			LONF	LONF		

Note. HIF = high-practice fan; HINF = high-practice no fan; LOF = low-practice fan; LONF = low-practice no fan; NEWF = new fan; NEWNF = new no fan.

from the appropriate condition. Thus, on Days 2, 3, 5, and 6 there were 480 reaction-time trials, and on all other days there were 960 trials. Trials in the recognition reaction time phase were divided into blocks of 24.

Results and Discussion

As in the previous experiments, mean reaction times for correct responses to targets and foils in each condition on each day were computed for each subject. The mean error rate was 10.6% in this study. These mean reaction time data, collapsed across targets and foils, for the 10 days of the experiment are presented in Figure 5.

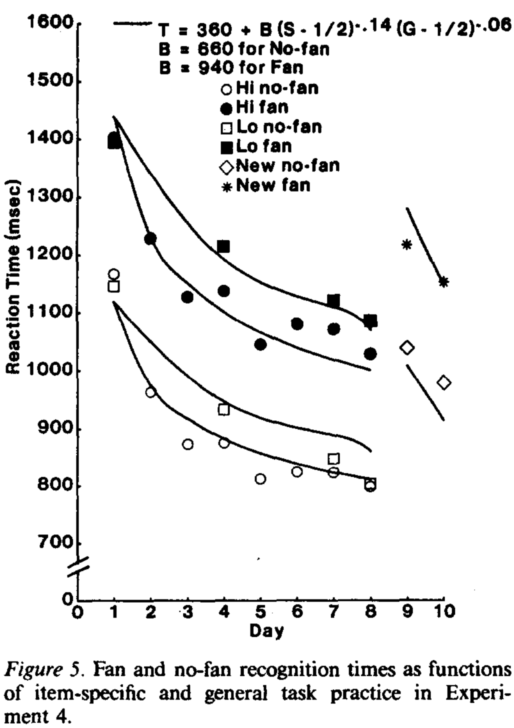

Figure 5. Fan and no-fan recognition times as functions of item-specific and general task practice in Experiment 4.

To examine general task practice effects, we looked at the Days 1 and 9 data. Recall that item-specific practice is the same (i.e., all materials are new), whereas general task practice increases from Day 1 to Day 9. On Day 1, the mean reaction time to fan items was 1,406 ms and to no-fan items it was 1,150 ms. On Day 9, the mean reaction time to fan items was 1,232 ms and to no-fan items it was 1,048 ms. A within-subjects analysis of variance of days (1, 9), fan (fan, no-fan), and item (target, foil) showed that the fan versus no-fan reaction time difference was significant $F(1, 11) = 148.25$, $MS_e = 8$, $p < .01$, as was the main effect of days of practice, $F(1, 11) = 8.76$, $MS_e = 50$, $p = .01$. Target responses were significantly faster than foil reaction times, $F(1, 11) = 43.21$, $MS_e = 10$, $p < .01$. Of special interest was the significant interaction between the effects of fan and days practice, $F(1, 11) = 10.10$, $MS_e = 3$, $p < .01$. This suggests that general practice in a memory task, as opposed to practice with specific items, can reduce the effects of long-term memory interference. This result also supports the ACT* prediction of a multiplicative relation between fan and general practice. However, looking beyond the predictions of the specific theory, this finding suggests a view of memory in which we should consider the interaction of structural attributes (e.g., strength) of stored traces and the general processes that operate on them.

To examine the relation between fan and item-specific practice, we performed a within-subjects analysis of variance of days (1, 4, 7, 8), fan (fan, no-fan), Practice (high, low), and item (target, foil). This analysis also showed a significant fan effect, $F(1, 11) = 155.05$, $MS_e = 40$, $p < .01$, an effect of days of practice, $F(3, 33) = 42.73$, $MS_e = 60$, $p < .01$, and that target responses were faster than foil responses $F(1, 11) = 124.28$, $MS_e = 10$, $p < .01$. We also found that reaction times in the high-practice conditions (M = 1,068 ms) were faster than the low-practice reaction times (M = 1,033 ms), $F(1, 11) = 13.94$, $MS_e = 8$, $p < .01$. High- and low-practice reaction times were basically equal on Day 1 (M = 1,276 ms and M = 1,280 ms, respectively) but differed slightly by Day 8 (M = 910 ms for high practice; M = 938 ms for low practice). This interaction of practice conditions with days was significant, although not very large in absolute terms, $F(3, 33) = 3.89$, $MS_e = 4$, $p < .05$. We expected a decrease of the fan effect with item-specific practice, and we were surprised to find that the interaction between fan and practice was not significant, $F(1, 11) = 1.13$, $MS_e = 20$. A specific t test contrasting the fan effect for the low-practice conditions on Days 7 and 8 (280 ms) with the high-practice conditions on Days 7 and 8 (233 ms) was also nonsignificant, $t(11) = .10$. The failure to get an interaction between fan and amount of specific practice may be owing to the rather weak effect of amount of specific practice in this experiment.

Figure 5 shows the best-fitting power functions (estimated using STEPIT) to the data. We fit this assuming that item-specific practice increased over the initial 8 days of the experiment and was then reduced to 0 on Day 9. Note that these functions underpredict reaction time for the fan conditions on Day 1 and overpredict reaction time for the non-fan condition. This corresponds to our failure to get a significant reduction in the fan effect with item-specific practice. These power-function estimates also indicate that the effect of general practice in Experiment 4 (f, the exponent reflecting general speed-up, is .06) was substantially less than that found in Experiment 3 (f = .20). Despite this weaker general practice effect, we did find a reliable decrease in the fan effect with general practice.

General Discussion

We make the following conclusions from the experiments:

1. Practice produces power-function improvements of speed of fact retrieval.
2. The fan effect diminishes in direct proportion to the overall speed-up in retrieval but does not disappear.
3. There is a relatively small effect of amount of practice within a day relative to amount of practice across days. There is no evidence for a benefit of more than four trials per day.
4. About 30% (Experiment 4) to 50% (Experiment 3) of the overall improvement seems to be owing to general practice factors, and these general practice factors reflect central processes such as comparison of the probe to memory.
5. General practice appears to reduce the fan effect.

As derived in Anderson (1982), the ACT* power-law speed-up prediction is principally based on the assumption that each strengthening of a stored production rule or fact decays as a power function of time. This assumption is empirically based on work from our lab as well as that of Wikelgren (1976). There is also some suggestive evidence in the neuroscience literature (McNaughton, undated) indicating that neural enhancement decays in a power-law manner.

An alternative practice model predicting speed-up of power law form is the *chunking model* of Newell and Rosenbloom (1981). This model assumes that improvement owing to practice results from the acquisition of knowledge regarding stimulus and response patterns in a task environment. The chunking model assumes that patterns are learned at a constant rate, with lower order patterns being combined into higher order ones. However, it also assumes that large patterns (e.g., a particular state of a chess board during play) reoccur less frequently than small patterns (e.g., a rook on a particular square). Thus, rate of improvement slows down as the subject has to learn more complex patterns. This slow-down in rate of improvement produces the power function. This model is intuitively appealing for tasks that have a combinatorial structure, such as chess (Chase & Simon, 1973) or responding to 10 lights each of which may be on or off (Seibel, 1963). However, the fact retrieval experiments reported in this article do not have an obvious combinatorial structure[3] because the stimulus and response patterns were constant each day. Thus, these results stand as a modest challenge to that theory.

We were somewhat perplexed at the discrepancy between our current results and those of Hayes-Roth (1977). We found that interference decreases in proportion to practice but does not disappear. Although it is possible that the Hayes-Roth conclusion was an instance of falsely failing to reject the null hypothesis, we should consider other explanations for the differences. One possibility, pointed out earlier, is that the Hayes-Roth subjects may have hit upon a strategy choice that diminished the impact of the fan effect to levels that could not be picked up in her experimental design.

Another possible explanation is suggested by the research on automaticity (Schneider & Shiffrin, 1977; Shiffrin & Dumais, 1981; Shiffrin & Schneider, 1977). Generally, those studies have shown that reaction times in visual scanning and short-term memory search become independent of target set size. If one considers the fan effect to be the long-term memory analog to the short-term memory set size effect (or vice versa), then one would expect to see the practice results obtained by Hayes-Roth (1977). However, there is evidence in the visual scanning literature which suggests that the set size effect does not always disappear with practice (Schneider & Eberts, 1980; Schneider & Fisk, 1980). The reduction appears to be inhibited by inconsistent or varied mappings of stimulus features onto responses. For example, Schneider and Fiske impeded the effects of practice on the set size effect by using distractors containing either the shape or color of targets, which were defined as a conjunction of shape and color. Perhaps in comparison to our studies, the Hayes-Roth material was less varied (recall that there was only one type of asymmetric fan) or the foils were more discriminable from targets.

In Experiments 1, 2, and 3, we found that increasing the degree of practice above a few repetitions within a day produced no detectable impact on reaction time. The ACT* theory does not predict such an effect. It seems that the increments of strength for a fact or production somehow become depleted with repeated use over a short span of time, and some period of time is required to recoup that depletion. Such an attenuation of strengthening effects over repetitions appears to violate the *total time law* (Bugelski, 1962; Cooper & Pantle, 1967), which states that the memorability of an item is a direct function of the total time spent studying the item. Our results are, however, generally consistent with the habituation model (Hintzman, 1974; Posner & Warren, 1972) of spacing effects, which states that the beneficial effects of repetitions on memory deteriorate (owing to habituation) with decreases in the time between repetitions. Currently, we are simply attributing this effect to features of the underlying neural hardware, and indeed there seems to be some evidence from the neurosciences indicating that such a fatigue effect does occur with synaptic transmitters (Eccles, 1972).

A particularly interesting and important notion suggested by the current results is that central processes may speed up with task practice. Initially one might be tempted to attribute speed-up due to general task practice to peripheral factors involving in response execution. However, it seems unlikely that such an account could address the rather substantial decrease in the fan effect found in Experiment 4.

[3] But one might assume that contextual patterns may have a combinatorial nature.

References

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

Anderson, J. R. (1981). Interference: The relationship between response latency and response accuracy. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 326–343.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89,* 369–406.

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

Anderson, J. R. (1983b). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22,* 261–295.

Atkinson, R. C., & Juola, J. F. (1973). Factors influencing speed and accuracy of word recognition. In S. Kornblum (Ed.), *Attention and performance IV.* New York: Academic Press.

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, S. D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology: Learning, memory and thinking* (pp. 243–293). San Francisco: Freeman.

Bray, C. W. (1948). *Psychology and military proficiency.* Princeton, NJ: Princeton University Press.

Bugelski, B. R. (1962). Presentation time, total time, and mediation in paired associate learning. *Journal of Experimental Psychology, 63,* 409–412.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction.* Hillsdale, NJ: Erlbaum.

Chandler, P. J. (1965). *Subroutine STEPIT: An algorithm that finds the values of the parameters which minimize a given continuous function* [Computer program]. Bloomington: Indiana University, Quantum Chemistry Program Exchange.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55–81.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82,* 402–428.

Cooper, E. H., & Pantle, A. J. (1967). The total-time hypothesis in verbal learning. *Psychological Bulletin, 68,* 221–234.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Eccles, J. C. (1972). Possible synaptic mechanisms subserving learning. In A. G. Karyman & J. C. Eccles (Eds.), *Brain and human behavior.* New York: Springer-Verlag.

Hayes-Roth, B. (1977). Evolution of cognition structures and processes. *Psychological Review, 84,* 260–278.

Hintzman, D. L. (1969). Recognition time: Effects of recency, frequency and the spacing of repetitions. *Journal of Experimental Psychology, 79,* 192–194.

Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium* (pp. 77–99). Potomac, MD: Erlbaum.

Hintzman, D. L., & Rogers, M. K. (1973). Spacing effect in picture memory. *Memory & Cognition, 1,* 430–434.

Lewis, C. H. (1979). *Speed and practice.* Unpublished manuscript, IBM Research Center, Yorktown Heights, NY.

McNaughton, B. L. (undated). *Activity dependent modulation of hippocampal synaptic efficacy: Some implications for memory processes.* Unpublished manuscript, Cerebral Functions Group, Department of Anatomy and Embryology, University College, London, England.

Mohs, R. C., Wescourt, K. T., & Atkinson, R. C. (1975). Search processes for associative structures in long-term memory. *Journal of Experimental Psychology: General, 104,* 103–121.

Neves, D. M., & Anderson, J. R. (1981). Knowledge compilation: Mechanisms for the automization of cognitive skill. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 57–84). Hillsdale, NJ: Erlbaum.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness. *Journal of Experimental Psychology Monographs, 76*(1, Pt. 2).

Posner, M. I., & Warren, R. E. (1972). Traces, concepts and conscious constructions. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory.* Washington: Winston.

Postman, L. (1969). Experimental analysis of learning to learn. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 3). New York: Academic Press.

Postman, L. (1982). An examination of practice effects in recognition. *Memory & Cognition, 10,* 333–340.

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

Reed, A. V. (1976). List length and the time course of recognition in immediate memory. *Memory & Cognition, 4,* 16–30.

Schneider, W., & Eberts, R. (1980). *Automatic processing and the unitization of two features* (Tech. Rep. 8008). Urbana-Champaign: Human Attention Research Laboratory, University of Illinois.

Schneider, W., & Fisk, A. D. (1980). *Degree of consistent training and the development of automatic processing* (Tech. Rep. 8005). Urbana-Champaign: Human Attention Research Laboratory, University of Illinois.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review, 84,* 1–66.

Seibel, R. (1963). Discrimination reaction time for a 1023 alternative task. *Journal of Experimental Psychology, 66,* 215–226.

Shiffrin, R. C., & Dumais, S. T. (1981). The development of automatism. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 111–140). Hillsdale, NJ: Erlbaum.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Shoben, E. J., Wescourt, R. T., & Smith, E. E. (1978). Sentence verification, sentence recognition, and the semantic-episodic distinction. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 304–317.

Smith, E. E. (1981). Organization of factual memory. In J. H. Flowers (Ed.), *Cognitive processes* (pp. 163–209). Lincoln: University of Nebraska Press.

Wikelgren, W. A. (1976). Network strength theory of storage and retrieval dynamics. *Psychological Review, 83,* 466–478.

Received January 9, 1984
Revision received June 25, 1984 ■