Journal of Experimental Psychology:
Learning, Memory, and Cognition
1991, Vol. 17, No. 1, 81-93

0278-7393/91/$3.00

Forgetting and the Mirror Effect in Recognition Memory: Concentering of Underlying Distributions

Murray Glanzer
New York University

John K. Adams
Montclair State College

Geoffrey Iverson
University of California at Irvine

The mirror effect is a strong regularity in recognition memory: If there are two conditions, A and B, with A giving higher recognition accuracy, then old items in A are recognized as old better than old items in B, *and also* new items in A are recognized as new better than new items in B. The mirror effect is explained by attention/likelihood theory, which also makes several new, counterintuitive predictions. One is that any variable, such as forgetting, that affects recognition changes the responses to new as well as old stimuli. In terms of underlying distributions, forgetting produces concentering, the bilateral movement of distributions, both new (noise) and old (signal), toward a midpoint. Data from two forced-choice experiments are reported that support the prediction of concentering and other predictions drawn from the theory. It is argued that current theories of memory, which are strength theories, cannot handle these regularities.

In this article, we examine implications of a theory of the mirror effect in recognition memory. In the course of this examination, we demonstrate support for a paradox: When subjects forget, they forget new items as well as old items. This means that when forgetting occurs, subjects' performance on new items declines, not just performance on old items.

The mirror effect, which is our starting point, refers to the following strong regularity in recognition memory. If there are two conditions, A and B, that result in differences in recognition accuracy, then the underlying old (signal) and new (noise) distributions can be shown to be arranged as in Panel 1 of Figure 1. The difference in accuracy is reflected not only in the differential placement of the old distributions, AO and BO, but also in the differential placement of the new distributions, AN and BN. When AO is higher than BO, AN is lower than BN. The constraint on the positions of the new distributions is surprising and requires explanation.

A meta-analysis of 80 experiments has shown that the mirror effect is a general characteristic of recognition memory (Glanzer & Adams, 1985). In those experiments, the effect was produced by using different sets of stimulus items (e.g., high-frequency vs. low-frequency words), that is, by stimulus variables. Further work has shown that the effect holds when differences in recognition accuracy are produced by transformations of words (normal vs. reverse presentation; Glanzer & Adams, 1990) and by encoding variables (classification tasks that induce positive vs. negative responses; Adams & Glanzer, 1989). These results generalize the effect beyond stimulus variables.

The mirror effect requires that recognition memory be viewed as more complex than it has been. Up to now the key element in a subject's memory performance was assumed to be some measure of the strength of the trace of the item. The regularities of the mirror effect call for the subject to make an additional evaluation of that information. This point was originally made by Brown (1976). Glanzer and Adams (1990) presented and tested a theory—attention/likelihood theory—that includes an evaluation beyond simple strength measure to account for the effect. This article extends that theory to cover forgetting effects and tests implications of that extension. The implications are novel and contradict expectations on the basis of existing theories.

Existing theories relevant to the data fall into two groups. One is an older group: theories of forgetting (Bower, 1972; Estes, 1955; and the theories summarized by Hintzman, 1978). The other is a more recent group of theories of recognition memory that have been concerned with the mirror effect (Brown, 1976; Gillund & Shiffrin, 1984; Hintzman, 1988; Hockley & Murdock, 1987). Both groups consist of strength theories with the exception of Brown's (1976). They view memory decisions as based directly on some measure of the strength of the trace of an old item. Because they are strength theories, they cannot handle the mirror effect unless they are radically restructured. The forgetting theories, moreover, because they antedate the demonstration of the effect, have not addressed the effect. The recent recognition theories that have addressed the mirror effect have not been extended to cover the interaction of forgetting with the mirror effect. Both groups of theories are covered further in subsequent discussion sections.

The theory presented next differs from strength theories in that the key information on which subjects base their memory

This research was supported by National Institute of Mental Health Grant MH 44938, and Air Force Office of Scientific Research Contract F49620-8-K-0004.

We thank Louis Varricchio, Amy Wolff, and Leslie Sherman for assistance in collection and analysis of data. We also thank Bennet B. Murdock, Jr., Arthur M. Glenberg, and William E. Hockley for their constructive comments.

Correspondence concerning this article should be addressed to Murray Glanzer, Department of Psychology, New York University, New York, New York 10003.

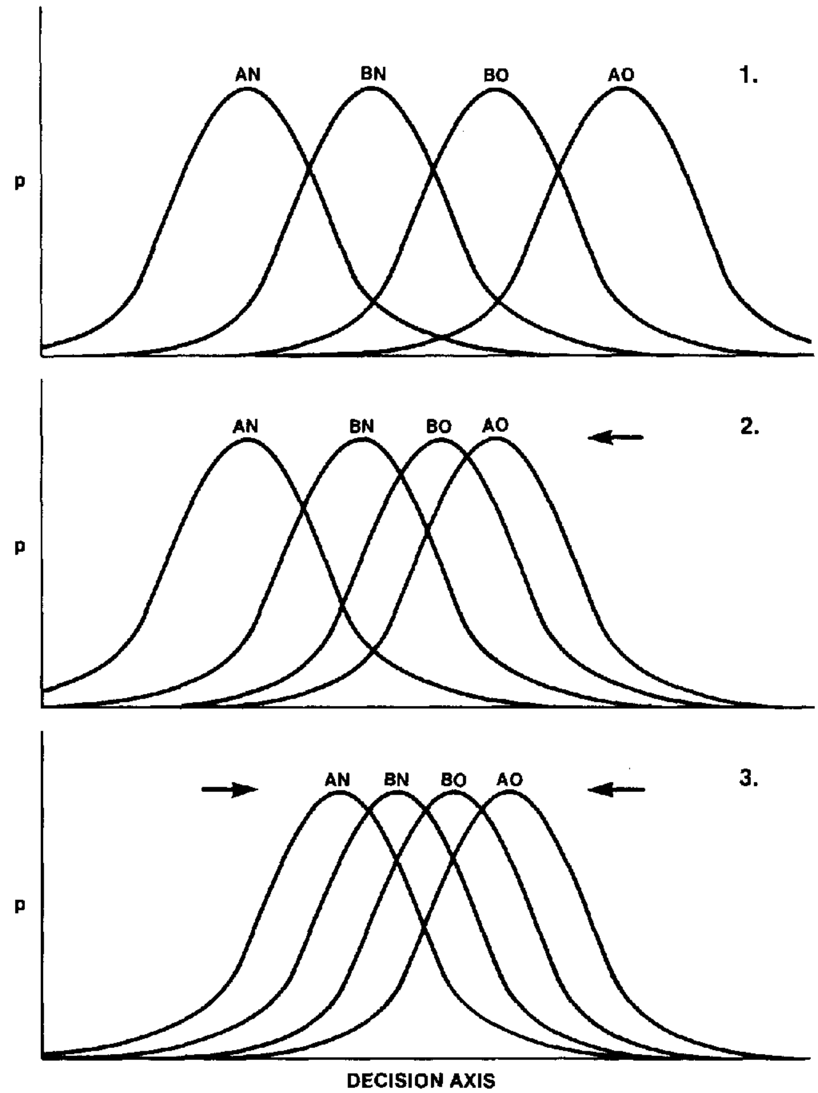

Figure 1. Arrangement of underlying old (signal) and new (noise) distributions immediately after study and after delay (two views). p = probability; *AO* = *A* old; *BO* = *B* old; *BN* = *B* new; *AN* = *A* new; with experimental condition A producing more efficient recognition than B. Panel 1: Array of underlying distributions immediately after study. Panel 2: Change in array after forgetting, according to conventional strength theories. Panel 3: Change in array after forgetting according to attention/likelihood theory.

decisions is not the strength of the trace but a more complex evaluation: its likelihood.

Attention/Likelihood Theory

The theory is a feature-sampling theory with two special mechanisms: an attention mechanism and a decision mechanism. It derives from a model first presented by Estes (1955) for conditioning and later developed by Bower (1972) for memory. It departs from that tradition in abandoning strength (amount of marking) as the basic information on which recognition decisions are made. It substitutes likelihood for strength in that role (see Assertion 5 below).

The assertions of the theory are the following:

1. Stimuli are sets of features. The number of such features, *N*, is assumed to be constant for all stimuli. There is no reason to assume that one stimulus has more or fewer features than another.

2. Some proportion of features is marked[1] in new stimuli. This proportion is *p*(new), which is the noise level. This is

[1] Each feature is either marked or unmarked. A stimulus could be viewed therefore as a string of binary digits. This statement applies to both new and old stimuli as in Assertion 4.

also considered the same for all conditions. There is no reason, at this point, to assume that one new stimulus enters with greater noise marking than another.

3. Different conditions elicit different amounts of attention by the subject. This is translated into differences in the number of features, $n(i)$, examined (sampled). The sampling of features for a given stimulus is assumed to be random.

4. When features are examined, they are marked. The proportion of features marked is $\alpha(i) = n(i)/N$. Therefore, the state of stimuli that were in some condition, i, during study is given by the following equation,

$$p(i,\text{old}) = p(\text{new}) + \alpha(i)\cdot[1 - p(\text{new})]. \tag{1}$$

This states that the subject samples a proportion, $\alpha(i)$, of the features. Those already marked, p(new), remain marked; those not yet marked, $1 - p$(new), are subject to marking. Conditions that evoke examination of a larger proportion of features result in the marking of a larger proportion of features: the learning constant $\alpha(i)$ is larger, and learning is faster.

5. During a recognition test, the subject uses the standard mechanisms of signal-detection theory in making responses. This means that likelihood ratios are computed and decisions are made on the basis of those likelihood ratios. However, the likelihood ratios are computed for every stimulus item and are the basis of the decision made about every item. In this, the present theory differs from other applications of signal-detection theory to recognition memory and differentiates it from a strength theory. In the usual applications of signal detection to a yes/no test in recognition memory, the subject is viewed as generating a likelihood ratio once to define a point on the decision axis in terms of the original strength measure. Thereafter, the subject reverts to strength measures for each stimulus item and uses the comparison with the likelihood-ratio-defined criterion (in strength measure) to make decisions in a yes/no task. For a forced-choice test, the likelihood ratios are, in the usual applications of signal detection, completely dispensed with (see Bower, 1972).

According to attention/likelihood theory, the generation of likelihood ratios for every stimulus presented and their use for every recognition judgment is what produces the mirror effect. Even though p(new) is assumed the same for two conditions (such as A and B discussed previously here), when distributions of likelihood ratios are plotted, the AN and BN distributions of likelihood ratios separate (as in Panel 1 of Figure 1). The transformation of the distributions of number of marked features to distributions of likelihood ratios produces the configuration seen in Panel 1.

The initial distributions used in the theory are binomials. The parameters are $n(i)$ and p(new) for the distribution of marked features for new items, given condition i. The same $n(i)$ together with $p(i$, old) make up the parameters for the distribution of marked features for old items, given condition i. These distributions are transformed, however, into likelihood-ratio distributions as follows. Each pair of distributions (new and old) for a given experimental condition (e.g., high frequency) is used to generate log-likelihood ratios for each binomial value.

$$\begin{aligned}\ln L(x) &= \ln\left(\frac{p(i,\text{old})^x\cdot q(i,\text{old})^{n(i)-x}}{p(\text{new})^x\cdot q(\text{new})^{n(i)-x}}\right)\\ &= n(i)\cdot\ln\left(\frac{q(i,\text{old})}{q(\text{new})}\right) + x\cdot\ln\left(\frac{p(i,\text{old})\cdot q(\text{new})}{p(\text{new})\cdot q(i,\text{old})}\right).\end{aligned} \tag{2}$$

The values of ln $L(x)$ are seen to be related to those of x itself, which has a binomial distribution depending on experimental condition i and state j. Explicitly, the distribution of ln $L(x)$ is as follows:

$$P(\ln L(x) = \lambda) = \binom{n(i)}{x}\cdot p(i,j)^x\cdot q(i,j)^{n(i)-x}, \tag{3}$$

where x is the number of marked features, i is the experimental condition (e.g., high frequency or low frequency), $n(i)$ is the number of features sampled in experimental condition i, j is the state (old or new), and λ is the right-hand side of Equation 2. There will be one distribution for the old items and another distribution for the new items. This does not mean that the subject knows beforehand whether an item is new or old, just that new and old items present a different distribution of likelihood ratios to the subject.

After noting that the current test item has x marked features observed out of $n(i)$, the subject estimates the probability of obtaining that value, x, given that the item is old (and in condition i) and given that it is new. The ratio of these probabilities is the quantity inside the logarithmic term in Equation 2. The greater the ratio, the greater the subject's belief that the item is old.

Equation 2 reflects the subject's model of the situation. It is the basis for the subject's evaluation of each item. Equation 3 reflects the distribution of marking of items presented to the subject by the experimental situation, that is, the actual test situation. As written here, the subject's model as reflected in the probabilities in Equation 2 corresponds with the actual situation (Equation 3). This correspondence is not necessary. There can be cases in which the subject's model differs from the actual situation. In such cases, different ps and qs would be used in Equations 2 and 3. As noted in Glanzer and Adams (1990), we can have those differences and still produce the mirror effect within the general structure of the theory.

The process the subject carries out, according to the theory, is the following. The subject views a test item, and notes both the total number of features, $n(i)$, and x, the number of those features that are marked. The subject then estimates two probabilities: the probability that an old item of that type will have that proportion of marked features and the probability that a new item will have that proportion of marked features. Combining these estimates, a likelihood ratio is computed for the item. The likelihood ratios computed for all the items produce a pair of likelihood-ratio distributions, one distribution for old test items and one for new test items, which are binomially distributed. These distributions of likelihood ratios determine the proportions of choices. In the case of two-alternative forced choice, the subject views two stimuli, generates a likelihood ratio for each stimulus, and then chooses the one with the larger value. As noted previously here, Equation 2 reflects the subject's model of the situation. It plays the same role as Brown's (1976) memorability evaluation.

The major role assigned to the likelihood ratios distinguishes this theory from strength theories. In those theories, the subject decides on the basis of strength or its equivalent: amount of marking, or familiarity. The decision axis for those theories is a strength axis. The decision axis here is a likelihood axis.

Attention/likelihood theory generates the main findings associated with the mirror effect for all three standard recognition paradigms: yes/no, confidence ratings, and two-alternative forced choice. Computations demonstrating this are reported in Glanzer and Adams (1990). Implications of the theory are also tested in that study.

The theory, when extended to cover forgetting, implies a number of additional regularities. The two-alternative forced-choice recognition test is of particular interest with respect to those regularities. As is shown shortly, forced choice permits us to evaluate fully the changes that take place in forgetting. We will, therefore, extend the theory to cover the effects of forgetting associated with a delay between study and test. This is done in the next assertion.

6. Forgetting that occurs with a delay between study and test results from a loss of marking. The loss consists of a proportion β of recently marked elements, $0 < \beta < 1$, in each of t successive units of time. This loss occurs repeatedly over successive units of time. At the end of the study trial, time 0, the proportion of marked elements is $p(i,\text{old},0)$. If forgetting results from repetitive losses over t delay intervals, in each of which the recently marked elements are reduced by β, then after t units of time

$$p(i,\text{old},t) = p(\text{new}) + [1-\beta]^t \cdot \alpha(i) \cdot [1 - p(\text{new})]. \quad (4)$$

A strong prediction about forgetting comes out of attention/likelihood theory. The ordinary expectation would be that forgetting consists of the unilateral movement of the old distributions toward the new distributions. The new distributions, in this view, remain fixed. This is depicted in Panel 2 of Figure 1.

Attention/likelihood theory predicts something quite different. It predicts that with forgetting there is concentering: Both the new and the old distributions move toward a midpoint in the set of distributions. Moreover, the theory predicts that the two old distributions move closer to each other as they move down the decision axis, and the two new distributions move closer to each other as they move up. This collapsing of all distributions on each other and on the midpoint is depicted in Panel 3 of Figure 1. The Appendix presents a proof of the concentering that occurs with forgetting. The Appendix also presents other implications of the theory that will be checked against the data.

To test these predictions, we used a two-alternative forced-choice recognition test with certain additional, special test conditions. In a forced-choice test with two experimental conditions A and B, there are ordinarily four comparisons required of the subjects: AO versus AN, AO versus BN, BO versus AN, and BO versus BN. The comparisons in which an old (O) item is always paired with a new (N) item are referred to as standard choices. The subjects' preference in each of these comparisons is given as a proportion: $P(AO, AN)$, $P(AO, BN)$, $P(BO, AN)$, and $P(BO, BN)$. The first argument in each term indicates the preferred alternative. The mirror effect in Figure 1 produces the following inequalities for the standard choices:

$$P(BO, BN) < P(AO, BN),\ P(BN, AN) < P(AO, AN).$$

These proportions correspond to four of the distances between the distributions depicted in Figure 1: the distance between BO and BN, between AO and BN, and so on.

There are also two distances of special interest that are not measured directly by the proportions just mentioned: the distance between AO and BO and the distance between AN and BN. Those distances are, however, measured directly by two comparisons introduced by Glanzer and Bowles (1976): the forced choice between AO and BO and the forced choice between AN and BN. These are called null choices to indicate that there is no apparent reason for the subject to choose one over the other. The subject is faced with two old items in one case or two new items in the other. They are, however, critical choices because they give direct information about the distance AO to BO and the distance AN to BN in Figure 1. When the mirror effect holds, both $P(AO, BO)$ and $P(BN, AN)$ are greater than .5. Furthermore, according to attention/likelihood theory, when forgetting occurs the concentering of the underlying distributions lowers both $P(AO, BO)$ and $P(BN, AN)$ toward .50.

As an example of the way the theory works, we compute the proportion of choices, both standard and null, obtained for a given set of parameters. We use the binomial distributions implied by the theory. The example presented is typical of results from exploration of the system with a wide range of parameters.

Let us consider two conditions, A and B, with A giving more accurate recognition than B. The $p(\text{new})$ is assumed identical for A and B (see Assertion 2) and set for this example at .10. The N is assumed identical for A and B (see Assertion 1) and is set here at 1,000. The parameters for B, the less accurate condition, are set at $n(B) = 50$ *with, therefore,* $\alpha(B) = .05$ (because $N = 1{,}000$). This $\alpha(B)$ gives, on the basis of Equation 1, $p(B,\text{old}) = .145$. The parameters for A, the more effective condition, are set at $n(A) = 100$ with, therefore, $\alpha(A) = .10$. The latter gives, on the basis of Equation 1, $p(A,\text{old}) = .19$. These values hold for $p(A,\text{old},0)$ and $p(B,\text{old},0)$ immediately after study at time $t = 0$. During forgetting there is a reduction β, of marked features during each successive time interval; β is assumed here equal to .10 per unit of time. For example, Equation 4 can be used to give the effect of forgetting after two units of time: $p(A,\text{old},2) = .173$ and $p(B,\text{old},2) = .136$. The parameters $p(\text{new}) = .10$ and $n(i)$ are assumed to remain the same during the test, that is, after forgetting.

These two pairs of values for $p(A,\text{old},0)$, $p(B,\text{old},0)$ and $p(A,\text{old},2)$, $p(B,\text{old},2)$ can then be used to compute the six proportions of forced choice for each time interval, $t = 0$ and $t = 2$. Those proportions are determined in the following way. We set up the following succession of distributions: (a) the four binomial distributions of marked features for AO, BO,

AN, BN for $t = 0$ and the four for $t = 2$; (b) the four distributions of likelihood ratios based on them for $t = 0$ and the four for $t = 2$; (c) the six bivariate distributions of likelihood ratios for each pair of possible choice conditions (e.g., $AO \times AN$, $AO \times BN$, $AO \times BO$, and so on) for $t = 0$ and also the six for $t = 2$. Using the bivariate distributions for each time condition, we sum the probabilities of all pairs of likelihood values for which the likelihood for one of the conditions (e.g., AO) is greater than that for the other (e.g., AN). Ties in likelihood are handled by assigning half the probability to each condition. This sum gives us the probability of choice such as $P(AO, AN)$ from the $AO \times AN$ bivariate distribution. Details concerning this type of computation are presented in Bower (1972), Glanzer and Bowles (1976), and Green and Swets (1966/1974).

When this computation is performed for the parameters just discussed in the standard comparisons, we obtain the following before forgetting (at $t = 0$):

$$P(BO, BN) = .75 < P(BO, AN) = .92, P(AO, BN) = .91 < P(AO, AN) = .97.$$

This follows the mirror order.

The computed values for the null choices also conform to the mirror effect with both greater than .50.

$$P(AO, BO) = .84, P(BN, AN) = .86.$$

To show the effect of forgetting, we recompute the proportions of choice for the same parameters with $p(A,\text{old},t)$ and $p(B,\text{old},t)$ reduced by $(1-\beta)^t$. After two units of time ($t = 2$), we obtain the following. For the standard choices,

$$P(BO, BN) = .71 < P(BO, AN) = .87, P(AO, BN) = .87 < P(AO, AN) = .93.$$

These conform again to the mirror order, but have moved toward the neutral point, .50. The null choices also conform to the mirror effect after forgetting, with both greater than .50. In addition, they also show the predicted reduction toward .50.

$$P(AO, BO) = .80, P(BN, AN) = .81.$$

A more complete picture of the effect of forgetting is obtained by computing the predicted proportions of choice over an extended delay period for $t = 0, 1, 2, 3$, and so on. In Figure 2 the proportions of choice over 20 successive delay units are plotted for the parameters given earlier ($N = 1{,}000$, $p(\text{new}) = .10$, $n(B) = 50$, $n(A) = 100$, and $\beta = .10$). The proportions for both null and standard choices are shown. All approach .50, showing the predicted concentering.

The Appendix presents a formal proof of the concentering effect. It also presents a proof that the members of two pairs of forced choices come together in value, becoming approximately equal as forgetting occurs if they are not equal initially. These two pairs are the standard choices

$$P(BO, AN) \approx P(AO, BN)$$

and the null choices

$$P(AO, BO) \approx P(BN, AN).$$

This approximate equality can be seen for the computations presented previously here. It is checked in the data of the experiments that follow.

These several predictions given previously are different from what can be inferred from conventional theories of forgetting. Most of the older theories of forgetting are summarized by Hintzman (1978). They include the following: autonomous trace change, acid bath, interference, and disuse. They are all strength theories. Because they are strength theories, they would not be able, without radical revision, to cope with the mirror effect. They all focus, moveover, solely on the trace laid down by an old item and how this trace is impaired by subsequent events. These theories have in common a silence about changes in the response to new items. They could not, therefore, cope with the concentering demonstrated here. The previous statements hold as well for fluctuation theory first presented by Estes (1955) and developed further by Bower (1972). In more recent theorizing, the importance of retrieval conditions is emphasized. Strength of the cuing conditions has the role that strength of the trace had in earlier theories. That change does not enable such theories to cope with the mirror effect. The basic similarity of strength of retrievability and strength of trace is seen in the close relation of Bower's (1972) model to the Estes (1955) model from which it derives.

All of the forgetting theories mentioned previously focus on the trace of the old item. When pictured in terms of underlying distributions as in Figure 1, those theories would have the distributions move down toward the new distributions (on an axis labeled strength, or its equivalents: amount of marking, familiarity, veridicality, and retrievability). The new distributions would, presumably, stay fixed.

As noted previously, there is, in the theorizing about forgetting, a general silence concerning new distributions. One exception is in a theory presented by Bowles and Glanzer (1983) in which they propose that forgetting produced by proactive or retroactive interference involved upward movement on the decision axis of the distributions for new items. The theory was, however, written specifically for word-frequency effects and could not, therefore, handle the general mirror effect (e.g., the mirror effect produced by transformations; see Glanzer & Adams, 1990). In the final discussion, we consider some other, more recent theories that attempt direct coverage of the mirror effect.

We now report two experiments that test the predictions derived from attention/likelihood theory: concentering with forgetting and the approximate equality of critical pairs of forced choices.

Experiment 1: Within-Subjects Design

The subjects studied a list of words, half of which were high-frequency words (H) and half, low-frequency (L). The subjects were then tested immediately by two-alternative forced choice on 120 pairs of words. This immediate test included half the study words. The remaining study words were tested 1 week later. Except for the introduction of delay for half the items, the procedure and materials were basically the same as those used in Glanzer and Bowles' (1976) study.

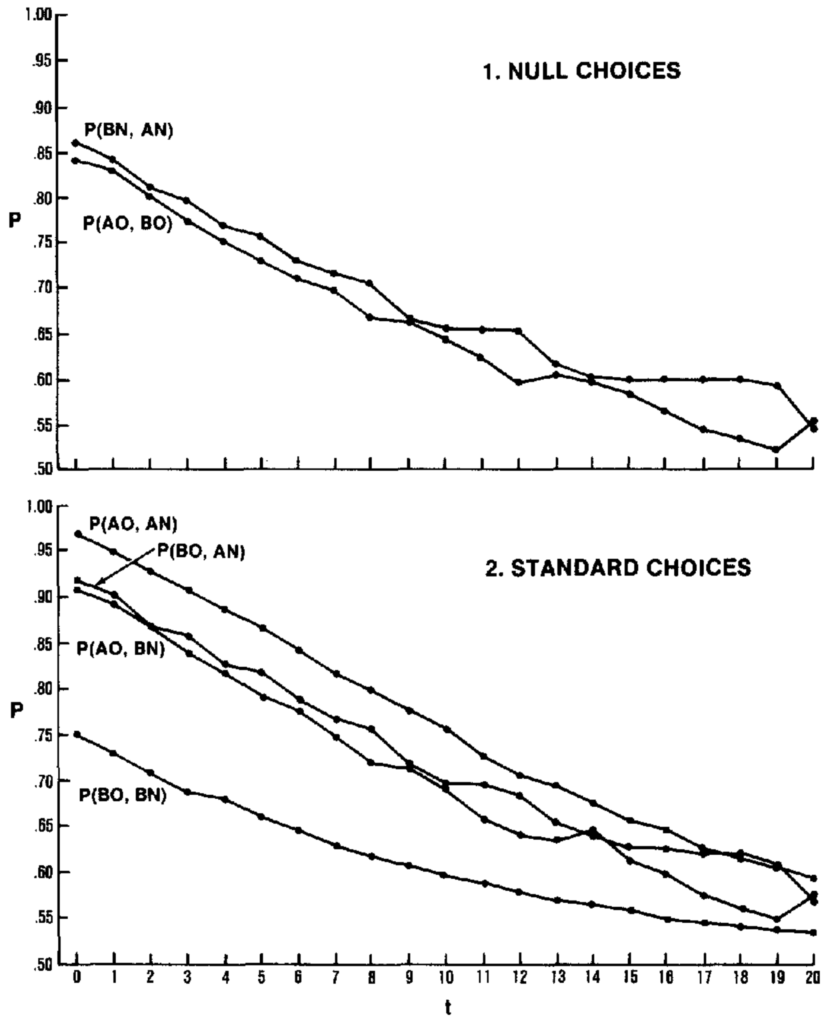

Figure 2. The proportion of choices over 21 time units according to the attention/likelihood theory. (The choices for both A and B immediately after study are at $t = 0$. P = proportion; $BN = B$ new; $AN = A$ new; $AO = A$ old; $BO = B$ old.)

All six possible types of choices were used in the forced-choice test including the null choices: *LO* (low frequency old) versus *HO* (high frequency old) and *HN* (high frequency new) versus *LN* (low frequency new).

Materials

There were four sets of words, previously used by Glanzer and Bowles (1976). They consisted of two sets of 120 high-frequency words (> 40 per million) and two sets of 120 low-frequency words (< 2 per million). The frequencies were based on summed Kučera and Francis (1967) counts. The words ranged in length from 3 to 10 letters. One of each of the two frequency sets furnished the study words. The other two sets furnished the new words for the test. The assignment of lists was counterbalanced so that across subjects the same words appeared an equal number of times in both the immediate and delay conditions and as new and old items. The study lists were individually randomized for each subject. Twelve different randomizations of the test lists were used, with lists assigned at random to subjects. Each test list was used twice across the 24 subjects.

The 120 words in each set were assigned to six subsets of 20 words each to form test pairs. The subsets were matched on concreteness and, within each frequency condition (*H* and *L*), were matched on mean frequency. These sets were rotated through the experimental conditions so that each word appeared an equal number of times as a new or old item in test pairs. The test pairs were furthermore constructed so that each pair of test words was matched on word length and concreteness. If the pair was of the same frequency class (e.g., as in the *HO*, *HN* choice), it was also matched on frequency (i.e., a word of normative frequency 50 was paired with another word of frequency 50). The procedure for evaluating concreteness is reported in Glanzer and Bowles (1976).

Subjects

The subjects were 24 undergraduate students. They participated to meet an introductory psychology course requirement.

Procedure

All subjects were tested individually. For study, they were given a set of 272 cards each bearing a single word: 240 subsequently tested words preceded by 16 filler words and followed by 16 filler words. The subjects were told to look at each word because they would be given a recognition test later. Immediately after the study period, they were tested on half the items with 120 forced-choice pairs.

The test pairs were listed in a booklet. The subjects circled the member of each pair they considered "old." After completing the immediate test, the subjects were asked to return 1 week later for the remainder of the test. The remaining 120 items were tested in that second session using the same procedure. Both study and test were self-paced.

Results

The mean proportions of choices for the six types of choice are presented in Table 1. All aspects of the mirror order are present in both the immediate and the delay conditions. In the standard pairs,

$$P(HO, HN) < P(HO, LN), P(LO, HN) < P(LO, LN).$$

In the null pairs,

$$.50 < P(HN, LN), P(LO, HO).$$

Moreover, as predicted, all the proportions decrease with delay.

For statistical analysis, the proportions in both Experiments 1 and 2 were first transformed to arc sines. The MS_es reported are therefore for the transformed data. On the overall effect of delay, analysis of variance gives $F(1, 23) = 35.421$, $p < .0001$, $MS_e = .213$. There is an interaction of choice type with delay, $F(5, 115) = 3.068$, $p = .012$, $MS_e = .082$. The interaction reflects a tendency for the larger proportions to move down further with delay than the small ones. It does not reflect a change in the ordering of means. This interaction can also be derived from the theory. It will be seen in the theoretical values computed later to fit the data.

Our main concern is with the effect of delay on the critical proportions $P(HO, HN)$, $P(LO, HO)$, and $P(HN, LN)$. Declines in these proportions are the three direct measures of concentering. All three give evidence of concentering with delay.

First, the difference between the immediate and delay conditions for $P(HO, HN)$ is statistically significant, $t(115) = 4.86$, $p < .0005$. Second, the difference between the immediate and delay conditions for $P(LO, HO)$ shows the predicted decline in the delay condition, although the effect is weak, $t(115) = 1.50$, $p < .10$. (The data of Experiment 2 furnish stronger support for this decline.) Third, and of critical importance, is the comparison of the immediate and delay conditions for $P(HN, LN)$. This also shows a statistically significant decline in the delay condition, $t(115) = 2.01$, $p < .025$. The standard error for all three tests is .082. Tests of these declines here and in Experiment 2 were one-tailed, because the direction of the difference is predicted.

These three declines differ in their importance. The decline of $P(HO, HN)$ is not critical because that could be interpreted as a unilateral movement of an old distribution toward a new distribution, expected on the basis of the conventional theories listed earlier.

The decline for the null choice $P(LO, HO)$ is of greater importance because, with one exception, only the attention/likelihood theory explicitly predicts it. Theories of forgetting generally have not addressed the question of how different old distributions move down on the decision axis in relation to each other during forgetting. Fluctuation theory (Bower, 1972; Estes, 1955) is the only one that has been developed to the point of making an explicit prediction. It would give an initial movement of old distributions toward each other as forgetting began, but then, in contrast to attention/likelihood theory, their means would maintain a constant distance as they settled at different asymptotes. (See Estes, 1955, Figure 2.)

The decline for the null choice $P(HN, LN)$ is absolutely critical. The idea that new distributions move during forgetting falls completely outside the scope of conventional theories. It cannot, therefore, be predicted by them.

An objection could be raised about the view presented here that forgetting involves movement of underlying distributions. An alternative approach could be based on the view that the variances of the underlying distributions increase with forgetting. Such an approach would require the development of a principled and explicit basis for variance in-

Table 1
Mean Proportions of Choices for the Immediate and Delay Conditions of Experiment 1 (N = 24)

	Choice					
	Null pairs		Standard pairs			
Condition	$P(HN, LN)$	$P(LO, HO)$	$P(HO, HN)$	$P(LO, HN)$	$P(HO, LN)$	$P(LO, LN)$
Immediate	.65	.60	.74	.78	.80	.83
Delay	.57	.55	.56	.62	.63	.64

Note. P = proportion; HN = high frequency, new; LN = low frequency, new; HO = high frequency, old; LO = low frequency, old.

creases. Those increases would, moreover, have to occur for new as well as old distributions.

Another objection could be raised concerning the distances estimated on the basis of the null conditions. Those involve unusual choices. We can, however, use the standard conditions to obtain a separate estimate of the critical distance, that between HN and LN. If we use the d' equivalent of $P(LO, LN)$ and the d' equivalent of $P(LO, HN)$, and subtract the second from the first, we have another estimate of the distance between HN and LN. Similarly, if we subtract the d' equivalent of $P(HO, HN)$ from the d' equivalent of $P(HO, LN)$, we obtain still another estimate of the distance between HN and LN. These differences were computed for both the immediate and the delay conditions for this and the next experiment. The results are presented in Table 2. As predicted, and in conformity with the data on the null choices, the distance between LN and HN, on the basis of both estimates, contracts with forgetting. The contraction as measured by the $d'(LO, LN) - d'(LO, HN)$ estimate is statistically significant, $t(115) = 2.214$, $p < .025$, $SE = .114$; as measured by the $d'(HO, LN) - d'(HO, HN)$ estimate, it is in the right direction but not statistically significant, $t < 1$. Overall, however, there is concurrence. The new distributions are moving as predicted during forgetting.

There are two further predictions that derive from the theory. One is that the two null choices, $P(HN, LN)$ and $P(LO, HO)$, should be approximately equal when forgetting occurs. That is indeed the case (see Table 1). Statistical analysis shows that in the immediate test condition comparison of those choices gives $t(115) = 1.19$, $p > .20$; in the delay condition, $t < 1$. The other prediction is that $P(HO, LN)$ and $P(LO, HN)$ should also be approximately equal when forgetting occurs. That is also the case. Statistical analysis shows that, in the immediate condition,[2] comparison of those choices gives $t(115) = 1.00$, $p > .20$; in the delay condition, $t < 1$. The standard error of all four tests is .082. The tests here are two tailed. The derivation of the predictions is given in the Appendix.

The adequacy of the theory can be further examined by choosing five parameters and approximating the obtained data. Those parameters are $p(\text{new}) = .15$, $n(\text{high}) = 53$, $n(\text{low}) = 73$, and the number of features, $N = 1{,}000$. The forgetting parameter was set at .146. In other words, the subjects were assumed to lose .146 of the marking of the features for old items each time unit. (Because the time unit, t, was taken as days, here 7, $(1 - \beta)^t = .33$.) Using Equation 1, these parameters give $p(\text{old,high,0}) = .195$ and $p(\text{old,low,0}) = .212$.

With the theoretical structure for forced choice described earlier, the following proportions of choice are then obtained for the immediate condition:

$P(HN, LN) = .65$, $P(LO, HO) = .62$, $P(HO, HN) = .73$,
$P(LO, HN) = .78$, $P(HO, LN) = .80$, and $P(LO, LN) = .83$.

When the effect of forgetting is evaluated with $(1 - \beta)^7 = .33$, the following are obtained using Equation 4: $p(\text{high,old,7}) = .165$ and $p(\text{low,old,7}) = .170$. The $p(\text{new})$ remains the same: .15. The recomputation of binomial distributions, likelihood distributions, and then proportions of choices gives the following:

$P(HN, LN) = .55$, $P(LO, HO) = .55$, $P(HO, HN) = .58$,
$P(LO, HN) = .61$, $P(HO, LN) = .61$, and $P(LO, LN) = .63$.

A χ^2 evaluation of the fit was made using

$$\chi^2(f) = \sum_{k=1}^{12} \frac{[\hat{P}(k) - P(k)]^2}{\sigma^2(k)}$$

where k is the choice condition (e.g., comparison of HN and LN at time 0), $\hat{P}(k)$ is the theoretical proportion, $P(k)$ is the observed proportion, and $\sigma^2(k)$ is the variance of the observed proportion in condition k. This gives $\chi^2(7) = 0.0893$. The degrees of freedom, f, here are 7 because five parameters were estimated: N, $n(\text{low})$, $n(\text{high})$, $p(\text{new})$, and β. The probability of a value this small or smaller by chance is less than .005.

Discussion

As predicted, forgetting causes all the underlying distributions to concenter, to collapse on neighbors and toward a midpoint. In particular, forgetting produces changes in the positions of new distributions as well as in the positions of old distributions. The new distributions move toward a central point, becoming more like the old distributions, and becoming less distinctively new. The paradoxical introductory statement—that when subjects forget they forget new items as well as old items—is supported.

Ordinarily, forgetting is viewed as old items changing and becoming indistinguishable from new items. Here we have

Table 2
Indirect Estimates of the Distance Between HN and LN for Experiments 1 and 2

Variable	$d'(LO, LN) - d'(LO, HN)$	$d'(HO, LN) - d'(HO, HN)$
Experiment 1		
Immediate	.30	.42
Delay	.05	.31
Experiment 2		
Immediate	.45	.65
Delay	.02	.07

Note. LO = low frequency, old; LN = low frequency, new; HN = high frequency, new; HO = high frequency, old.

[2] The theory predicts that, within the specified pairs of choices, the proportions will approximate each other during the course of forgetting. It does not predict whether these choices should differ in the immediate condition. With some parameter settings they do; with others they do not. The amount of the initial difference will determine the extent of the change in the relation between such pairs. If they are near equality in the immediate condition, not much change is predicted. In this experiment, the initial differences are not statistically significant; in Experiment 2 they are. For the conditions represented in Figure 2, $P(BO, AN)$ and $P(AO, BN)$ are nearly equal at $t = 0$. The same holds for $P(BN, AN)$ and $P(AO, BO)$. They, therefore, remain nearly equal with forgetting. If the parameters are changed so that $n(A) = 50$ and $n(B) = 25$, then at $t = 0$, $P(AO, BO) = .63$, and $P(BN, AN) = .69$. At $t = 20$ they are both .52. For the same parameters at $t = 0$, $P(AO, BN) = .67$ and $P(BO, AN) = .72$. At $t = 20$ they are both .53.

demonstrated that new items are also, on their own, changing and becoming indistinguishable from old items. The reason for the symmetry of movement of underlying distributions is that the subject evaluates every item, both new and old, with reference to a model of both new and old items. Any change that affects one has to affect the other.

One objection that could be raised about the results and their interpretation is the following. Subjects may have become aware of the special character of the null choices during their first test. This awareness may have affected their responses during the second, delayed test. This objection can be met by changing the design of the experiment: using different groups of subjects for immediate and delayed testing. Delay is then a between-subjects variable, and the immediate and delay conditions are made independent.

Experiment 2: Between-Subjects Design

The main purpose of this experiment was to replicate the findings of Experiment 1 with a between-subjects design. The general procedure was identical with that of Experiment 1, except that two groups of subjects, not one, were tested. One group was tested immediately after study. The other group was tested after a week's delay.

Subjects

Subjects were 48 undergraduate students who participated to meet an introductory psychology course requirement.

Procedure

The only difference between this experiment and Experiment 1 was that the subjects in each group received one set of 120 test pairs in their single testing session. Half the subjects were tested immediately and the remainder were tested 1 week later. All were told that they would be tested. Subjects were assigned to the immediate- or delay-testing group at random. The assignment of lists was counterbalanced so that each word appeared an equal number of times as new and old and in immediate or delay conditions. The study lists were randomized individually for each subject. As in Experiment 1, 12 different randomizations of the test lists were used, assigned at random to each subject.

Results

The mean proportions of choice for each of the six types of pairs are shown in Table 3. The data for the standard choices, with one exception, conform to the mirror pattern

$$P(HO, HN) < P(HO, LN), P(LO, HN) < P(LO, LN),$$

and for the null choices

$$.50 < P(HN, LN), P(LO, HO).$$

The exception is the reversal $P(LO, LN) < P(HO, LN)$ in the immediate condition. We consider that small reversal to be a chance variation because it does not appear with the same material in either Experiment 1 or in the delay condition of this experiment.

As before, all proportions decrease with delay. The overall effect of delay gives $F(1, 46) = 47.22$, $p < .0001$, $MS_e = .221$. There is again the expected interaction of choice type with delay, $F(5, 230) = 4.200$, $p < .002$, $MS_e = .082$.

Our main concern again is with the effect of delay on the proportions $P(HO, HN)$, $P(LO, HO)$, and $P(HN, LN)$, whether the means of the underlying distributions come together. Test of the decline in $P(HO, HN)$ gives $t(230) = 4.07$, $p < .0005$. The effect of delay on the null choice $P(LO, HO)$ is also statistically significant, $t(230) = 1.795$, $p < .05$. Finally, test of the decline on the critical null choice $P(HN, LN)$ shows it to be statistically significant, $t(230) = 4.07$, $p < .0005$. The standard error for all three tests is .083.

We can again use the standard conditions to give additional measures of the change in the critical distance between *HN* and *LN*. These estimates are presented in Table 2. Statistical tests of the decline in these estimates find concurring results. The decline in the delay condition using the $d'(LO, LN) - d'(LO, HN)$ estimate is statistically significant, with $t(230) = 3.71$, $p < .0005$; using the $d'(HO, LN) - d'(HO, HN)$ estimate, the decline is also statistically significant, $t(230) = 5.07$, $p < .0005$. The standard error for both tests is .114.

As in Experiment 1, there is clear evidence of the predicted concentering. Again as in that experiment, there is clear evidence that the new distributions, as well as the old, move when forgetting occurs.

The data in Table 3 also support the prediction that the null choices $P(HN, LN)$ and $P(LO, HO)$ should be approximately equal when forgetting occurs. Statistical analysis shows that, in the immediate condition, comparison of those choices gives $t(230) = 2.07$, $p < .05$; in the delay condition, $t < 1$.

Table 3
Mean Proportions of Choices for the Immediate and Delay Conditions of Experiment 2 (N = 24 in Each Condition)

	Choice					
	Null pairs		Standard pairs			
Condition	P(HN, LN)	P(LO, HO)	P(HO, HN)	P(LO, HN)	P(HO, LN)	P(LO, LN)
Immediate	.69	.61	.70	.75	.83	.82
Delay	.53	.54	.55	.61	.57	.62

Note. P = proportion; HN = high frequency, new; LN = low frequency, new; LO = low frequency, old; HO = high frequency, old.

The same pattern holds for $P(HO, LN)$ and $P(LO, HN)$. Comparison of those choices to the immediate condition gives $t(230) = 2.79, p < .01$; in the delay condition, $t(230) = -1.04, p > .20$. The standard error for these tests is .083.

Because the procedure and materials used in this experiment are essentially the same as those in Experiment 1, we might expect that the parameters fitted to the data of Experiment 1 would also fit the data of this experiment. The theoretical values given earlier do indeed give a reasonable although less close fit to the data of Table 3. A χ^2 evaluation gives $\chi^2(12) = 0.462$. The probability of a fit this close or closer by chance is less than .005. (We used 12 degrees of freedom because we did not estimate any parameters in this case. Using only 7 degrees of freedom does not make much difference in the statistical evaluation.)

Discussion

The experiments support the predictions drawn from attention/likelihood theory. In particular, they support the assertion that forgetting affects the means of the underlying new distributions, not just the means of the underlying old distributions.

There are also data on another type of forgetting that parallel the pattern of results found here. Those data are from an earlier study by Bowles and Glanzer (1983) on the effects of proactive and retroactive interference on recognition. (We follow the tradition of labeling conditions involving additional items interference conditions. The mechanism proposed here to handle the effects of additional items is, however, not an interference explanation if interference implies a loss of information about studied items.)

In the Bowles and Glanzer (1983) study, 80 subjects were tested, 20 assigned to each of four conditions: proactive, proactive control, retroactive, and retroactive control. In the interference conditions, there were additional studied (but not tested) words inserted before the main list (proactive condition) or after the main list (retroactive condition). All study lists consisted of equal numbers of high-frequency and low-frequency words. The subjects were tested in a two-alternative forced-choice task that included null choices. The mean proportions of choices are presented in Table 4. The means in Table 4 for all conditions show the mirror effect for forced-choice data in both the standard choices (the mirror inequalities hold) and also in the null choices (both greater than .50).

Moreover, when the interference conditions are compared with the control conditions, the interference condition means all decline toward .50. The overall decline in the standard pairs gives $F(1, 76) = 8.266, p = .006$. The decline in the null pairs is marginally significant with $F(1, 76) = 3.009, p = .087$. The operations used in this experiment were sufficient to produce the predicted pattern of forgetting effects. The effects were, however, not strong.

A question may be raised as to whether forgetting produced by additional items as in this experiment should be conceptualized in the same way as forgetting produced by delay, as in Experiments 1 and 2. In attention/likelihood theory, it probably is more appropriate to view additional items as raising the noise level, p(new), rather than reducing p(old). Within the theory, that change would produce the same general pattern of results as that found in Experiments 1 and 2.

To show that point, the data for the control and proactive conditions in Table 4 were approximated by fixing the following parameters: $N = 1{,}000$, as before; $n(\text{high}) = 60$; $n(\text{low}) = 80$; and $p(\text{new}) = .10$ for the control condition. For the proactive condition, all parameters were set at the same values as the control except for p(new), which was raised to .116. These parameters give the following proportions of choices for control:

$P(HN, LN) = .68$, $P(LO, HO) = .66$, $P(HO, HN) = .81$, $P(LO, HN) = .87$, $P(HO, LN) = .88$, and $P(LO, LN) = .91$.

For proactive,

$P(HN, LN) = .66$, $P(LO, HO) = .64$, $P(HO, HN) = .73$, $P(LO, HN) = .79$, $P(HO, LN) = .80$, and $P(LO, LN) = .84$.

The fit is good, although not as close as that obtained for the data in Table 1. The theoretical values used to fit the proactive data can be used as well for the retroactive data with only a slight increase in deviation of theoretical from observed values.

Several points should be underscored. One is that the underlying distributions we are considering are the result of a considerable amount of processing by the subject. The subject perceives the stimulus, notes the number of features, notes the number of marked features, computes a likelihood ratio

Table 4
Mean Proportions of Choices for the Control and Interference Conditions of Bowles and Glanzer (1983) (N = 20 in Each Group)

	Choice					
	Null pairs		Standard pairs			
Condition	$P(HN, LN)$	$P(LO, HO)$	$P(HO, HN)$	$P(LO, HN)$	$P(HO, LN)$	$P(LO, LN)$
Control	.68	.68	.79	.87	.86	.92
Proactive	.64	.60	.75	.80	.80	.82
Control	.68	.64	.80	.82	.84	.90
Retroactive	.67	.62	.75	.80	.82	.86

Note. P = proportion; HN = high frequency, new; LN = low frequency, new; LO = low frequency, old; HO = high frequency, old.

for each item in the recognition test, and decides on the basis of that evaluation. The difference between attention/likelihood theory and other theories of recognition is in the inclusion of information about likelihood into each decision. Other theories assume that decisions are made on the basis of the strength of the item or something equivalent to it, for example, the number of marked features.

We have not considered alternative theories that have explicitly addressed the mirror effect because they have not been sufficiently extended to cover the effects of forgetting on the mirror effect. One theory, presented by Brown (1976) and his associates (Brown, Lewis, & Monk, 1977), is centrally concerned with the mirror effect. Although that theory is similar to the attention/likelihood theory presented here in that it gives a central role to an evaluation mechanism, it would have to be considerably extended to cover forgetting and the effects of forgetting on the mirror effect. Other more recent theoretical approaches to the mirror effect would also have to be explicitly extended to cover the effects considered in this article.

Hintzman (1988) discussed ways in which the mirror effect could be approached with his multiple-trace model, MINERVA 2. The approach has been briefly outlined but not worked out. Hintzman noted the following (p. 543) about what he outlined: "... this particular version of the model is sufficiently ad hoc to discourage further development until favorable experimental evidence of a more direct nature is in hand." Hintzman's view of recognition memory as a special case of memory for frequency would make effects of forgetting in combination with the mirror effect (as with word frequency) of particular interest. The extension has not, however, been made.

Hockley and Murdock (1987) considered the mirror effect, but simply assumed that underlying distributions are ordered as in Figure 1. Their concern was in developing a general framework relating speed to accuracy in a variety of tasks, not in explaining the basis for the mirror effect. They did not discuss forgetting effects.

Gillund and Shiffrin (1984) presented a detailed approach to the mirror effect in their extension of the SAM (search of associated memory) model. Although the theory is a strength theory, they added special mechanisms, variable criterion placement, and differential scaling of items to cover the mirror effect. They assumed that the subject uses different criteria for each experimental condition. The subject, in addition, scales the familiarity of items separately on the basis of estimates of the standard deviations of the distractor distributions. These special mechanisms play the evaluation role that the likelihood ratios play in our theory. Gillund and Shiffrin (1984) are constrained, moreover, to assume that subjects use the same mechanisms, separate criteria, and scaling by standard deviations in forced-choice tests. This produces a complex picture of forced-choice performances compared with a simple strength theory such as Bower's (1972). The same work (Gillund & Shiffrin, 1984) contains explicit theorizing on the nature of forgetting. However, their analysis of the mechanisms of forgetting is not extended to cover the way in which those mechanisms combine with the mechanisms that produce the mirror effect.

In summary, recognition memory is more complex than once thought. Its complexity is signaled by the regularities we have discussed: the mirror effect and concentering. Attention/likelihood theory explains the mirror effect. It predicts the phenomenon of concentering and other regularities that appear with forgetting. It does this by having the subject base judgments on the likelihood associated with the stimulus rather than its strength. Although other theories could be adapted to cover the regularities discussed, these adaptations have not yet been made. They will, when they are made, require inclusion of an evaluation process beyond strength measures in memory decisions. They will also require the specification of the effects of forgetting on this evaluation process.

References

Adams, J. K., & Glanzer, M. (1989, April). *Generalization of the mirror effect in recognition memory to encoding.* Paper presented at the meeting of the Eastern Psychological Association, Boston, Massachusetts.

Bower, G. H. (1972). Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 85–123). New York: Wiley.

Bowles, N. L., & Glanzer, M. (1983). An analysis of interference in memory. *Memory and Cognition, 11,* 307–315.

Brown, J. (1976). An analysis of recognition and recall and of problems in their comparison. In J. Brown (Ed.), *Recall and recognition* (pp. 1–35). New York: Wiley.

Brown, J., Lewis, V. J., & Monk, A. F. (1977). Memorability, word frequency and negative recognition. *Quarterly Journal of Experimental Psychology, 29,* 461–473.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review, 62,* 145–154.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–67.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory and Cognition, 13,* 8–20.

Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 5–16.

Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 21–31.

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics.* Huntington, NY: Robert E. Krieger. (Original work published 1966)

Hintzman, D. L. (1978). *The psychology of learning and memory.* San Francisco: Freeman.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95,* 528–551.

Hockley, W. E., & Murdock, B. B., Jr. (1987). A decision model for accuracy and response latency in recognition memory. *Psychological Review, 94,* 341–358.

Kučera, F., & Francis, W. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Appendix

Proof That Forgetting Produces Concentering and the Equality of Certain Pairs of Proportions

The proof that follows shows, among other things, that concentering will occur with forgetting. It can easily be shown that the means of the several likelihood distributions all move together with forgetting.

The mean of the log-likelihood ($\ln L$) distribution is

$$M \ln L(x) = n(i)\cdot p(i,j)\cdot \ln(p(i,\text{old})/p(\text{new})) + n(i)\cdot q(i,j)\cdot \ln(q(i,\text{old})/q(\text{new})),$$

where i is the experimental condition, such as stimulus set A or B, j is the stimulus state, either new or old, $p(i,j)$ is the proportion of marked elements, $q(i,j)$ is its complement, and $n(i)$ is the number of features examined in condition i. Note that $q(\text{new}) = 1 - p(\text{new})$. It is obvious that as $p(i,\text{old})$ decreases toward $p(\text{new})$ and as $q(i,\text{old})$ increases toward $q(\text{new})$ during forgetting, $M \ln L(i,j)$ approaches zero for all conditions. However, the variances of the log likelihoods, where

$$\text{Var} \ln L(x) = n(i)\cdot p(i,j)\cdot q(i,j)\cdot \{\ln(p(i,\text{old})q(\text{new})/p(\text{new})q(i,\text{old}))\}^2,$$

also change with forgetting. It is necessary, therefore, to prove, as is done in the following, that the approach of the means toward each other is not countervailed by the declines in the variances that also occur with forgetting.

All forced-choice predictions involve a pairwise comparison of random variables (log-likelihood ratios) of the form

$$\delta + \epsilon \cdot r,$$

where δ, $\epsilon > 0$ are constants depending on stimulus-state condition (old [O] vs. new [N]) and experimental condition (A vs. B); r is binomially distributed with parameters n, p (whose values are, again, fixed by condition). We write q for $1 - p$.

We focus discussion on the forced-choice probability for AO versus BN. The same assertions hold for the other choices with an appropriate change of indices.

The forced-choice probability is

$$P(AO, BN) = P[\delta(AO) + \epsilon(AO)\cdot r(AO) > \delta(BN) + \epsilon(BN)\cdot r(BN)] \qquad \text{(A1)}$$

with

$$\begin{aligned}
\delta(AO) &= n((A)\cdot \ln(q(AO)/q(N)),\\
\epsilon(AO) &= \ln(p(AO)\cdot q(N)/q(AO)\cdot p(N)),\\
\delta(BN) &= n(B)\cdot ln(q(BO)/q(N)),\\
\epsilon(BN) &= \ln(p(BO)\cdot q(N)/q(BO)\cdot p(N)).
\end{aligned}$$

The random variables $r(AO)$, $r(BN)$ are independent and binomially distributed with parameters $n(A)$, $p(AO)$ and $n(B)$, $p(N)$, respectively; that is, $E(r(AO)) = n(A)\cdot p(AO)$, $\text{Var}\,(r(AO)) = n(A)\cdot p(AO)\cdot q(AO)$, and similarly for the moments of $r(BN)$.

There is no simple exact expression for $P(AO, BN)$, because the distribution of $\epsilon(AO)\cdot r(AO) - \epsilon(BN)\cdot r(BN)$ is complicated. However to a good approximation $\epsilon(AO)\cdot r(AO) - \epsilon(BN)\cdot r(BN)$ is normally distributed, with

$$\text{M} = n(A)\cdot p(AO)\cdot \epsilon(AO) - n(B)\cdot p(N)\cdot \epsilon(BN),$$
$$\text{Var} = n(A)\cdot p(AO)\cdot q(AO)\cdot \epsilon(AO)^2 + n(B)\cdot p(N)\cdot q(N)\cdot \epsilon(BN)^2.$$

With that approximation, we have

$$P(AO, BN) = \Phi[\underline{\delta(AO) + n(A)\cdot p(AO)\cdot \epsilon(AO) - [\delta(BN) + n(B)\cdot p(N)\cdot \epsilon(BN)]]} \,/[\sqrt{n(A)\cdot p(AO)\cdot q(AO)\cdot \epsilon(AO)^2 + n(B)\cdot p(N)\cdot q(N)\cdot \epsilon(BN)^2}], \qquad \text{(A2)}$$

where Φ is the distribution function of the unit normal.

Under forgetting (i.e., the limit as $t \to \infty$, or $\beta \to 1$) $q(AO) \to q(N)$, $p(AO) \to p(N) = 1 - q(N)$, and thus $\delta(AO) \to 0$, $\epsilon(AO) \to 0$. Similarly, for the same reasons, $\delta(BN)$, $\epsilon(BN)$ each converge to 0. It is not immediately apparent what the limit of $P(AO, BN)$ is, because both numerator and denominator in Equation A2 are converging to zero.

We show that the limit is in fact zero (because the numerator approaches 0 faster than the denominator—in fact, as the square of the denominator). In the following, we write θ for the term $(1 - \beta)^t$ in Equation 4 of the text.

As $\theta \to 0$, the limiting behavior of $\delta(AO) + n(A)\cdot p(AO)\cdot \epsilon(AO)$ is dictated by that of

$$\ln(1 - \theta\cdot\alpha(A)) \text{ and } \ln(1 + \theta\cdot\alpha(A)\cdot[q(N)/p(N)]).$$

In fact, using the Taylor expansion $\ln(1 + x) = x - x^2/2 + x^3/3 - \ldots$, we see that

$$\begin{aligned}
\delta(AO) + n(A)\cdot p(AO)\cdot \epsilon(AO) &= n(A)\cdot[\ln(q(AO)/q(N))\cdot[1 - p(AO)]] + n(A)\cdot p(AO)\cdot \ln(p(AO)/p(N))\\
&= n(A)\cdot q(AO)\cdot \ln(1 - \theta\cdot\alpha(A)) + n(A)\cdot p(AO)\cdot \ln(1 + \theta\cdot\alpha(A)\cdot[q(N)/p(N)])\\
&= n(A)\cdot q(AO)\cdot[-\theta\cdot\alpha(A) - \theta^2\cdot[\alpha(A)^2/2] - \ldots] + n(A)\cdot p(AO)\cdot [\theta\cdot\alpha(A)\cdot[q(N)/p(N)] - \theta^2\cdot[\alpha(A)^2/2]\cdot[q(N)^2/p(N)^2] + \ldots]\\
&= n(A)\cdot q(N)\cdot[1 - \theta\cdot\alpha(A)]\cdot[-\theta\cdot\alpha(A) - \theta^2\cdot[\alpha(A)^2/2] - \ldots]\\
&\quad + n(A)\cdot[p(N) + \theta\cdot\alpha(A)\cdot q(N)]\cdot[\theta\cdot\alpha(A)\cdot[q(N)/p(N)] - \theta^2[\alpha(A)^2/2]\cdot[q(N)^2/p(N)^2] + \ldots].
\end{aligned}$$

The terms linear in θ cancel, and we obtain

$$\delta(AO) + n(A)\cdot p(AO)\cdot \epsilon(AO) = [\theta^2\cdot n(A)\cdot\alpha(A)^2/2]\cdot[q(N)/p(N)] + \text{terms in } \theta \text{ of higher order.}$$

In the same way,

$$\delta(BN) + n(B)\cdot p(N)\cdot \epsilon(BN) = [-\theta^2\cdot n(B)\cdot\alpha(B)^2/2]\cdot[q(N)/p(N)] + \text{higher order terms.}$$

The variance can be treated in exactly the same way. We obtain

$$n(A)\cdot p(AO)\cdot q(AO)\cdot \epsilon(AO)^2 + n(B)\cdot p(N)\cdot q(N)\cdot \epsilon(BN)^2 = \theta^2\cdot[n(A)\cdot \alpha(A)^2\cdot[q(N)/p(N)] + n(B)\cdot\alpha(B^2\cdot[q(N)/p(N)])$$

plus higher order terms and so (to an excellent approximation, for small θ)

$$P(AO,BN) = \Phi[[\theta^2/2][n(A)\cdot\alpha(A)^2 - n(B)\cdot\alpha(B)^2]\cdot[q(N)/p(N)]] \,/\sqrt{\theta^2\cdot[n(A)\cdot\alpha(A)^2 + n(B)\cdot\alpha(B)^2]\cdot[q(N)/p(N)]}.$$

We see that $P(AO, BN) > \frac{1}{2}$ because the argument of Φ is > 0, and that as $\theta \to 0$, $P(AO, BN) \to \Phi[0] = \frac{1}{2}$.

To summarize, for small θ, ln likelihood ratio (L) is approximately normal for all conditions with

$$\begin{aligned}
&M \ln L(AO) = [\theta^2/2]\cdot n(A)\cdot\alpha(A)^2,\ \text{Var} \ln L(AO) = \theta^2\cdot n(A)\cdot\alpha(A)^2\ (= 2 \times \text{mean}),\\
&M \ln L(AN) = [-\theta^2/2]\cdot n(A)\cdot\alpha(A)^2,\ \text{Var} \ln L(AN) = \theta^2\cdot n(A)\cdot\alpha(A)^2,\\
&M \ln L(BO) = [\theta^2/2]\cdot n(B)\cdot\alpha(B)^2,\ \text{Var} \ln L(BO) = \theta^2\cdot n(B)\cdot\alpha(B)^2,\\
&M \ln L(BN) = [-\theta^2/2]\cdot n(B)\cdot\alpha(B)^2,\ \text{Var} \ln L(\text{BN}) = \theta^2\cdot n(B)\cdot\alpha(B)^2.
\end{aligned}$$

In this approximation then, AO and AN are symmetrically placed about 0, as are BO and BN.

In this approximation (that is, when forgetting is dominant), we also have the predictions

$$P(AO, BN) \approx P(BO, AN)$$
$$\text{and } P(AO, BO) \approx P(AN, BN).$$

The following two points should be noted.

1. The previous approximation can be improved, of course, by taking more terms in the Taylor expansion.

2. The exact theory, based on the binomial, predicts only approximate equality between $P(AO, BO)$ and $P(AN, BN)$, and between $P(AO, BN)$ and $P(BO, AN)$. The data show small but systematic violations of equality between these quantities in agreement with the exact theory.

The predictions $P(AO, AN) > P(AO, BN)$, and so on, follow immediately. They follow from the general form

$$\text{Choice prob} = \Phi[X - Y]/[\sqrt{2} \cdot \sqrt{X + Y}],$$

which is easily shown to be monotone increasing in X, decreasing in Y. For example,

$$P(AO, AN) = \Phi[[X - [-X]]/[\sqrt{2} \cdot \sqrt{X + X}]]$$

with $X = \frac{1}{2}\,\theta^2 \cdot \alpha(A)^2 \cdot n(A) \cdot [q(N)/p(N)]$,

and because $-X < -Y$, $Y = \frac{1}{2} \cdot \theta^2 \cdot \alpha(B)^2 \cdot n(B)$ (for $n(A) > n(B)$ and $\alpha(A) > \alpha(B)$), we have

$$P(AO, AN) > \Phi[[X - [-Y]]/\sqrt{2} \cdot \sqrt{X + Y}] = P(AO, BN).$$

Received October 24, 1989
Revision received March 21, 1990
Accepted May 22, 1990 ■