Journal of Experimental Psychology:
Learning, Memory, and Cognition
1985, Vol. 11, No. 2, 397–406

0278-7393/85/$00.75

Observations

Evaluating Forgetting Curves

Geoffrey R. Loftus
University of Washington

A new method is described for determining the effect of original learning (or any other variable) on forgetting. The major question is, How much forgetting time is required for memory performance to fall from any given level to some lower level? If this time is the same for different degrees of original learning, then forgetting is not affected by degree of original learning. If this time is greater for higher degrees of original learning, then forgetting is slower with higher original learning. Application of the method to a variety of forgetting data indicated that forgetting is slower for higher degrees of original learning.

Slamecka and McElree (1983) performed three experiments to determine how degree of original learning affects forgetting from long-term memory. Subjects learned verbal material to one of two levels of proficiency and were then tested at a delay interval ranging from 0.0 to 5.0 days. A variety of different kinds of information were tested by free recall, cued recall, and recognition. Very regular data were obtained; the degree of original learning did not interact with delay interval. Slamecka and McElree concluded that forgetting was independent of degree of original learning.

This conclusion follows if "forgetting" is operationally defined to be the slope of the forgetting function between any two delay intervals. It is not clear, however, that this definition will ultimately prove to be the most useful in illuminating the processes that underlie forgetting. Consider a physical analogy, that of radioactive decay. Imagine two chunks of radioactive material, identical except in size; the smaller chunk weighs 10 units, and the larger chunk weighs 20 units. Suppose the two chunks decay exponentially with identical decay parameters. If the decay parameter is 1.0, then decay functions could be defined as

$$\text{Small chunk: } W(t) = 10e^{-t}$$

and

$$\text{Large chunk: } W(t) = 20e^{-t},$$

where $W(t)$ is the weight (e.g., W in g) of remaining material[1] at time t (e.g., t in days).

The curves resulting from these two equations are shown in Figure 1. It is evident that there is an interaction between delay and amount of original material; the weight difference between large and small chunks is greater at shorter delays. A physicist, analyzing these data by conceptualization analogous to that of Slamecka and McElree, would conclude that radioactive decay depends on the original chunk size. At a descriptive level, this conclusion would be entirely valid. But it would not capture the simple underlying process, which is that the decay parameter does not depend on chunk size.

The physicist analyzing the radioactive decay data of Figure 1 enjoys two advantages over a psychologist analyzing forgetting data of the sort provided by Slamecka and McElree. First, the physicist has both a clearly defined underlying variable (amount of radioactive material in the chunk of material) and an empirical measure (weight) that is linearly related to it. The psychologist, in contrast, typically has only a vaguely defined underlying variable (e.g., "information") and an empirical measure (performance on some memory test) that can be assumed to be only monotonically related to it.

The physicist's second advantage is that, on the basis of experience and/or fundamental physical

This article was supported by National Science Foundation Grant BNS82-09617 to the author. I thank Donald Bamber, Robert Bolles, Earl Hunt, Elizabeth Loftus, Thomas Nelson, Walter Nelson, John Palmer, Richard Shiffrin, Norman Slamecka, and Willem Waggoner for helpful comments on earlier drafts.

Requests for reprints should be sent to Geoffrey R. Loftus, Department of Psychology, University of Washington, Seattle, Washington 98195.

[1] To be technically correct, a chunk of radioactive material does not actually decay to nothing; rather, it decays to some radio-inert substance. In the present discussions, *chunk* refers to that portion of the radioactive chunk that actually decays away. This caveat does not alter the logic of any of the arguments.

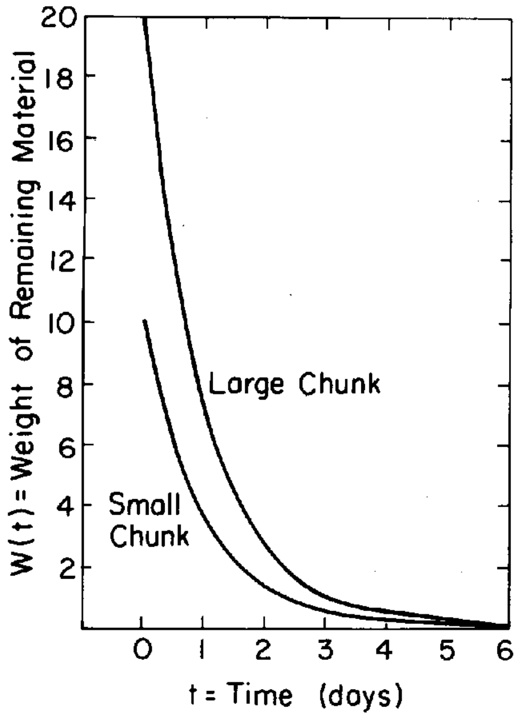

Figure 1. Decay curves (weight as a function of time) for large (20-unit) and small (10-unit) chunks of radioactive material.

considerations, he or she has quantitative laws relating the relevant variables to one another. In particular, it is reasonable to postulate that weight is linearly related to the amount of radioactive material which, in turn, is exponentially related to time. This means that exponential functions can be fit to the observed decay curves, and the relations among curves can be characterized in terms of differences among parameters of these functions. The psychologist, in contrast, typically does not have uniformly justifiable quantitative laws relating time, amount of information, and performance on a memory test.

Despite these difficulties, it turns out that certain general models of forgetting, and of the influence of variables such as degree of original learning on forgetting, making surprisingly strong predictions about relations among forgetting curves. In this article, I describe these predictions. For ease of discourse, the predictions are first illustrated within the context of a very specific model of information decay and of the relation between information and performance. It will then be shown that this specific model is a member of a much more general class of models and that the predictions also hold for this more general class. It is this latter result that I wish to emphasize.

The arguments will be applied to degree of original learning, the variable of concern to Slamecka and McElree. It should be noted, however, that they apply, more generally, to any variable that affects forgetting.

Theory

The Specific Model

Suppose that forgetting is characterized by the following assumptions:

1. Original learning produces some amount of information in memory. The higher the original learning, the greater the amount of information.

2. Following learning, the amount of retrievable information decays exponentially over time. (The term *decay* is used simply for ease of discourse and does not imply a particular theory of forgetting).

3. Performance (number of items recalled or recognized) is a linear function of information. For simplicity, assume that $P(t) = I(t)$ where P and I are performance (e.g., number of items recalled) and the amount of information, respectively, at time t following learning.

As noted, the model is made this specific in order to facilitate the illustration of predictions. However, the model is one that has been proposed in the past (e.g., Atkinson & Shiffrin, 1968; see also Murdock & Cook, 1960).

Consider a low-learning situation in which 10 units of information are originally stored in memory. If the information decays with a decay parameter of 1.0, the forgetting function is

$$P(t) = 10e^{-t} \quad (1)$$

(t in days).

Now consider a high-learning situation in which 20 units of information are originally stored. How does degree of original learning affect forgetting? Two possible theories are embodied in two versions of Assumption 4.

4a. Forgetting, as reflected by the information decay parameter, is unaffected by degree of original learning.

4b. Forgetting, as reflected by the information decay parameter, is slower in a high-learning condition relative to a low-learning condition.

Illustrative high-learning forgetting functions that correspond to these two theories are

$$\text{Decay rate unaffected: } P(t) = 20e^{-t} \quad (2a)$$

and

$$\text{Decay rate decreased: } P(t) = 20e^{-t/2}. \quad (2b)$$

Figure 2 shows the predictions of these two theories. Figure 2a shows low- versus high-learning forgetting curves, assuming that the degree of original learning does not affect forgetting (Equa-

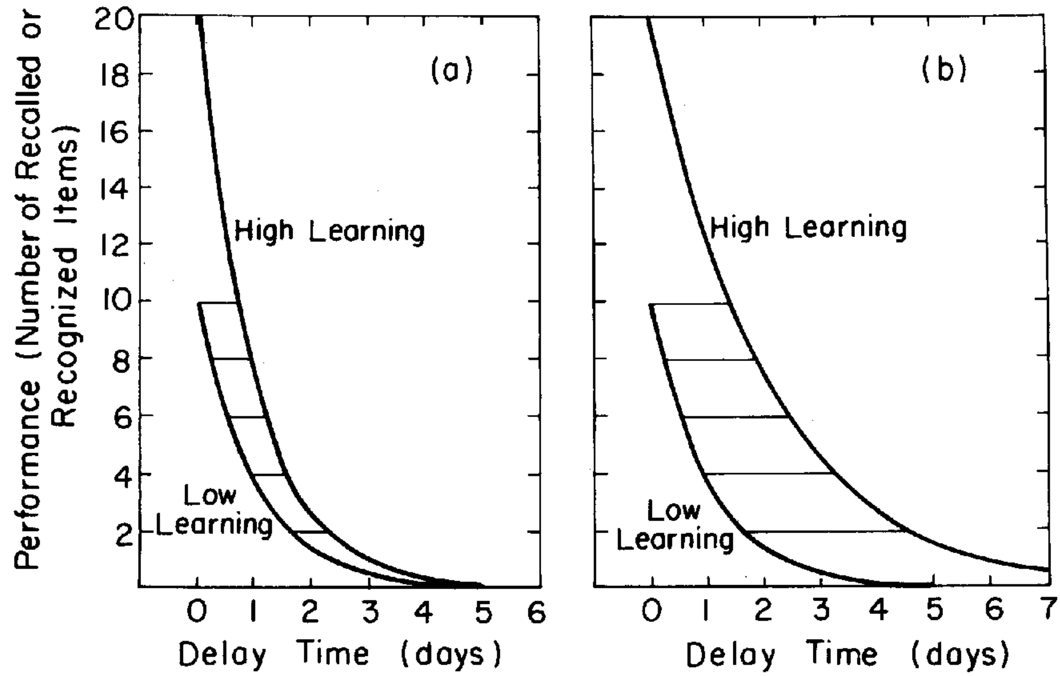

Figure 2. Hypothetical forgetting curves (performance as a function of time) for high original learning (20 units of information) and low original learning (10 units of information) conditions. (Panel a represents forgetting curves resulting from functions with the same decay parameters. Panel b represents forgetting curves resulting from functions with different decay parameters. Decay is slower following high learning than following low learning. In both panels, the horizontal lines illustrate the magnitudes of the horizontal differences between the curves.)

tions 1 and 2a), and Figure 2b shows low- versus high-learning forgetting curves assuming that higher learning slows forgetting (Equations 1 and 2b).

The predictions shown in Figures 2a and 2b differ in several ways. Of importance for the present argument, however, is the *horizontal relation* between the high- and low-learning forgetting curves. Consider first the curves shown in Figure 2a (which are identical to those shown in Figure 1). They are horizontally parallel.[2] This property is proven in Appendix 1a, and can be most easily understood by returning to the radioactive decay analogy. At some decay interval (specifically, 0.69 days) the 20-unit chunk will have decayed to the point where only 10 units remain. At that point, the larger chunk has become physically identical to what the smaller chunk had been at time zero. Thus, large-chunk decay from time 0.69 days onward is the same as small-chunk decay from time zero onward. Except for a lateral shift, the two decay curves must therefore be identical.

The logic of the forgetting situation is the same. High-learning information will, after 0.69 days, decay to the starting amount of low-learning information. Thus, the high-learning curve in Figure 2a is identical to the low-learning curve except that it is shifted to the right by 0.69 days. This property is indicated by the constant-length horizontal curve connections in Figure 2a.

The two curves shown in Figure 2b derive from the theory that higher learning slows decay rate. These curves diverge horizontally; that is, the horizontal difference between them increases as performance level decreases. This property is proven in Appendix 1b. Intuitively, it can be seen that, because of the slower forgetting, it takes longer in the high-learning condition relative to the low-learning condition for performance to fall from any given level, x, to any lower level, y.

To summarize thus far, the two theories of how original learning affects forgetting make very spe-

---

[2] It is worth emphasizing the distinction between horizontally and vertically parallel curves. Vertically parallel curves are those associated with lack of interaction in the analysis of variance model; for example, those presented by Slamecka and McElree. In contrast, two curves are horizontally parallel if the horizontal difference between them does not depend on performance level.

cific predictions about the horizontal relation between the high- and low-learning forgetting curves. If original learning does not affect forgetting, then the two curves must be identical except for a constant lateral shift. If original learning does affect forgetting, then the horizontal difference between them is not constant. It must either increase as performance level decreases (if, as in the present example, decay is slower for the high-learning condition relative to the low-learning condition) or decrease as lower performance level decreases (if decay is slower for the low-learning condition relative to the high-learning condition).

The General Model

The logic so far has been predicated on the model defined by Assumptions 1–3, above, along with one of the two versions of Assumption 4. But this model can be generalized considerably without affecting the predictions that have been derived. In particular, the notion that information is a quantitative substance, whose amount in memory is increased by learning and decreased by forgetting, can be replaced by the more general notion that learning and forgetting simply consist of passage by the cognitive system through a series of qualitative states. With this foundation, Assumptions 1–3 can be replaced by the following:

1. There is a one-to-one correspondence between state of the cognitive system and memory performance.[3]

2. The order in which the cognitive system passes through states is unaffected by degree of original learning.

3. There exists a delay interval, tj, such that $PL(0) = PH(tj)$, where $PL(t)$ and $PH(t)$ are low- and high-learning performance, respectively, at delay interval, t.

This new model is more general in that it is implied by, but does not imply, the original model.

Consider now a person in a high-learning condition. Original learning may be characterized as achievement by the cognitive system of some state, $S0$, at a delay interval of 0. Whatever activity is engaged in by the system at delay intervals, $t1$, $t2$, . . . , tj, $tj + 1$, $tj + 2$, . . . , corresponds to its progression through a series of subsequent states, $S1$, $S2$, . . . , Sj, $Sj + 1$, $Sj + 2$, ... In the normal course of events, this progression constitutes forgetting; memory performance based on a later state is lower than memory performance based on an earlier state.

Whereas initial high learning is achievement by the cognitive system of state $S0$, initial low learning is achievement only of some state, Sj, a state that, in the high-learning condition, does not occur until delay interval tj following learning. Subsequently, following low learning, the system passes through the same states, $Sj + 1$, $Sj + 2$, . . . , that are also passed through in the high-learning condition. In short, the activity engaged in from time 0 onward in the low-learning condition is identical to the activity engaged in from time tj onward in the high-learning condition.

Within the context of this more general model, we can now reconsider the two forgetting hypotheses described earlier—that degree of original learning (a) does not affect forgetting or (b) does affect forgetting. Briefly, the predictions made by these two hypotheses remain the ones indicated in Figure 2. This assertion is proven in Appendix 2. Intuitively, the reasons for it are as follows.

Suppose first that original learning does not affect forgetting. In that case, the system will proceed at the same rate through states, Sj, $Sj + 1$, $Sj + 2$, . . . , beginning immediately in the low-learning condition, but beginning at delay interval tj in the high-learning condition. Thus, low-learning performance at any time t must always be the same as high-learning performance at time $t + tj$. This implies that the high- and low-learning forgetting curves are the same except they are laterally shifted relative to one another, as depicted in Figure 2a.[4]

Now suppose that original learning affects forgetting in that it takes longer for the system to transit from any given state to any subsequent state in the high-learning condition relative to the low-learning condition. In that case, the forgetting time required for performance to drop from any level to any lower level is greater for the high-learning condition than it is for the low-learning condition. This implies that the high- and low-learning curves diverge as depicted in Figure 2b.[5]

---

[3] The term, "state of the cognitive system" is used in a restricted way. The cognitive system referred to is whatever subset of the system that is affected by the learning of the new material under consideration, along with retrieval strategies engendered by the memory test.

[4] Mathematically, this prediction may be expressed by the equation, $PL(t) = PH(t + tj)$, where PL and PH represent performance in low- and high-learning conditions, respectively, and tj is the time, following high learning, for the cognitive system to reach state Sj, the initial state achieved following low learning.

[5] Mathematically, this prediction may be expressed by the equation, $PL(t) = PH(f(t) + tj)$, where PL, PH, and tj are defined as in Footnote 1, $f(t)$ is monotonic with t, and $f(0) = 0$.

There is a stronger model of how original learning influences forgetting, which makes a correspondingly stronger prediction. The model is that the time required to transit from any state to any subsequent state differs by a factor of k in a high-learning condition relative to a

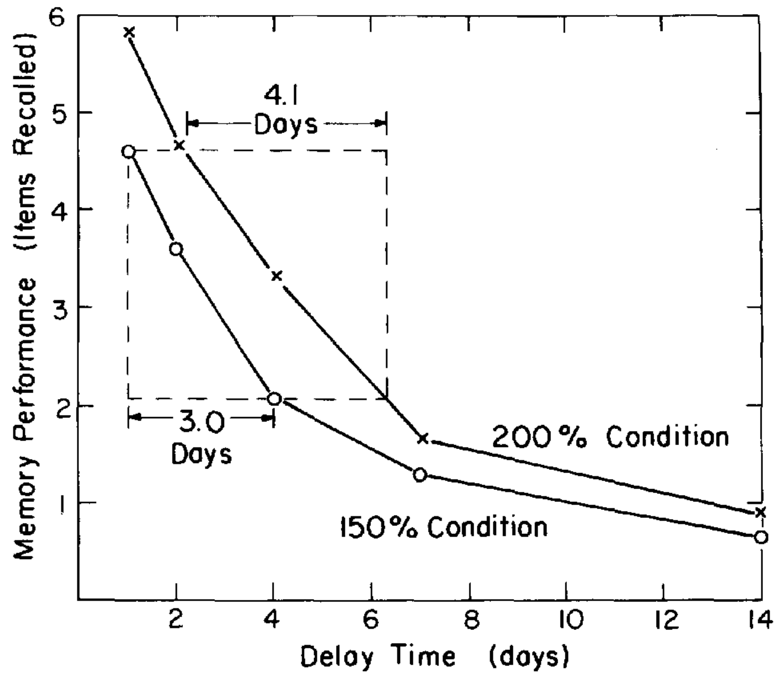

Figure 3. Forgetting curves from Krueger's (1929) study. (For simplicity, only two of the three initial-learning conditions are included.)

Data

Within this theoretical framework, evaluation of forgetting data is relatively straightforward. To be appropriately compared, however, any pair of forgetting curves must have the following properties. First, performance must be tested at a sufficient number of delay intervals that the shapes of the curves become reasonably apparent. Second, the curves must include overlapping performance ranges. Forgetting curves presented in the literature have these properties to varying degrees, but inspections of them leads to a reasonably firm conclusion.

Higher Learning Produces Slower Forgetting

For purposes of illustration, consider Krueger's (1929) 200% and 150% conditions (Slamecka & McElree, 1983, Figure 1, top panel), which are reproduced in Figure 3. These two conditions are overlearning conditions in which subjects received 200% and 150%, respectively, of the number of list-learning trials required for bare mastery of a serial list. The curves indicate, for example, that a performance level of about 4.65 items is reached by 1.0 day in the 150% condition and by 2.2 days in the 200% condition. How long does it take for performance to fall to a level of, say, 2.05 items? A 2.05-item performance level is reached at about 4.0 days in the 150% condition, but not until about 6.3 days in the 200% condition. Thus, the drop takes 3.0 days in the 150% condition (4.0 − 1.0 = 3.0 days), but 4.1 days in the 200% condition (6.3 − 2.2 = 4.1 days). The additional 1.1 days required for the 4.65- to 2.05-item performance drop in the 200% condition relative to the 150% condition indicates that forgetting is slower in the 200% condition.

Figure 4 shows similar forgetting data from four other studies: Hellyer (1962), Postman and Riley (1959), Slamecka and McElree (1983, Experiment 3), and Youtz (1941). In all cases, the same effect is found: Forgetting is slower under higher-learning conditions than it is under lower-learning conditions. These four studies incorporate a variety of different procedures and ranges of retention intervals. The divergence of high- and low-learning forgetting curves appears to be a quite robust phenomenon.

---

low-learning condition (where $k > 1$ corresponds to slower forgetting in the high-learning condition and $k < 1$ corresponds to slower forgetting in the low-learning condition). The prediction of this model becomes, $PL(t) = PH(kt + tj)$. The curves shown in Figure 2b satisfy this stronger model, with parameter values, $tj = 1.39$ and $k = 2$.

Jost's Law

This conclusion is not new. On the basis of relearning data, Jost (1897, cited in Woodworth & Schlosberg, 1961, p. 730) hypothesized that, "If two associations are now of equal strength but of different ages, the older one will lose strength more slowly with the further passage of time." Indeed, Youtz (1941) provided an empirical test of Jost's law by using a technique that incorporated the same kind of horizontal comparison that is described here.

Different Forgetting Definitions

The present definition of how forgetting is affected by the degree of original learning differs both from the one suggested by Slamecka and McElree (1983) and from others described by Slamecka and McElree (e.g., Underwood & Keppel, 1963). Given these latter definitions, evaluation of different hypotheses rests on consideration of standard statistical interactions; on a determination of whether the high-learning/low-learning performance difference depends on delay interval. Given the present definition, however, evaluation of different hypotheses rests on consideration of a different kind of interaction; "a horizontal interaction," so to speak. The principal question posed in this evaluation method is, Does the time required for performance to drop from level x to level y depend on whether original learning was high or low?

Returning now to Slamecka and McElree's original question, does forgetting depend on degree of original learning or not? Given Slamecka and McElree's definition of how forgetting is affected by original learning, the answer is *no;* given the present definition, the answer is *yes.*

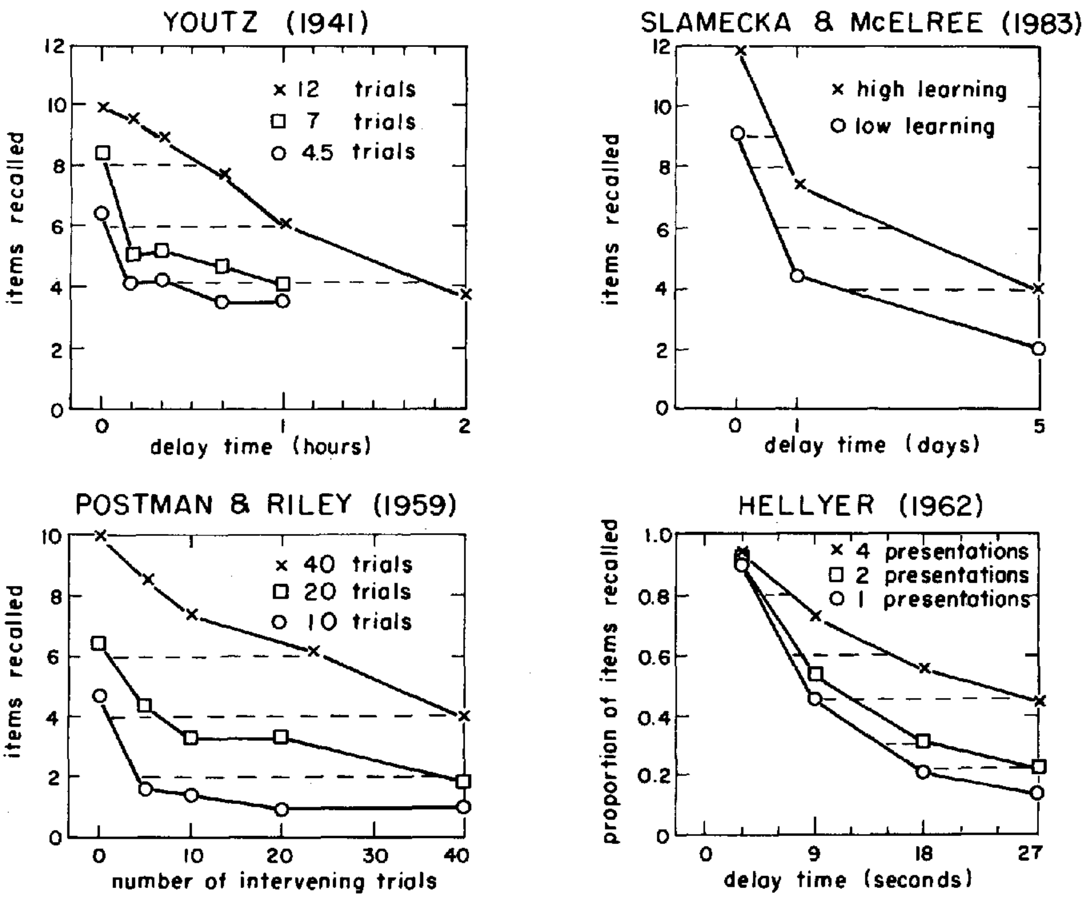

Figure 4. Forgetting curves from four studies in the literature. (For simplicity, only three of Hellyer's four initial-learning conditions are included in the bottom right panel. In all cases, the curve parameter is a measure of degree of original learning. Slamecka and McElree, 1983, provide details about the experimental procedures.)

Ultimately, as Slamecka and McElree point out, the choice of definition will probably depend more on taste and preference than on anything else. These are, however, several advantages, as well as one disadvantage, of the present definition and the corresponding evaluation method in comparison with those proposed by Slamecka and McElree and others.

Scaling

As has been noted elsewhere (e.g., Anderson, 1961; Bogartz, 1976; Krantz & Tversky, 1971; Loftus, 1978), a nonordinal interaction can be made to appear or disappear at will by applying suitable nonlinear transformations to the dependent variable. This unfortunate fact raises a problem for any theory designed to formally account for the lack of interaction on which Slamecka and McElree's conclusion depends: If such a theory were to make the correct prediction for some particular dependent variable (for example, response probability in a recognition test), then it could not make the correct prediction for any other nonlinearly related, dependent variable (for example, d').

In contrast, the present method is relatively indifferent to the choice of dependent variable. A conclusion based on one dependent variable will hold for (a) any other monotonically related, dependent variable and (b) any unobservable, unidimensional construct (e.g., information) that is assumed to be monotonic with the observed dependent variable.

Floor and Ceiling Effects

Slamecka and McElree (1983) candidly point out that they do not consider situations in which performance is either relatively high or relatively low. Such situations produce high-learning/low-learning performance differences that are dismissed as "artifically constricted." In contrast, the present method is not so restricted by such floor and ceiling effects. Consider, again by analogy, the two radioactive decay curves in Figure 1. In theory, the horizontal difference between them remains constant, even at long delay intervals where they both approach the floor. If, indeed, the delay interval were extended to the point at which both curves hit the floor—to the point at which the radioactive material in both the small and large chunks had completely disappeared—the two curves would, nonetheless, still have the same shape; they could still be made to completely overlap by shifting one horizontally relative to the other.

The principal difficulty raised by floor and ceiling effects is a practical one. As forgetting curves approach floor, for example, they typically become flatter; thus, a fixed amount of vertical noise implies progressively more horizontal noise. This means that a test of whether the horizontal difference between two curves is or is not constant becomes progressively less reliable as floor is approached.

Methodological Considerations

If one wishes to evaluate forgetting data by using standard interaction methods, then a relatively straightforward methodology is appropriate. One constructs a factorial design, with delay as one factor and something else, such as degree of original learning, as the other factor.

The appropriate methodology is not this simple if the present evaluation method is to be used. Ideally, one should choose various performance levels and measure the times required to reach these performance levels for different levels of, say, degree of original learning. In principle, this can be done, but in practice, it is difficult. In practice, one will usually end up interpolating between data points, as was shown earlier in the examples of Figures 3 and 4.

The Domain of the General Model

A data evaluation method should be founded on some reasonably explicit model, or models, of the phenomena being evaluated. Accordingly, the appropriateness of the evaluation method depends on the validity of the models.

The present evaluation method is implied by various models, including the general model of learning and forgetting that was described earlier. In some experimental paradigms, this general model—in particular, the assumption that there is a one-to-one correspondence between performance and state of the cognitive system—is probably false. Suppose, for instance, that high and low learning were obtained by imagery and rote repetition instructions, respectively. Intuitively, the proposition that equal performance implies identical states of the system seems implausible. A hint of such implausibility is the finding that two dependent variables, response probability and reaction time, often assumed to both measure an underlying unidimensional memory structure, are affected differently by a rote repetition/imagery manipulation (Peterson, Rawlings, & Cohen, 1977). As MacLeod and Nelson (in press) point out, explanation of these (and other) results requires postulation of a multidimensional memory structure, rather than a unidimensional structure, such as amount of information.

That memory must be multidimensional does not pose a problem for the general model as long as the different dimensions affect different dependent variables. Multidimensionality would render the model invalid if different combinations of multiple memory attributes led to the same values of the same dependent variable (e.g., to the same value of response probability.) In such a situation, however, comparing two forgetting curves in an experiment that incorporated only a single memory measure would not be very illuminating, no matter what the underlying theoretical framework. It would seem in the best interests of researchers studying forgetting processes to design experiments such that, as far as possible, a particular value of performance does imply a unique state of the cognitive system.

References

Anderson, N. H. (1961). Scales and statistics: Parametric and nonparametric. *Psychological Bulletin, 58,* 305–316.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. T. Spense & J. T. Spence (Eds.), *The Psychology of learning and motivation, Vol. 2.* New York: Academic Press.

Bogartz, R. S. (1976). On the meaning of statistical interactions. *Journal of Experimental Child Psychology, 22,* 178–183.

Hellyer, S. (1962). Frequency of stimulus presentation and short-term decrement in recall. *Journal of Experimental Psychology, 64,* 650.

Krantz, D. H., & Tversky, A. (1971). Conjoint-measurement analysis of composition rules in psychology. *Psychological Review, 78,* 151–169.

Krueger, W. C. F. (1929). The effects of overlearning on retention. *Journal of Experimental Psychology, 12,* 71–78.

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition, 6,* 312–319.

MacLeod, C. M. & Nelson, T. O. (in press). Response latency and response accuracy as measures of memory. *Acta Psychologica.*

Murdock, B. B. Jr., & Cook, C. D. (1960). On fitting the exponential. *Psychological Reports, 6,* 63–69.

Peterson, L. R., Rawlings, L., & Cohen, C. (1977). The internal construction of spatial patterns. In G. H. Bower (Ed.), *The Psychology of learning and motivation, Vol. 11.* New York: Academic Press.

Postman, L., & Riley, D. A. (1959). Degree of learning and interserial interference in retention. *University of California Publications in Psychology, 8,* 271–396.

Slamecka, N. J., & McElree, B. (1983). Normal forgetting of verbal lists as a function of their degree of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 384–397.

Underwood, B. J. & Keppel, G. (1963). Retention as a function of degree of learning and letter-sequence interference. *Psychological Monographs, 77* (Whole No. 567).

Woodworth, R. S., & Schlosberg, H. (1961). *Experimental Psychology (Revised Edition).* New York: Holt, Reinhart and Winston.

Youtz, A. C. (1941). An experimental evaluation of Jost's laws. *Psychological Monographs* (No. 238).

(*Appendix follows on next page*)

Appendix 1

a. The two decay equations are

$$\text{Low learning: } PL(t_L) = Le^{-t_L}$$

and

$$\text{High learning: } PH(t_H) = He^{-t_H},$$

where $H > L$. Solving for t,

$$\text{Low learning: } t_L = \ln(L) - \ln(PL(t_L)), \quad (1)$$

$$\text{High learning: } t_H = \ln(H) - \ln(PH(t_H)) \quad (2)$$

To compute the horizontal difference between the two forgetting curves, we let

$$PL(t_L) = PH(t_H).$$

Subtracting Equation 1 from Equation 2,

$$t_H - t_L = \ln(L) - \ln(H). \quad (3)$$

Note that the right side of Equation 3 is a constant; hence, $t_H - t_L$ is also constant.

b. The two decay equations are

$$\text{Low learning: } PL(t_L) = Le^{-k_L t_L}$$

and

$$\text{High learning: } PH(t_H) = He^{-k_H t_H},$$

where $H > L$ and $k_H < k_L$. Solving for t,

$$\text{Low learning: } t_L = [\ln(L) - \ln(PL(t_L))]/k_L, \quad (1)$$

$$\text{High learning: } t_H = [\ln(H) - \ln(PH(t_H))]/k_H \quad (2)$$

To compute the horizontal distance between the two learning curves, we let

$$PL(t_L) = PH(t_H) = P.$$

Subtracting Equation 1 from Equation 2,

$$t_H - t_L = [\ln(H) - \ln(L)] - [\ln(P)/k_L + \ln(P)/k_H]. \quad (3)$$

The first term in square brackets on the right-hand side of Equation 3 is a constant. As P decreases, the second term in square brackets on the right-hand side of Equation 3 decreases; hence, $t_H - t_L$ increases.

Appendix 2

a. Add to the 3 assumptions of the general model:

4a. The time to progress from any state to any subsequent state is unaffected by the degree of original learning.

Let the cognitive system be in state *Sj* at delay interval 0 following low learning. By Assumptions 1 and 3, the system must similarly be in state *Sj* at delay interval *tj* following high learning.

Consider any state, *Sk*, subsequent to *Sj*. By Assumption 2, the system must pass through the same intervening states enroute from *Sj* to *Sk* following both low and high learning. By Assumption 4a, the time to transit from each state to its successor will be the same following low and high learning. Therefore, it will take the same time to transit from state *Sj* to state *Sk* following low and high learning. Let this transition time be *ta*.

By Assumption 1, because the state of the system is the same at time *ta* following low learning, and at time $tj + ta$ following high learning, performance at time *ta* following low learning must be identical to performance at time $tj + ta$ following high learning.

This reasoning holds true for any *Sk* that is chosen, hence over all possible *ta* values. Therefore, the forgetting curves must be horizontally parallel. They must always differ by *tj*.

b. Add to the 3 assumptions of the general model:

4b. The time required to progress from any state to any subsequent state is greater following high learning than following low learning.

Let the cognitive system be in state *Sj* at delay interval 0 following low learning. By Assumptions 1 and 3, the system must similarly be in state *Sj* at delay interval *tj* following high learning.

Consider any state, *Sk*, subsequent to *Sj*. By Assumption 2, the system must pass through the same intervening states enroute from *Sj* to *Sk* following both low and high learning. By Assumption 4b, the time to transit from each state to its successor will be longer following high than following low learning. Therefore, it will take longer to transit from state *Sj* to state *Sk* following high versus low learning. Let the transition times be *ta* for low learning and $ta + tb$ for low learning.

By Assumption 1, because the state of the system is the same at time *ta* following low learning, and at time $tj + ta + tb$ following high learning, performance at time *ta* following low learning must be identical to performance at time $tj + ta + tb$ following high learning. This means that the horizontal difference corresponding to state *Sk*,

$$(tj + ta + tb) - ta = tj + tb$$

is greater than the horizontal difference corresponding to state *Sj*,

$$(tj + ta) - ta = tj.$$

Thus, the curves diverge from the performance level corresponding to state *Sj* to the performance level corresponding to state *Sk*.

Finally, consider some state, *Sm*, subsequent to *Sk*. Let the time to transit from state *Sk* to state *Sm* following low learning by *tc*. By Assumption 4b, it must take longer than *tc* to transit from state *Sk* to state *Sm* following high learning. Let this time be $tc + td$. Thus, the system will be in state *Sm* at time $ta + tc$ following low learning and at time $tj + ta + tb + tc + td$ following high learning. By Assumption 1, performance will also be the same at these times. This means that the horizontal difference corresponding to state *Sm*,

$$(tj + ta + tb + tc + td) - (ta + tc) = tj + tb + td$$

is greater than the horizontal difference corresponding to state *Sk*,

$$(tj + ta + tb) - ta = tj + tb.$$

Thus the curves diverge from the performance level corresponding to state *Sk* to the performance level corresponding to state *Sm*. Because *Sk* and *Sm* can be chosen to be any two states in the sequence (as long as *Sm* is subsequent to *Sk*), the curves must continually diverge.

Received July 20, 1984
Revision received October 5, 1984 ■

Berscheid Appointed Editor, 1986–1991

The Publications and Communications Board of the American Psychological Association announces the appointment of Ellen S. Berscheid, University of Minnesota, as Editor of *Contemporary Psychology* for a 6-year term beginning in 1986.

Publishers should note that books should *not* be sent to Berscheid. **Beginning January 1, 1985, publishers should send two copies of books to be considered for review plus notices of publication to:**

Contemporary Psychology
American Psychological Association
c/o PsycINFO
1400 North Uhle Street
Arlington, Virginia 22201

Please note that all reviews are written by invitation.
