Journal of Experimental Psychology:
Learning, Memory, and Cognition
1990, Vol. 16, No. 5, 916-926

0278-7393/90/$00.75

OBSERVATIONS

Learning–Forgetting Independence, Unidimensional Memory Models, and Feature Models: Comment on Bogartz (1990)

Geoffrey R. Loftus
University of Washington

Donald Bamber
Naval Ocean Systems Center

In his recent articles, Bogartz offered a definition of what it means for forgetting rate to be independent of degree of original learning. He showed that, given this definition, independence is confirmed by extant data. Bogartz also criticized Loftus's (1985b) proposed method for testing independence. In this commentary, we counter Bogartz's criticisms and then offer two observations. First, we show that Loftus's horizontal-parallelism test distinguishes between two interesting classes of memory models: unidimensional models wherein the memory system's state can be specified by a single number and multidimensional models wherein at least two numbers are required to specify the memory system's state. Independence by Loftus's definition is implied by a unidimensional model. Bogartz's definition, in contrast, is consistent with either model. Second, to better understand the constraints on memory mechanisms dictated by the mathematics of the models under consideration, we develop a simple but general feature model of learning and forgetting. We demonstrate what constraints must be placed on this model to make learning and forgetting rate independent by Loftus's and by Bogartz's definitions.

Is forgetting rate independent of degree of original learning? There has been an ongoing debate, not merely over the answer to this question, but more fundamentally over the method that should be used to answer the question.

Slamecka and McElree (1983; see also, Slamecka, 1985) proposed a method for answering this question. Loftus (1985a, 1985b) criticized Slamecka and McElree's method and proposed a method of his own. Now, Bogartz (1990) criticized Loftus's method and proposed a method of his own.

This article is designed to provide perspective on the entire set of issues and to defend Loftus's (1985b) method against Bogartz's criticisms.

Methods for Testing Independence

There are three steps—one optional and two mandatory—in devising a method to answer the question "Is forgetting rate independent of degree of original learning?" The first, optional step is to construct a memory mechanism embodying one's intuitive notion of learning–forgetting independence. The second step is to formulate a pattern to which empirical data must conform when there is learning–forgetting independence. (If the optional first step has been implemented, this pattern will be diagnostic of whether or not empirical data are consistent with the proposed memory mechanism.) This pattern constitutes a definition of learning–forgetting independence. The third step is to devise a *statistical test* to determine whether empirical data conform to the definition.

Loftus (1985b) went through all three of these steps. Bogartz (1990) skipped the first step and went through only the last two.

Loftus (1985a, 1985b)

Loftus formulated a memory mechanism embodying his concept of learning–forgetting independence. Briefly, the assumptions of this mechanism were that (a) learning leads to stored information in memory, (b) there is a one-to-one correspondence between stored information and memory-test performance, and (c) during forgetting, the time required for information to drop from one level to another depends only on the two levels (i.e., not on degree of original learning).

As Loftus demonstrated, this mechanism implies a definition of independence: that forgetting curves issuing from different degrees of original learning should be separated by a constant amount of time (i.e., be horizontally parallel), as illustrated in Figure 1. In Appendix A, Theorem 1, we prove that this independence definition is satisfied if and only if performance, P, degree of original learning, O, and forgetting time, T, are related by the equation[1]

$$P = m[g(O)e^{-T}], \quad (1)$$

This research was supported by a National Institute of Mental Health grant to Geoffrey R. Loftus. Keith Rayner, Sandy Pollatsek, and an anonymous reviewer provided valuable comments and suggestions in response to an earlier version of this article. We are especially indebted to Richard Bogartz, both for his very extensive review of our earlier manuscript and for stimulating our thinking as much as he has on this entire set of issues.

Correspondence concerning this article should be addressed to Geoffrey R. Loftus, Department of Psychology, University of Washington, Seattle, Washington 98195. Electronic mail may be sent to gloftus@milton.u.washington.edu.

[1] Bogartz (1990) proved this same theorem using much stronger assumptions than we have. Whereas we used the assumption of horizontal parallelism to derive Equation 1, Bogartz used that assumption together with Equation 2 to derive Equation 1.

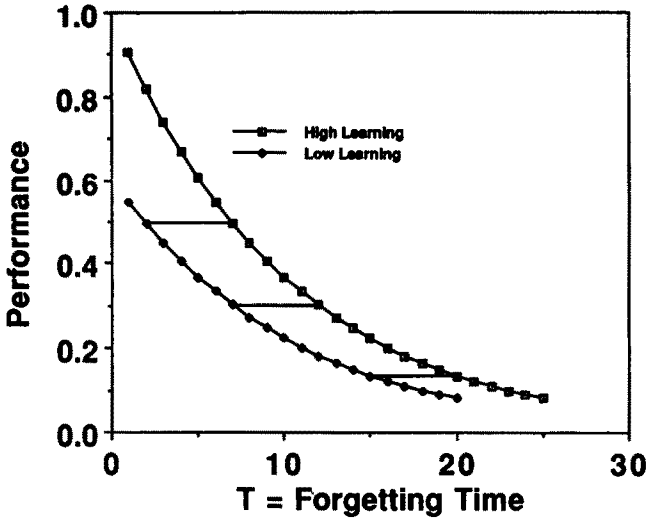

Figure 1. Example of horizontally parallel curves.

where m and g are monotonically increasing[2] functions. (Thus, for Loftus, Equation 1 is an alternative independence definition.) Loftus's proposed test of horizontal parallelism is rather primitive: It consists of interpolating smooth curves between observed points on empirical forgetting graphs and measuring the horizontal separation between different forgetting curves (ignoring considerations of error variance). Because extant forgetting curves are not horizontally parallel, Loftus (1985b) concluded that, by his definition of independence, forgetting rate depends on degree of original learning.

Bogartz (1990)

Bogartz (1990) proposed a different definition of independence strictly in terms of a mathematical equation: Degree of original learning and forgetting rate are independent if and only if P, O, and T are related by the equation

$$P = m[g(O)e^{-f(T)}], \tag{2}$$

where m and g are again monotonically increasing and f is also monotonically increasing. Bogartz's test to determine whether Equation 2 is consonant with existing data is rather sophisticated; it uses the machinery of functional measurement to determine (a) if Equation 2 can be satisfied and (b) given that it can, what the form of f is. Because Bogartz could not reject the null hypothesis embodied in Equation 2, he concluded that, by his definition of independence, forgetting rate is independent of degree of original learning.

In summary, Bogartz (1990) made two valuable contributions. First, he provided a specific equation that fits existing data. Second, he provided the statistical machinery required for evaluating his equation's fit using traditional statistical procedures. We wish to emphasize that this machinery allows, ipso facto, a statistical test of Loftus's horizontal-parallelism definition, which boils down to a test of whether $f(T)$ in Equation 2 is linear (see Appendix A, Theorem 2).

Goals of This Article

We have three goals in the remainder of this article. Our first goal is to answer Bogartz's (1990) criticisms of Loftus's (1985a, 1985b) independence definition. Our second goal is to characterize Loftus's independence definition in a new form and to contrast it with Bogartz's definition. In particular, we show that by Loftus's independence definition, independence is implied by a special case of what we term a unidimensional memory model, whereas Bogartz's independence definition is consistent with what we term a *multidimensional memory model.* Our third goal is to contrast the difference between Loftus's and Bogartz's definitions of independence by examining what kinds of forgetting mechanisms have the property of forgetting–learning independence as defined by Loftus and by Bogartz. Bogartz's independence definition is expressed in terms of mathematical descriptions of forgetting, whereas Loftus's independence definition is expressed in terms of mechanisms of forgetting. (The difference between mathematical descriptions and mechanisms is analogous to the difference between Kepler's descriptive planetary orbit laws and Newton's mechanistic gravitational law.) To provide a common frame of reference within which Bogartz's and Loftus's definitions may be compared, we develop feature-based memory mechanisms that conform to both of their independence definitions, respectively.

Bogartz's Criticism of the Horizontal-Parallelism Method

Bogartz (1990) provided four criticisms of Loftus's (1985a, 1985b) horizontal-parallelism method for testing learning–forgetting independence. We summarize these criticisms and provide responses. We believe that the first three criticisms bespeak a possible misapprehension of Loftus's definition (Equation 1). If Equation 1 were incorrectly regarded as resulting from (a) adopting Bogartz's independence definition (Equation 2) and (b) making assumptions that $f(T)$ is linear, then Bogartz's first three criticisms would be valid. However, this is not the genesis of Equation 1. In short, Loftus's definition of independence does not consist of Bogartz's definition plus assumptions; Loftus's definition stands on its own and is different from Bogartz's. The fourth criticism, based on statistical grounds, is entirely on target; however, Bogartz himself provided a solution to it.

Bogartz's (1990) criticisms follow:

1. The horizontal-parallelism test is invalid as a means of testing learning–forgetting independence. For example Bogartz (1990) described Wickelgren's (1974) data showing nonparallel forgetting curves and remarked that

> Wickelgren's (1974) theory and data . . . pose a significant problem for Loftus's approach. According to Loftus's analysis, Wickelgren's results indicate dependence of forgetting rate on initial level of learning . . . . But the underlying assumptions of Wickelgren's model entail independence. (p. 145)

---

[2] For ease of discourse, we use the term *monotonic* to mean strictly monotonic.

What Bogartz seemed to say here is that the horizontal-parallelism test is inappropriate as an evaluation of independence by Bogartz's definition. This is quite correct. However, Loftus's independence definition is different from Bogartz's definition (cf. Equations 1 and 2). The horizontal-parallelism test is appropriate for evaluating independence by Loftus's definition.

2. Confirmation of horizontal parallelism implies exponential forgetting. For example, Bogartz (1990) remarked,

> It is shown in [Bogartz's] Appendix B that under rather general conditions (additive or multiplicative decomposition of the function describing amount of information retained), horizontal additivity implies an exponential forgetting function .... Thus although Loftus starts with the notion of independence of rate of forgetting from amount of original learning, he winds up essentially equating the idea of independence with the existence of an exponential forgetting function because his definition does not admit independence in the absence of exponential forgetting. (p. 143)

Bogartz is certainly correct, but it is difficult to understand why he viewed this property of the horizontal-parallelism method as a liability. Any set of data that fails the horizontal-parallelism test allows one to rule out a large and particularly interesting class[3] of forgetting models. A method's ability to rule out a large class of models is customarily viewed as an asset.

3. Loftus assumes a one-to-one correspondence between memory information and performance. For example, Bogartz remarked,

> But Loftus's approach amounts to preserving a simple decay function at the cost of abandoning additivity and independence.[4] I believe this occurred because he adopted too strong an assumption: the assumption that each level of performance uniquely identifies one and only one cognitive state .... (p. 143)

Both Bogartz and Loftus proposed models in which there is a one-to-one correspondence between retained information, r, and performance. For example, Bogartz stated that "m is a monotone increasing function relating observed performance, R, to amount of information retained, r" (p. 139) and characterized this assumption (among others) as being "abstract ... and noncontroversial." In essence, then, Bogartz and Loftus adopted the same one-to-one assumption.

As we see, there is a more subtle interpretation of the term *one-to-one correspondence*, which is intimately related to the difference between Loftus's and Bogartz's independence definitions. Independence by Loftus's definition implies a one-to-one correspondence between retained information at any time T_1 and performance at any later time, $T_2 > T_1$. Independence by Bogartz's definition, in contrast, carries no such implication.

4. The statistical test of horizontal parallelism is inadequate. For example, Bogartz (1990) remarked,

> It is not clear ... how Loftus would statistically test for the presence or absence of horizontal interaction. In the data that he has cited, the error variance estimates will all be based on vertical variability within a cell .... Performance is measured and varies at fixed delay times; delay times are ordinarily not measured at fixed performances. There will ordinarily be no estimate of horizontal error variance against which to compare horizontal interaction effects. (p. 145)

This criticism was certainly valid when Loftus (1985a, 1985b) made his original arguments. As we have noted, Bogartz (1990) made the very valuable contribution of providing a powerful horizontal-parallelism test: As we show by proving Theorem 2 in Appendix A, it amounts to testing the null hypothesis that $f(T)$ in Equation 2 is linear in T.

A New Characterization of Loftus's Definition

Our goal in this section is to recast Loftus's forgetting definition in a new light. In particular, we show how memory models in general can be divided into unidimensional models on the one hand and multidimensional models on the other. In a unidimensional model, the state of the memory system can be characterized by a single memory dimension, whereas in a multidimensional model more than one dimension is needed. We show that learning–forgetting dependence by Loftus's definition (i.e., failure of the horizontal-parallelism test) disconfirms all unidimensional memory models. We end this section with a comment on Bogartz's notion of "psychological time" and relate this notion to the multidimensional character of Bogartz's independence definition.

Unidimensional Versus Multidimensional Memory Models

There exists a class of models whose fundamental tenet is that the state of the memory system is describable as a point in a unidimensional space. Learning and forgetting may both be described as movements of that point through the unidimensional space. In other words, the memory system's state can be completely specified by a single spatial coordinate (say r, the amount of stored information).

To understand the nature of such unidimensional models, consider the system at some time, T_0, following original learning, O. Stored information at that time is $r(O, T_0)$. Now consider the system at a later time ($T_0 + \Delta T$), when some forgetting has occurred. The critical property of unidimensional models is that stored information at time ($T_0 + \Delta T$) depends only on the memory system's state at time T_0 and on the length of the interval ΔT. Now, the memory system's state at time T_0 is completely specified by the value of r at time T_0. Hence,

$$r(O, T_0 + \Delta T) = u[r(O, T_0), \Delta T],$$

where u is monotonically increasing in the first argument and monotonically decreasing in the second argument.

In contrast, in a multidimensional model, the state of the memory system is describable as a point in a multidimensional space. Consequently, the state of the memory system

---

[3] For the present discussion, by "class of models" we refer to any model that posits exponential loss of retained information, plus a monotonic function, m, mapping retained information into performance. As we see shortly, we actually refer more generally to the class of all unidimensional memory models.

[4] We must emphasize that this is independence by Bogartz's definition, not by Loftus's.

cannot be completely specified by a single spatial coordinate. The process of learning and subsequent forgetting may be described by a trajectory through multidimensional memory space. The essential difference between unidimensional memory models and multidimensional models is that a wider variety of learning–forgetting trajectories are possible in the multidimensional space. In a unidimensional model, learning and subsequent forgetting is like climbing up and then down a ladder: The path taken down is identical to the path taken up. However, in a multidimensional model, learning and subsequent forgetting is like climbing up and then down a mountain: The path down may differ from the path up. Thus, by a multidimensional model, two people studying the same list for differing numbers of learning trials are like two people taking the same path up a mountain but one person going further than the other. Once learning stops and forgetting begins, the two people may take different paths down the mountain.

A variety of investigators suggested multidimensional models in which the state of the system is completely described by some measure of retained information and some function of forgetting time, T (e.g., Wickelgren, 1972, 1974; Yost, cited in Woodworth & Schlosberg, 1961; Youtz, 1941). More generally, in a multidimensional model, the memory system's state cannot be completely specified by the amount, r, of retained information. At least one more memory-space coordinate is needed. Suppose that such a coordinate is denoted by s. We will not attempt to specify the nature of s. In different multidimensional models, it might be various things (e.g., in the models just mentioned, s is a function of amount of forgetting time). Now the amount, r, of information retained at time $(T_0 + \Delta T)$ is a function of the state of the system at time T_0 and of the length of the interval ΔT. Because the state of the system at time T_0 is completely specified by the values of the coordinates r and s at time T_0, we may write,

$$r(O, T_0 + \Delta T) = u[r(O, T_0), s(O, T_0), \Delta T]. \quad (3)$$

We assert that models characterized by Equation 3 exhibit a dependence of forgetting on degree of original learning. To see why this is, suppose that two combinations of O and T—(O_1, T_1) and (O_2, T_2)—produce equal performance (and thus r) values. We refer to the r value as r_0; thus,

$$r(O_1, T_1) = r(O_2, T_2) = r_0.$$

To characterize the state of the system after some interval, ΔT, it is necessary to specify the value of s. The equations are

$$r(O_1, T_1 + \Delta T) = u[r_0, s(O_1, T_1), \Delta T]$$

and

$$r(O_2, T_2 + \Delta T) = u[r_0, s(O_2, T_2), \Delta T],$$

where $s(O_1, T_1)$ and $s(O_2, T_2)$ can be, and usually are, different from one another.

Forgetting rate may now be identified with the decrease in r as one progresses over the time interval ΔT. The size of this decrease will be different for different values of s. Because the value of s depends on whether the forgetter was in learning condition O_1 or O_2, we conclude that such multidimensional models exhibit dependence between original learning and forgetting.

Bogartz's and Loftus's Definitions

As we show in Appendix B, a unidimensional memory model implies horizontal parallelism of high- and low-learning forgetting curves (i.e., implies Loftus's independence definition). Accordingly, if forgetting curves are not horizontally parallel, the entire class of unidimensional models may be ruled out. In the next section, we give examples from the literature of how this feature of the horizontal-parallelism method could be applied.

Bogartz (1990) amply demonstrated that nonparallel forgetting curves are consistent with his independence definition, provided his $f(T)$ is nonlinear. Again by the logic provided in Appendix B, independence by Bogartz's definition is consistent with a multidimensional memory model.

Two Examples

Forgetting models can be classified as unidimensional or multidimensional. Here we briefly consider two forgetting models: those of Atkinson and Shiffrin (1968) and Wickelgren (1972). In each case, the investigators evaluated their models by positing a specific relation between an internal construct (e.g., amount of information or trace strength) and the dependent variable (e.g., recall probability or d') used in a memory test, and then determining the model's fit to memory data. The point we wish to make here is that this evaluation strategy requires an unnecessarily strong model. One must formulate both a model of how the contents of memory vary with study and forgetting time and a model of how performance on memory tests depends on the contents of memory. Essentially, one must formulate a performance submodel even if one is principally interested only in acquisition and loss of information. Use of the horizontal-parallelism test makes it unnecessary to formulate a performance model to test the information-acquisition/loss model.

Atkinson-Shiffrin model. Atkinson and Shiffrin (1968) explicitly incorporated both short- and long-term memory processes in their model (see also Waugh & Norman, 1965). Accordingly, to specify the system's state requires two elements: the probability that the necessary information is present in short-term memory, and the amount of task-relevant information present in long-term memory. This makes the model multidimensional.

However, each memory store is individually unidimensional; indeed, the Atkinson-Shiffrin forgetting equations are explicitly exponential. This means that if one can arrange experimental situations in which use of only one kind of memory or the other can be reasonably assumed, then each of the model's components can be tested individually by the horizontal-parallelism test.

There are many paradigms in the literature in which memory performance can be assumed to be based only on long-term memory; this, for example, was true of all paradigms considered by Slamecka and McElree (1983), Loftus (1985a), and Bogartz (1990) except the Hellyer paradigm. Demonstrations that the low- and high-learning–forgetting curves are not horizontally parallel is therefore sufficient to disconfirm Atkinson and Shiffrin's long-term forgetting assumption.

It is more difficult (in fact, probably impossible) to arrange situations in which recall is based on short-term memory only (the best one can do is to try to correct short-term forgetting data via some model, e.g., in which retrieval occurs independently from short- and long-term memory). In this context, it is worthwhile to make an observation about the Hellyer (1962) data that were reanalyzed by both Loftus (1985b) and Bogartz (1990). These data are uncorrected;[5] hence, recall is undoubtedly based on both short- and long-term memory. It is therefore not surprising that they disconfirm a unidimensional model. The question of whether short-term forgetting can be described by a unidimensional model is still open.

Wickelgren model. Wickelgren (1972, 1974) presented a forgetting theory in which, essentially, forgetting rate depends on both amount of retained information and on time since learning. This is an example of a forgetting model that, in contrast to the Atkinson-Shiffrin model, is multidimensional in long-term forgetting. However, to test his theory (and alternative theories to explain his data), Wickelgren made the strong assumption that memory performance (measured as d') is a linear function of trace strength. Using this strategy, Wickelgren confirmed his model and disconfirmed others, including exponential decay. The disconfirmation of the exponential-decay model, however, could have been accomplished without the strong hypothesis linking strength and d'; all that was required was that the horizontal-parallelism test fail.

In short, Loftus's horizontal-parallelism test provides valuable information about the memory system's structure that goes beyond resolution of an issue (learning–forgetting independence) whose definition is subject to disagreement. Although we do not demand that memory investigators accept Loftus's independence definition, we nonetheless believe that it would be a suboptimal scientific strategy to refuse to test for horizontal parallelism as Bogartz appeared to suggest.

Nonlinear Transformations and "Psychological Time"

Bogartz posited a transformation of physical forgetting time, $f(T)$, which he called "psychological time." In this section, we show that (a) Bogartz's psychological time is different from the common view of psychological time, (b) inclusion of psychological time within Bogartz's model violates certain common-sense ideas, and (c) this violation occurs because Bogartz insisted that his multidimensional memory model incorporates the assumption that forgetting rate is independent of original learning.

We have no quarrel with the proposition that psychological time is a useful construct in some instances. Intuitively, time seems to run faster in some situations and slower in others. This intuition has been confirmed in numerous laboratory experiments (e.g., Ornstein, 1969).

However, this is not the meaning of psychological time as Bogartz incorporated it in his model. Instead, Bogartz's use of the construct allows some psychological process to run at different physical rates depending on when in physical time it begins. According to Bogartz's model, for instance, forgetting of some stored information that begins, say, on Wednesday could progress at a faster rate than forgetting of the same stored information when the forgetting begins on Tuesday instead.

Example

To see why this is, let us consider hypothetical data generated by Bogartz's model. Such data, shown in Figure 2 (top panel), issue from the equations

$$\text{High learning: } P = \sqrt{10.00e^{-T^{0.5}}}$$

and

$$\text{Low learning: } P = \sqrt{3.68e^{-T^{0.5}}},$$

where T is in days.

Suppose that two people, Heather and Lois, participate in the two different forgetting conditions. Heather is in the high-learning condition, and Lois is in the low-learning condition. Suppose further that Heather and Lois are identical clones of one another, at least with respect to learning and forgetting processes.

At exactly noon on Tuesday, Heather begins forgetting her high-learned list. At exactly noon on Wednesday, Lois begins forgetting her low-learned list. Note that at noon on Wednesday, Lois and Heather are identical with respect to performance and thus with respect to retained information; for both, $P = 1.92$, and $r = m^{-1}$ (1.92), where m^{-1} is the inverse of the monotonic function mapping retained information into performance. This situation is illustrated in Figure 2 (bottom panel), wherein the forgetting curves in the top panel of Figure 2 are depicted in physical time. Within the context of Bogartz's independence definition, an oddity becomes apparent at noon Wednesday: Although Heather and Lois are identical people with identical amounts of retained information who, according to Bogartz, are forgetting at identical rates, they will still have different performances (and thus different amounts of retained information) at any later point. At noon on Thursday, for example, Heather will have a performance of 1.56, whereas Lois will have a performance of 1.16.

Bogartz's Model Is a Multidimensional Model

The reason for this apparently strange prediction is that Bogartz's model is a multidimensional model that nonetheless insists on assuming independence of original learning and forgetting rate. In Bogartz's model, two memory coordinates are required to completely specify the state of the memory system at time T. They are the amount, $r(O, T)$, of information retained and the elapsed psychological time $f(T)$. Thus, as proved in Appendix C,

$$r(O, T + \Delta T) = u[r(O, T), f(T), \Delta T]. \quad (4)$$

[5] The lack of very long retention intervals in Hellyer's experiment precludes correcting them for long-term memory effects.

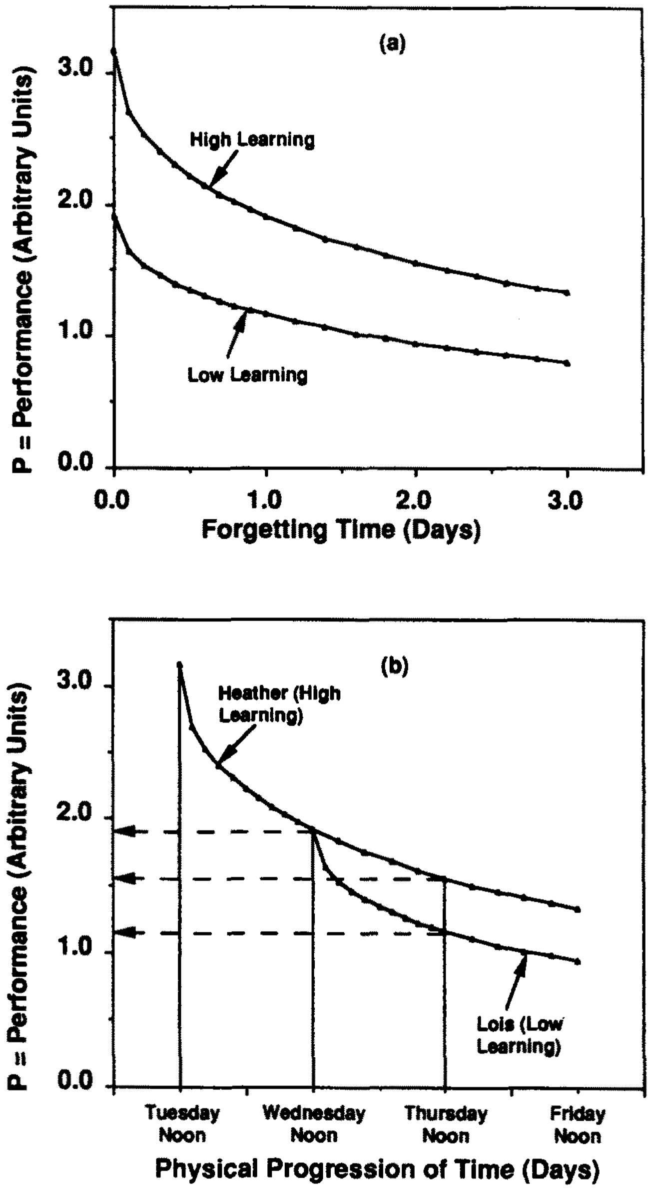

Figure 2. The top panel shows high- and low-learning forgetting curves derived from Bogartz's model. In the bottom panel, the same forgetting curves are shown adjusted so that they reflect forgetting in physical time.

Equation 4 demonstrates that, by the multidimensional nature of the model, retention at any given time ($T_0 + \Delta T$) is dependent not only on the amount of stored information, r_0, at time T_0, but also on T_0, how long forgetting has been progressing. As of Wednesday noon, Heather has been forgetting for a day, whereas Lois has only begun to forget. It is this difference between Heather and Lois that causes the subsequent performance differences.

Thus, beginning at noon Wednesday, identical information stores are eroded at rates that depend on when the list was originally learned. Given this state of affairs, it makes little sense to claim that forgetting rates are the same; such a claim contradicts a basic, common-sense notion of what it means for two processes to progress at the same rate.

Feature Models

Bogartz defined independence of forgetting rate and learning strictly in terms of mathematical properties and not in terms of any learning–forgetting mechanisms that would embody those properties. To better understand what constraints Bogartz's definition places on possible memory mechanisms, we now develop a simple feature model of learning and forgetting. We describe three versions of this model: (a) the general feature model, (b) a special case of the model that we call the Bogartz feature model, and (c) a special case of the Bogartz feature model that we call the Loftus feature model. We caution that it is not our purpose here to argue for the validity of this general feature model. Rather we use it as a forum to illustrate constraints on memory mechanisms imposed by formal mathematical models.

General Feature Model

The general model consists of the following assumptions. Each item in a list of to-be-remembered items consists of N features. Each feature is learned (and forgotten) in an all-or-none manner and independently of every other feature. After amount of study O, the probability that a randomly selected feature will be learned is $g(O)$. (Of course, the proportion of features actually learned after amount of study O is subject to chance variability. It is assumed, however, that the number of features per item is sufficiently large that any chance deviation of the proportion of learned features from $g(O)$ is small enough to be negligible.)

When a feature is learned, it is placed in a feature storage unit; each learned feature goes to a different unit. When a feature is forgotten, it is lost from the unit. Forgetting, like radioactive decay, is exponential. Thus, if a feature is learned at time zero, the probability that it will be still be retained at time T is $e^{-T/t}$, where t is the decay time for the storage unit containing the feature. (Decay time is analogous to "half-life." It denotes the time needed for probability of retention to fall to $1/e$.)

The storage units vary in quality. Higher quality units have longer decay times and hence tend to retain their features longer. (It should not be surprising that memory storage units could have varying quality. It is harder, for example, to manufacture resistors that have precisely the same resistance than resistors that deviate randomly from their nominal resistance.) Because the general feature model's assumptions are intended to be nonrestrictive, we assume that a storage unit's decay time is a function not only of the unit's quality, q, but also of the amount of study O. Thus, a unit's decay time may be expressed: $t(q, O)$. If $O_2 > O_1$, then $t(q, O_2) > t(q, O_1)$.

Let P_q denote the proportion of all storage units having quality q. It is assumed that when a feature is learned it is assigned a storage unit entirely at random. Let $r(O, T)$ denote the proportion of features retained after study O and forgetting time T. Then, assuming that the number of features is large enough to make sampling variability negligible,

$$r(O, T) = g(O)\Sigma P_q e^{-T/t(q,O)} \quad (5)$$

where the sum is over all q.

The probability that an item will be remembered at time T after study O is a function of the proportion of its features that are still retained at that time. Specifically, the probability, $P(O, T)$, of its being remembered is

$$P(O, T) = m[r(O, T)], \quad (6)$$

where $r(O, T)$ is given by Equation 5.

Bogartz Feature Model

The special characteristic of this model is that a storage unit's decay time is assumed to be unaffected by the amount O of study. Thus,

$$t(q, O) = t(q). \quad (7)$$

Using Equation 7 to rewrite Equation 5 yields

$$r(O, T) = g(O)\Sigma P_q e^{-T/t(q)}. \quad (8)$$

Combining Equations 6 and 8 yields

$$P(O, T) = m[g(O)\Sigma P_q e^{-T/t(q)}]. \quad (9)$$

Equation 9 may be simplified by defining the function

$$f(T) = -\ln[\Sigma P_q e^{-T/t(q)}].$$

Then,

$$P(O, T) = m[g(O)e^{-f(T)}]. \quad (10)$$

Equation 10 is Bogartz's independence definition. For that reason, we have named this specialized feature model the Bogartz feature model. We wish to emphasize, however, that this is not Bogartz's original model; it would be unfair to Bogartz to claim that it was. Bogartz's model is stated in abstract terms without any reference to features. The Bogartz feature model is our attempt to construct a memory mechanism that conforms to the mathematical specification of Bogartz's definition. Although it is not Bogartz's model, we have found it instructive to see what constraints must be placed on a feature model to make it conform to Bogartz's equations.

Note that the $f(T)$ in Equation 10, which corresponds to Bogartz's psychological time, is not a fundamental construct in our feature model. Rather it is a mathematical convenience introduced to simplify Equation 9 into the form of Equation 10.

Loftus Feature Model

As we have described earlier, Loftus's independence definition is stronger than Bogartz's; correspondingly, the Loftus feature model is a special case of the Bogartz feature model. In the Loftus feature model, there is no variation in storage-unit quality. An equivalent way of saying this is that any quality variation that does exist is irrelevant. Specifically, quality has no effect on decay time. Thus,

$$t(q) = t$$

and

$$P(O, T) = m[g(O)e^{-T/t}].$$

Then it follows from Appendix A, Theorem 2, that

$$P(O, T) = M[G(O)e^{-T}], \quad (11)$$

where M and G are monotonic. Equation 11 has the form of Loftus's independence definition (Equation 1).

Comparing Definitions

The Bogartz and Loftus definitions of independence of forgetting rate from learning are now compared.

Bogartz's Definition

According to Bogartz (1990), forgetting is independent from original learning if two functions, h_1 (of O, original learning) and h_2 (of T, forgetting time) can combine to express r, retained information, as follows:

$$r(O, T) = h_1(O)h_2(T).$$

The Bogartz feature model (Equation 8) conforms to this condition where

$$h_1(O) = g(O)$$

and

$$h_2(T) = \Sigma P_q e^{-T/t(q)}.$$

Conversely, according to Bogartz, forgetting is dependent on original learning if

$$r(O, T) = h_1(O)h_2(O, T).$$

The general feature model (Equation 5) conforms to this condition where

$$h_2(O, T) = \Sigma P_q e^{-T/t(q,O)},$$

except in the special case where Equation 7 holds; but that special case is precisely the Bogartz feature model.

Consider the conditional probability that a randomly selected feature will be retained at time T given that it was learned. In the general feature model, this conditional probability is obtained by dividing Equation 5, the joint probability that a feature is learned and retained, by $g(O)$, the probability that the feature is learned. Thus,

$$P(\text{feature retained at } T \mid \text{feature learned}) = \Sigma P_q e^{-T/t(q,O)} = h_2(O, T).$$

In the Bogartz feature model, however, this conditional probability is obtained by dividing Equation 8 by $g(O)$, or

$$P(\text{feature retained at } T \mid \text{feature learned}) = \Sigma P_q e^{-T/t(q)} = h_2(T).$$

Thus, within the context of the general feature model, Bogartz's definition of independence of forgetting from learning is satisfied if and only if the conditional probability of a feature being retained given that it was learned does not vary with O.

Loftus's Definition

Loftus's definition of original learning–forgetting rate independence is satisfied if and only if

$$r(O_1, T_1) = r(O_2, T_2) \qquad (12)$$

implies

$$r(O_1, T_1 + \Delta T) = r(O_2, T_2 + \Delta T).$$

In other words, Loftus's definition is satisfied if, in order to predict the proportion of features retained at time $(T + \Delta T)$, the only information needed is the proportion of features retained at time T. Specifically, if one knows $r(O, T)$, one does not need to know O to predict $r(O, T + \Delta T)$. Thus, the definition is satisfied if (a) the memory system is unidimensional and (b) r may be used as the coordinate of the unidimensional memory space.

The Loftus feature model satisfies Loftus's learning–forgetting independence definition. However, the Bogartz feature model does not satisfy Loftus's definition. To see intuitively why this is so, consider that, in the Bogartz feature model, the longer forgetting has continued, the higher the average quality of the storage units of the surviving features. Features stored in low-quality units having short decay times are disproportionately forgotten as time passes. Consequently, the surviving features tend to be concentrated in the higher quality units.

Why Lois Forgets Faster Than Heather

Suppose, to illustrate, that Lois had study O_1 and forgetting time T_1, whereas Heather has had study O_2 and forgetting time T_2. Suppose Equation 12 holds with $O_2 > O_1$ and $T_2 > T_1$. In other words, Lois and Heather retain the same proportion of features, but Heather has been forgetting longer. Because $T_2 > T_1$, the features retained by Heather tend to be in higher quality storage units than those retained by Lois. Consequently, after an additional time ΔT passes, Heather will retain a higher proportion of features than will Lois. In other words, knowing the proportion of features retained by a subject at a given time does not provide enough information to predict the proportion of features retained after an additional interval ΔT passes.

Bogartz's and Loftus's Questions

Within the context of a feature model, when Bogartz asks whether forgetting is independent of learning, he is essentially asking whether there is variation under O of the conditional probability of a feature being retained at time T given that it was learned. When Loftus asks whether forgetting is independent of learning, he is asking whether retention at time T is a perfect predictor of retention at time $T + \Delta T$.

Both these questions are reasonable to investigate empirically. We would not support any contention that one of these questions is legitimate to empirically investigate and the other one is not. However, for reasons detailed elsewhere in this article, it is our contention that it is Loftus's question that is best paraphrased "Is forgetting independent of original learning?"

Conclusions

Bogartz's characterization of different forgetting models provides an enormously useful foundation for explicitly comparing the relations among and predictions of such models. However, we are troubled that Bogartz (1990) dismissed the horizontal-parallelism tests associated with Loftus's independence definition, stating, "I reject the use of the horizontal interaction as a means of studying the dependence of forgetting rates on amount of learning" (p. 143). As we have tried to emphasize, the horizontal-parallelism test, although inappropriate for assessing independence by Bogartz's definition, is perfectly appropriate for testing independence by Loftus's definition.

We have seen that there is an intimate relationship between Loftus's independence definition on the one hand and unidimensional memory models on the other. All such models can be rejected by a set of data that fails the horizontal-parallelism test. We illustrated this ability using two models from the literature.

We developed the feature models to better understand the constraints placed on memory mechanisms by Bogartz's and Loftus's independence definitions. As we have shown, these two definitions place quite different constraints on memory mechanisms.

References

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

Bogartz, R. S. (1990). Evaluating forgetting curves psychologically. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 138–148.

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

Levine, M. V. (1970). Transformations that render curves parallel. *Journal of Mathematical Psychology, 7,* 410–443.

Loftus, G. R. (1985a). Consistency and confoundings: Reply to Slamecka. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 817–820.

Loftus, G. R. (1985b). Evaluating forgetting curves. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 396–405.

Ornstein, R. (1969). *On the experience of time.* New York: Penguin.

Slamecka, N. J. (1985). On comparing rates of forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 812–816.

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.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72,* 89–104.
Wickelgren, W. A. (1972). Trace resistance and the decay of long-term memory. *Journal of Mathematical Psychology, 9,* 418–455.
Wickelgren, W. A. (1974). Single trace fragility theory of memory dynamics. *Memory and Cognition, 2,* 775–780.
Woodworth, R. S., & Schlosberg, H. (1961). *Experimental psychology* (rev. ed.). New York: Holt, Reinhart and Winston.
Youtz, A. C. (1941). An experimental evaluation of Yost's laws. *Psychological Monographs, 53*(1, Whole No. 238).

Appendix A

Proofs of Two Theorems

The purpose of this appendix is to prove two theorems that are relevant to our horizontal-parallelism arguments.

Notation

There are three observable variables: They are O, amount of study (e.g., time or number of trials); T, forgetting time (e.g., days); and P, performance (e.g., recall probability). P is a function of O and T, that is, $P = R(O, T)$.

Assumptions

We make four assumptions that, except for Assumption 3, are sufficiently weak to be nondebatable. Assumption 3 is at least reasonable. The assumptions are as follows:

1. Continuity: The function $R(O, T)$ is continuous in both O and T.
2. Study never hurts: If $O_2 > O_1$, then $R(O_2, T) \geq R(O_1, T)$ for all $T \geq 0$.
3. Forgetting always progresses: If $T_2 > T_1$, then $R(O, T_2) < R(O, T_1)$ for all $O > 0$.
4. What has not been learned cannot be forgotten: Take amount of study $O = 0$. If $T > 0$, then $R(0, T) = R(0, 0)$.

Definition

The horizontal-parallelism condition is defined as follows. If $O_2 > O_1$, then there exists a $\Delta T \geq 0$ such that

$$R(O_1, T) = R(O_2, T + \Delta T), \tag{A1}$$

for all $T > 0$.

Theorem 1: Text Equation 1

Theorem

Given Assumptions 1 to 4, the horizontal-parallelism condition holds if and only if there exists a monotonic increasing function m and a positive-valued monotonic nondecreasing function, g, such that,

$$R(O, T) = m[g(O)e^{-T}], \tag{A2}$$

for all $O > 0$ and all $T \geq 0$.

Proof of Theorem

Because it is easy to show that Equation A2 implies horizontal parallelism, that part of the proof is left to the reader. We now show that horizontal parallelism implies Equation A2.

Suppose that horizontal parallelism holds. This condition implies that the time taken to fall from one performance level to another is the same for all forgetting curves that pass through those two levels. So, for all $P_1 \geq P_2$, define $t(P_1, P_2)$ to be the amount of time required for performance to fall from level P_1 to level P_2. For $P_1 < P_2$, define

$$t(P_1, P_2) = -t(P_2, P_1).$$

Clearly,

$$t(P_1, P_3) = t(P_1, P_2) + t(P_2, P_3). \tag{A3}$$

Choose an arbitrary value of P and denote it P_0. In Equation A3, make the following substitutions: replace P_1 with $R(O, 0)$, P_2 with P_0, and P_3 with $R(O, T)$. This yields

$$t[R(O, 0), R(O, T)] = t[R(O, 0), P_0] + t[P_0, R(O, T)]. \tag{A4}$$

By definition, the left-hand side of Equation A4 equals T. Now define the new function,

$$t_0(P) = t(P, P_0) = -t(P_0, P).$$

Then Equation A4 may be rewritten,

$$T = t_0[R(O, 0)] - t_0[R(O, T)]. \tag{A5}$$

Define the functions,

$$m(x) = t_0^{-1}(\ln x) \tag{A6}$$

and

$$g(O) = \exp\{t_0[R(O, 0)]\} \text{ for all } O > 0. \tag{A7}$$

We set x in Equation A6 to

$$x = g(O)e^{-T}$$

or, from the definition of $g(O)$ in Equation A7,

$$\ln x = t_0[R(O, 0)] - T.$$

It follows from Equation A5 that,

$$\ln x = t_0[R(O, T)].$$

Thus, inserting x into Equation A6 yields

$$m[g(O)e^{-T}] = t_0^{-1}\{t_0[R(O, T)]\},$$

which, in turn is equal to $R(O, T)$. This completes our proof that horizontal parallelism implies Text Equation 1.

Theorem 2: Implications of Bogartz's $f(T)$ Being Linear or Nonlinear

Notation

The functions m and M are monotonically increasing. The functions g and G are positively valued and monotonically increasing. The function f is differentiable and monotonically increasing.

Theorem

Given Assumptions 1 to 4 listed in Theorem 1, if $R(O, T)$ can be represented in the form

$$R(O, T) = m[g(O)e^{-f(T)}], \tag{A8}$$

and $f(T)$ is linear, then it can be represented in the form

$$R(O, T) = M[G(O)e^{-T}]. \tag{A9}$$

Conversely, if $R(O, T)$ can be represented in the form of Equation A8 where f is nonlinear, then it cannot be represented in the form of Equation A9.

Proof of Theorem

Suppose Equation A8 holds with $f(T) = aT + b$ where $a > 0$. Define the functions

$$M(x) = m(x^a)$$

and

$$G(O) = [g(O)e^{-b}]^{1/a}.$$

Then,

$$\begin{aligned} M[G(O)e^{-T}] &= m[g(O)e^{-(b+aT)}] \\ &= m[g(O)e^{-f(T)}] \\ &= R(O, T). \end{aligned}$$

This proves the first half of the theorem. To prove the second half of the theorem, it is shown that, if Equations A8 and A9 both hold, then $f(T)$ must be linear. Suppose Equations A8 and A9 both hold. Following Bogartz, the variable T is referred to as physical time, whereas $f(T)$ is referred to as psychological time and is denoted T_*. Thus, Equation A8 may be rewritten

$$R(O, T) = m[g(O)\exp(-T_*)]. \tag{A10}$$

It follows from Equation A9 and Theorem 1 that forgetting curves will be horizontally parallel when plotted with physical time on the horizontal axis. Similarly, it follows from Equation A10 and Theorem 1 that forgetting curves will be horizontally parallel when plotted with psychological time T_* on the horizontal axis. Consider the forgetting curves for O and $(O + \Delta O)$. These two curves will be separated by a constant ΔT when plotted using physical time. Likewise, they will be separated by a constant ΔT_* when plotted using psychological time. Now

$$\Delta T_* = f(T + \Delta T) - f(T). \tag{A11}$$

Because the separation ΔT_* between the pair of forgetting curves is the same everywhere, it does not vary with T. Let ΔO approach zero. As it does so, ΔT and ΔT_* also approach zero, and it follows from Equation A11 that their ratio $\Delta T_*/\Delta T$ approaches the derivative $df(T)/dT$. Because ΔT_* does not vary with T, neither does the derivative $df(T)/dT$. From this, it follows that $f(T)$ is linear.

For additional results related to Theorem 2, the reader is referred to Levine (1970).

Appendix B

Proof That a Unidimensional Model Implies Confirmation of the Horizontal-Parallelism Test

Suppose that a memory system is characterized by a single memory-space coordinate. In that case,

$$r(O, T_0 + \Delta T) = u[r(O, T_0), \Delta T]. \tag{B1}$$

Now consider two combinations of original learning and forgetting time—(O_1, T_1) and (O_2, T_2)—that produce the same stored information, r_0, that is,

$$r(O_1, T_1) = r(O_2, T_2) = r_0. \tag{B2}$$

By Equation B1, at time ΔT later,

$$r(O_1, T_1 + \Delta T) = u[r(O_1, T_1), \Delta T] = u(r_0, \Delta T) \tag{B3}$$

and

$$r(O_2, T_2 + \Delta T) = u[r(O_2, T_2), \Delta T] = u(r_0, \Delta T). \tag{B4}$$

Because the right sides of Equations B3 and B4 are equal, the left sides are equal as well, that is,

$$r(O_1, T_1 + \Delta T) = r(O_2, T_2 + \Delta T). \tag{B5}$$

Thus, Equation B2 implies Equation B5. Applying the function m to the quantities equated in Equations B2 and B5, and recalling that $m[r(x)] = P(x)$,

$$P(O_1, T_1) = P(O_2, T_2)$$

implies that

$$P(O_1, T_1 + \Delta T) = P(O_2, T_2 + \Delta T).$$

That is, equal performance at times T_1 and T_2 implies equal performance when ΔT is added to both times. This completes our proof that a unidimensional memory model implies horizontally parallel forgetting curves.

(Appendix C follows on next page)

Appendix C

Proof of Text Equation 4

Bogartz's model is captured by the equation

$$r(O, T) = g(O)e^{-f(T)}.$$

Hence,

$$\ln[r(O, T)] = \ln[g(O)] - f(T).$$

At time, ΔT later,

$$\ln[r(O, T + \Delta T)] = \ln[g(O)] - f(T + \Delta T)$$

or

$$\ln[r(O, T + \Delta T)] = \ln[r(O, T)] + f(T) - f(T + \Delta T).$$

Clearly, there exists a function h such that

$$f(T + \Delta T) = h[f(T), \Delta T].$$

So,

$$\ln[r(O, T + \Delta T) = \ln[r(O, T)] + f(T) - h[f(T), \Delta T]$$

and

$$r(O, T + \Delta T) = r(O, T)e^{[f(T) - h[f(T), \Delta T]]},$$

which has the form of Equation 4.

Received October 6, 1989
Revision received January 16, 1990
Accepted January 18, 1990 ■

1991 APA Convention "Call for Programs"

The "Call for Programs" for the 1991 APA annual convention will be included in the October issue of the *APA Monitor.* The 1991 convention will be held in San Francisco, California, from August 16 through August 20. Deadline for submission of program and presentation proposals is December 14, 1990. This earlier deadline is required because many university and college campuses will close for the holidays in mid-December and because the convention is in mid-August. Additional copies of the "Call" will be available from the APA Convention Office in October.

Call for Nominations for *Developmental Psychology*

The Publications and Communications Board has opened nominations for the editorship of *Developmental Psychology* for the years 1993–1998. Ross D. Parke is the incumbent editor. Candidates must be members of APA and should be available to start receiving manuscripts in early 1992 to prepare for issues published in 1993. Please note that the P&C Board encourages more participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. To nominate candidates, prepare a statement of one page or less in support of each candidate. Submit nominations to

Norman Abeles
Department of Psychology
Michigan State University
Psychology Research Building
Room 129, Bogue Street
East Lansing, Michigan 48824

Other members of the search committee are Frances D. Horowitz, University of Kansas; Anne Pick, University of Minnesota; Alexander W. Siegel, University of Houston; and Sheldon White, Harvard University. First review of nominations will begin January 15, 1991.