Journal of Experimental Psychology:
Learning, Memory, and Cognition
1985, Vol. 11, No. 4, 817–820

0278-7393/85/$00.75

Consistency and Confoundings: Reply to Slamecka

Geoffrey R. Loftus
University of Washington

Two issues are discussed. First, the vertical comparison method favored by Slamecka (1985) is subject to scaling problems that yield inconsistent conclusions, both within and across experiments. The existence of such inconsistencies is illustrated with data drawn from the literature, and the reason for such inconsistencies is illustrated with a hypothetical example. The second issue involves confoundings. Slamecka (1985) notes that the horizontal comparison method requires that degree of learning be confounded with list age. Similarly, the vertical comparison method requires that degree of learning be confounded with the performance level over which forgetting is assessed. The confounding noted by Slamecka is identified with a potential failure of Loftus's (1985) assumption that equal levels of performance imply identical states of the cognitive system. In situations where this assumption does fail, forgetting of high-learning and low-learning lists must occur for qualitatively different reasons, and any comparison of forgetting curves is therefore inappropriate unless made within the context of a strong theory.

There are (at least) two reasons to conduct experiments: to test a theory, or to answer a question. Slamecka (1985) makes it clear that the Slamecka and McElree (1983) experiments were not performed to test theory. Rather, they were performed to answer the experimental question: Does forgetting rate (forgetting per unit time) depend on degree of original learning?

Answering this question requires a definition of forgetting. Slamecka and McElree (1983) and Slamecka (1985) defined forgetting to be the drop in the value of some dependent variable from one retention interval to a subsequent retention interval. This will be termed the *observed performance definition* (OPD) of forgetting.[1] Given the OPD, the experimental question becomes: Does the magnitude of the performance drop from one retention interval to another depend on degree of original learning? Answering this question requires a vertical comparison of forgetting curves.

Loftus (1985), in contrast, conceptualized forgetting as a mental phenomenon. He defined forgetting to be the change from one mental state (that produces high performance) to a subsequent mental state (that produces lower performance). This will be termed the *mental state definition* (MSD) of forgetting. Assuming a one-to-one correspondence between mental state and memory performance, forgetting is reflected by the drop in performance from one level to a lower level. Given the MSD, the experimental question becomes: Does the time required for performance to drop from one level to another depend on degree of original learning? Answering this question requires a horizontal comparison of forgetting curves.

Which Definition is Most Useful?

As long as forgetting curves are monotonic, but nonlinear, Slamecka (1985) is correct in his assertion that, given any set of data, the two definitions can produce conflicting conclusions. If, by one of the definitions, forgetting is *not* affected by degree of original learning, then, by the other definition, for-

The writing of this article was supported by National Science Foundation Grant BNS82-09617 to the author. Discussions with Donald Bamber provided a good deal of focus for the ideas presented herein. Tom Nelson and Beth Loftus provided helpful critiques of an earlier version of the manuscript.

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

[1] Actually, Slamecka and McElree (1983) investigate the drop in performance over retention interval for each of six dependent variables. Since, according to Slamecka, these dependent variables do not reflect any common, underlying theoretical construct (such as "amount of information in memory"), forgetting as measured by any one of the dependent variables is a different phenomenon from forgetting as measured by any of the others. Technically, therefore, the OPD is a collection of six separate forgetting definitions, one corresponding to each of the dependent variables.

getting *is* affected by degree of original learning. Therefore, Slamecka argues, in order to have a consistent answer to the experimental question, one definition or the other should prevail.

I agree. However, there is a second requirement for having a consistent answer to the experimental question: The definition that does prevail must yield consistent conclusions both within and across experiments. The MSD meets this requirement in the sense that all experiments considered both by Slamecka and McElree (1983; see Figures 1–4) and by Loftus (1985; see Figures 3 and 4) result in horizontally diverging high-learning and low-learning forgetting curves. Thus, all experiments yield the conclusion that forgetting is slower with higher original learning.

The OPD does not meet this consistency requirement. Experimental data do not yield consistent vertical relationships among forgetting curves, either within or between experiments.[2]

Within most experiments, high-learning and low-learning forgetting curves are vertically parallel over some retention–interval ranges, but not over other retention–interval ranges. Slamecka dismisses these intra-experiment inconsistencies as being due to floor effects (e.g., as in Krueger, 1929; Postman & Riley, 1959) or to ceiling effects (e.g., as in Slamecka & McElree, 1983, Experiment 3). However, given Slamecka's definition of forgetting, floor and ceiling effects should be irrelevant, since such effects are problematical only insofar as they mask a relationship between a dependent variable and an underlying construct.[3] So, for example, by the OPD, a performance drop from 1.00 to 0.99 must be viewed simply as forgetting of 0.01; otherwise, the definition itself becomes inconsistent.

The vertical relationships among forgetting curves are inconsistent across experiments as well as within experiments. In some experiments forgetting curves converge vertically (e.g., Postman & Riley, 1959); in other experiments forgetting curves diverge vertically (e.g., Hellyer, 1962); in still other experiments, forgetting curves are vertically parallel (e.g., Underwood & Keppel, 1963).

This lack of consistency in vertical relationships is not surprising, given the scaling problems inherent in the vertical comparison method (see Loftus, 1985, for references) and can be explained as follows. Whether or not one is interested in underlying processes, such processes do exist. Moreover, the effect of some focal variable, such as degree of learning,

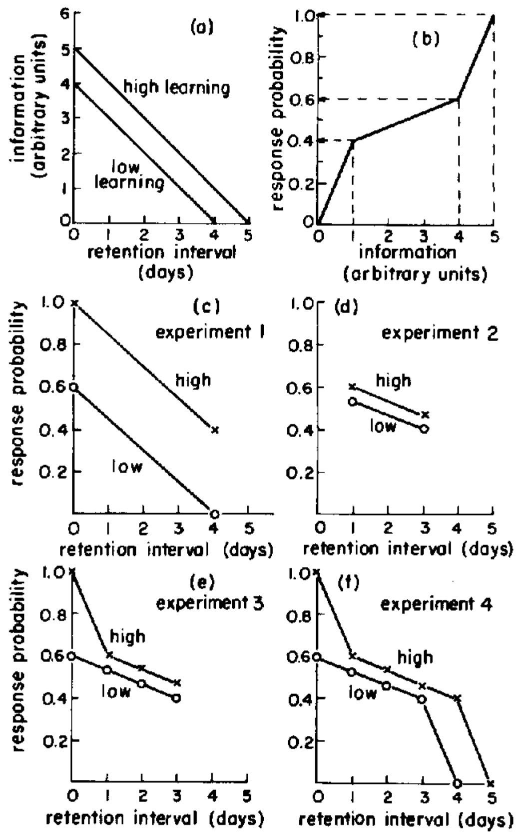

Figure 1. a: Hypothetical functions relating amount of information to original learning and retention interval. b: Hypothetical function relating response probability to amount of information. (The dashed lines show the mapping from information in memory to response probability for Experiment 1.) c–f: Hypothetical results for different experiments involving different retention intervals.

must be on the underlying process, not directly on performance, since the former determines the latter, and not vice versa. If there is any sort of complex relationship between an underlying process and performance, then a consistent effect of the focal variable on the former will lead to an inconsistent effect of the focal variable on the latter. The example provided in Figure 1 illustrates why this is so.

Consider the underlying construct of "amount of information in memory." Hypothetical effects of learning and retention interval on amount of information are represented in Figure 1a. The function relating probability correct, the observed dependent variable, to amount of information is represented in Figure 1b. The exact shape of this function is not important for the argument; what *is* important is that the function is nonlinear. Any nonlinear function would lead to the same conclusion.

Suppose now that each of four experimenters

[2] Contemporary methodologists have discussed numerous other problems with the "measurement by fiat" technique, of which the OPD is an example. See Blalock (1982, chap. 1) for a thorough discussion.

[3] I thank Tom Nelson for bringing this point to my attention.

performs an experiment. The four experiments are identical except in terms of the particular retention intervals that are selected. Consider first Experimenter 1, who has selected retention intervals of 0 and 4 days. The resulting data from Experiment 1 can be determined as follows. According to Figure 1a, the high-learning and low-learning conditions at a retention interval of 0 (designated H(0) and L(0), respectively) produce 5 and 4 units of information. As indicated by the dashed lines in Figure 1b, these 5 and 4 units of information yield response probabilities of 1.0 and 0.6, respectively. Likewise, the H(4) and L(4) conditions produce 1 and 0 units of information, respectively, which correspond to response probabilities of 0.4 and 0.0. These hypothetical data are plotted in Figure 1c. Since the curves are vertically parallel, Experimenter 1 would conclude that, by the OPD, degree of original learning does not affect forgetting rate.

The results of the other three experiments can be determined by the same process, and are shown in Figures 1d–1f. Experimenter 2 has selected retention intervals of 1 and 3 days. Experiment 2, like Experiment 1, produces vertically parallel curves, and Experimenter 2, like Experimenter 1, concludes that degree of original learning does not affect forgetting rate. Notice however that, by the OPD, the "effect" of degree of original learning (that is, the high learning/low learning difference) is 0.400 in Experiment 1, but only 0.067 in Experiment 2. Thus, even when the underlying process is the same, Experimenters 1 and 2 come to quantitatively different conclusions. Experimenter 3 has selected retention intervals of 0, 1, 2, and 3 days. Since there is an interaction in the data of Experiment 3, Experimenter 3, unlike Experimenters 1 and 2, concludes that forgetting rate *is* affected by degree of original learning. Finally, Experimenter 4 selects five retention intervals. Overall, Experimenter 4 finds an interaction between retention interval and degree of original learning and, like Experimenter 3, concludes that forgetting rate is affected by degree of original learning. The data from Experiment 4 also make it quite clear that an interaction would be obtained in some experiments, but not in others; and that an interaction would be in one direction in some experiments, but in the opposite direction in others, all depending on which retention intervals happened to be selected. It is a simple exercise to show that similar confusion would ensue from different experiments in which different degrees of original learning were selected.

In short, given the OPD and the associated vertical parallelism test, inconsistent conclusions are expected and obtained. This means that the OPD is not a very good definition. But is the MSD and the associated horizontal parallelism test free of problems? Slamecka (1985) thinks not.

Confoundings

Slamecka points out that when the horizontal comparison method is used, degree of original learning must covary with list age. This is correct. However, when the vertical comparison method is used, there is analogous covariation. With the vertical comparison method, two forgetting curves being compared for different learning levels, but at the same retention interval, differ in performance level. Slamecka suspects that the nature of forgetting from one performance level to another may depend on list age. Likewise, however, the nature of forgetting from one retention interval to another may depend on the performance levels over which forgetting occurs.

One might argue that these two types of "confoundings" are different, since the first involves two independent variables, whereas the second involves an independent and a dependent variable. Such an argument would require an unwarranted distinction between dependent and independent variables. Traditionally, the kinds of forgetting experiments under consideration here are organized such that retention interval is the independent variable and performance is the dependent variable. But that's done simply as a matter of convenience. One could just as legitimately make performance the independent variable and retention interval the dependent variable; that is, one could ask how long a retention interval is required to reach a performance level of X? This technique is common in psychophysical experiments. One sets performance at some level (e.g., 75% detection), and determines how much of something (e.g., number of quanta per unit time) is required to achieve this performance for different levels of another variable (e.g., for lights of different wavelengths).

The One-to-One Assumption

That the vertical comparison method also suffers from a confounding does not, of course, eliminate the problem for the horizontal comparison method. Within the context of the general model proposed by Loftus (1985, p. 399), Slamecka is questioning the validity of the assumption that there is a one-to-one correspondence between state of the cognitive system and memory performance. A consequence of this one-to-one assumption is that any performance level implies a unique state of the cognitive system.

Loftus (1985, pp. 402–403) points out that the one-to-one assumption is almost certainly false in some circumstances. An obvious example would be high and low learning obtained by imagery and rote repetition instructions, respectively. Failure of the one-to-one assumption, however, means that forgetting in the high-learning and low-learning conditions occurs for qualitatively different reasons. Under such circumstances, it would be inappropriate to compare forgetting curves either vertically or horizontally or in any other simple, qualitative way—it would be an apples-and-oranges comparison. Evaluation of the effects of learning on forgetting would, instead, require a stronger model of learning, forgetting, and performance. Wickelgren's single-trace fragility theory (e.g., 1974) represents one such model, and Slamecka's (1985) item-selection hypothesis (p. 815) represents the start of another such model. Both models explain why forgetting curves consistently diverge horizontally. Neither model (nor any other extant model), however, predicts anything simple about vertical relationships among forgetting curves. Slamecka and McElree (1983) themselves make this latter point, noting of their results that, "No current theories of memory predict these outcomes, but neither does the pattern of results disconfirm any theory" (p. 384).

Et Alia

The foregoing comprises the essentials of my reply to Slamecka (1985). There are, in addition, two minor points that I would like to address.

Simultaneously Equating Performance and List Age

Slamecka (1985, p. 815) notes that when examining the effects of most focal variables on forgetting, it is possible to equate performance for two (or more) levels of the focal variable at the time of initial acquisition. If forgetting from the time of initial acquisition to any subsequent retention interval is compared for different levels of the focal variable, then vertical and horizontal comparisons must yield the same conclusions. This is true essentially because any statistical interaction would then be invariant over monotonic transformations of the dependent variable (see Loftus, 1978).

Suppose, however, that forgetting curves do diverge in this situation. If one wished to compare different levels of the focal variable between any two retention intervals *subsequent* to the time of initial aquisition (retention intervals at which there existed performance differences between levels of the focal variable), then all the scaling problems that typically bedevil the vertical interaction method would immediately return. The horizontal comparison method, in contrast, would have none of these problems.

Negative Acceleration of Forgetting Curves

Slamecka (1985, p. 815) is incorrect in his assertion that slower forgetting of old lists relative to new lists is a necessary consequence of negatively accelerated forgetting curves. A young list, momentarily at the same level of performance as an old list, may indeed be on the steeper section of the forgetting curve, but it may also be on a section of equal steepness (see Loftus, 1985, Figures 2b and 2a). It may also be on a shallower section.

References

Blalock, H. M., Jr. (1982). *Conceptualization and measurement in the social sciences.* Beverly Hills: Sage Publications.

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

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

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

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

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. (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.

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

Wickelgren, W. A. (1974). Single-trace fragility theory of memory dynamics. *Memory & Cognition, 2,* 775–780.

Received March 12, 1985
Revision received April 5, 1985 ■