Journal of Experimental Psychology:
Learning, Memory, and Cognition
1993, Vol. 19, No. 6, 1450–1453

0278-7393/93/$3.00

List-Strength and List-Length Effects: Reply to Shiffrin, Ratcliff, Murnane, and Nobel (1993)

Bennet B. Murdock and Michael J. Kahana

R. M. Shiffrin, R. Ratcliff, K. Murnane, and P. Nobel (1993) claimed that TODAM (a theory of distributed associative memory) is unable simultaneously to predict an absent (or negative) list-strength effect (LSE) and a positive list-length effect (LLE). However, Shiffrin et al. failed to distinguish between situations in which lag (number of items intervening between study and test) is controlled and situations in which it is not. We stand by our previous conclusion; TODAM can explain why there is little or no LSE when at the same time there is an LLE when the LLE is studied under the standard conditions. To our knowledge there are no published studies where lag has been controlled. However, this simplified version of TODAM cannot explain an LLE when a scoring window is used. Whether such a result would be inconsistent with a more complete version of TODAM remains to be seen.

Murdock and Kahana (1993) derived TODAM (a theory of distributed associative memory) predictions for the list-strength effect (LSE), the list-length effect (LLE), and the d' from a study–test paradigm and showed that, with reasonable parameter values, the predicted results were in good agreement with the general pattern of data in the literature. They assumed that memory carried over from the preexperimental to the experimental situation and from list to list; they called this assumption the continuous memory assumption. Shiffrin, Ratcliff, Murnane, and Nobel (1993) claimed that they had analyzed this model and rejected it because it is unable, simultaneously, to predict an absent (or negative) LSE and a positive LLE. Shiffrin et al. reiterated this point with an informal analysis of TODAM, which was buttressed by a simulation using the Murdock and Kahana (1993) parameter values.

The main problem with the analysis of Shiffrin et al. (1993) is that, for the LLE, they failed to distinguish between the situations where lag is controlled and the situations where it is not. (By *lag* we mean the number of items intervening between the study of an item and its test.) Shiffrin et al.'s analysis applies to the former case, whereas our analysis applies to the latter case. It is important to understand that in studies of the LLE, lag is generally not controlled; instead, in the usual study–test paradigm all list items are tested, and there are an equal number of old and new items in the test phase. Items are randomly ordered in the test phase, so perforce lag and list length must be confounded. In fact, to our knowledge, there are no published studies of the LLE where lag is properly controlled.

Bennet B. Murdock and Michael J. Kahana, Department of Psychology, University of Toronto, Toronto, Ontario, Canada.

This research was supported by Grant APA 146 from the Natural Sciences and Engineering Research Council of Canada.

Correspondence concerning this article should be addressed to Bennet B. Murdock, Department of Psychology, University of Toronto, Toronto, Ontario, Canada M5S 1A1. Electronic mail may be sent to murdock@psych.toronto.edu.

This is the situation we analyzed in Murdock and Kahana (1993), and our analysis thus applies to the standard LLE. We showed that TODAM can predict a standard LLE as well as a negligible LSE with the same parameter values (see Tables 1, 2, and 3 of Murdock & Kahana, 1993). It does not much matter whether strength is varied by presentation duration or by repetition; we analyzed both cases. We also showed that TODAM can predict reasonable study–test d' values with these same parameter values (see Table 4 of Murdock & Kahana, 1993). As far as we can tell, Shiffrin et al. (1993) did not dispute these claims.

Shiffrin et al. (1993) analyzed a special case of the LLE (namely, where lag is controlled), but we do not know of any studies of the LLE in which lag has been experimentally controlled. In the various LSE studies of Shiffrin and his colleagues (e.g., Murnane & Shiffrin, 1991, especially Experiment 4), an interpolated distraction task was used to equate elapsed time between study and test across conditions, but elapsed time is not the same as lag. Furthermore, we have known since Waugh and Norman (1965) that the critical variable in recognition memory is the number of intervening items (i.e., lag), not the elapsed time.

As far as we can tell, Shiffrin et al. (1993) did not fault our arguments or dispute our claim that TODAM can handle the standard LLE. Instead, they focused on a special case of the LLE; namely, the LLE when lag is controlled. The critical evidence that they cited was an old experiment from our laboratory (Experiment 3 of Ratcliff & Murdock, 1976), which used a standard study–test procedure with five different list lengths (4, 8, 16, 32, 64). Five subjects were tested for 20 sessions each, with four sessions per list length, so list length was a between-sessions variable. The telling result that they reported is the clear presence of an LLE for List Lengths 16, 32, and 64 when lag is controlled.

Again, Shiffrin et al. (1993) spoke loosely. Lag was not controlled; instead, a scoring window was used. The only items scored in lists with 32 or 64 items were items in the last 16 input positions and the first 32 test positions. Thus, in addition to the fact that different numbers of items were studied, more items were tested on previous lists even though they were not scored. Thus, in this experiment, number of items tested and list length were confounded.

Why should this matter? In Murdock and Kahana (1993) we said that "*one* [italics added] reason the LLE comes about is because, on average, the study–test lag is greater in long lists than in short lists so the mean is different" (p. 692). However, there are certainly other possible reasons too. One is output interference. The more items tested, the more items are added to the memory vector and, by the continuous memory assumption, the more potential interference there is on subsequent trials. Another possible reason (mentioned in Shiffrin, Ratcliff, & Clark, 1990) is fatigue or decreased attention. Long test trials are more demanding, and this could decrease encoding efficiency on subsequent trials (modeled by encoding probability in TODAM). The first factor could be controlled by using the same number of test trials regardless of list length, and the second factor could be controlled by varying list length within session rather than between sessions. Because, to our knowledge, no studies have been conducted that use both of these controls, the presence of an LLE with lag controlled is still an open question.

Considering TODAM's ability to account for the standard LLE, the fact that the version of TODAM used by Murdock and Kahana (1993) cannot explain the LLE in the special case described by Shiffrin et al. (1993; namely, when a scoring window is used) is not crucial. Furthermore, output interference, which may play a crucial role in the LLE when lag is controlled, was explicitly not included in the Murdock and Kahana TODAM version. Perhaps an augmented version of TODAM could also handle this special case.[1]

Contrary to what was claimed by Shiffrin et al. (1993), Shiffrin et al. (1990) did not derive any results using the continuous memory assumption; they simply presented the standard TODAM analyses (e.g., Murdock, 1982, 1989; Murdock & Lamon, 1988) using a slightly different notation. Shiffrin et al. (1990) did consider the possibility that items from prior lists are present in the memory vector. They showed that for massed repetitions or presentation time the predicted ratio of ratios (ROR) should be greater than 1.0 because $\sigma_N^2(PS) > \sigma_N^2(PW)$ (per their Equation 21 in which PS = pure–strong and PW = pure–weak), but they did not say by how much because they did not work out the values of C_p, $\eta_1 C_i$, or $\eta_2 C_i$ in their Equation 20. The same comment applies to the values of C_p, β, δ_i, and $C_{k(i)j}$ in Shiffrin et al.'s Equation 22.

In fact, TODAM does predict that $\sigma_N^2(\mathrm{PS}) > \sigma_N^2(\mathrm{PW})$, so ROR should be greater than 1.0. However when one works out what the numbers should be, the ROR values are only marginally greater than 1.0 (see Tables 1 and 2 of Murdock & Kahana, 1993). Although Shiffrin et al. (1990) admitted that "this model [prior items in memory] deserves further exploration" (p. 187), they went on to say, "we will not pursue it for several reasons" (p. 187). One reason is that "the explanation requires enormous recency effects, which are not seen in the data" (Shiffrin et al., 1990, p. 187). This may or may not be the case.

We are not sure what "enormous" recency effects are, but the slope of the recency effect depends on α, the forgetting parameter in TODAM, and the data of Murdock and Hockley (1989) show that the value of α must be very close to 1.0. We used a value of .995 in Murdock and Kahana (1993), which will make the slope of the recency curve very gradual so the recency effect will be quite extensive. The analysis of the Ratcliff and Murdock (1976) experiment reported in Shiffrin et al. (1993) underestimates the slope. Shiffrin et al. used a common false-alarm rate for all lags. We reanalyzed these data using a false-alarm rate that corresponded to the matched lag block, and perhaps for this reason the observed recency effects are somewhat greater (Figure 1).

MINERVA 2 (Hintzman, 1988), CHARM (a composite holographic associative recall model; Metcalfe-Eich, 1982), and the Matrix model (Humphreys, Bain, & Pike, 1989) all lack forgetting parameters, though they could easily be added. This means that these models must rely on a buildup of variance to produce an LLE, and this buildup of variance will also produce a substantial LSE, contrary to data. For these models Shiffrin et al. (1993) are probably correct. Thus, not all the global matching memory models will be able to produce a negligible LSE coupled with an LLE even under standard conditions. In TODAM, however, the forgetting parameter (α) plays a critical role in producing the LLE.

The continuous memory assumption is not new; it was used in simulations reported by Murdock and Lamon (1988) and was discussed by Shiffrin et al. (1990). What was novel were the TODAM derivations that used this assumption and the demonstration (apparently not disputed) that with reasonable parameter values one could get ROR values that were very close to 1.0 and that did not change much over blocks of trials.

Ratcliff, Clark, and Shiffrin (1990) designed an experiment (Experiment 4) to test the possibility that subjects did not restrict the focus of retrieval to the current list alone, but rather to the entire session. Ratcliff et al. tested this possibility by using lures that had been studied on prior lists. "This procedure should encourage subjects to focus retrieval upon the most recently presented list only (assuming it is possible to do so)" (Ratcliff et al., 1990, p. 166). Because they did not find a positive LSE, they concluded that, "Experiment 4 provides evidence against the hypothesis that the failure to obtain a list-strength effect is due to failure to focus retrieval on the most recent list" (Ratcliff et al., 1990, p. 168). This is not quite the same thing as testing and rejecting the continuous memory assumption.

Finally, Shiffrin et al. (1993) pointed out that the version of TODAM presented and analyzed in Murdock and Kahana (1993) cannot handle some of the detailed findings in the myriad LSE experiments reported by Shiffrin and his colleagues. Of course it cannot—this is a very impoverished version of TODAM. As noted, neither output interference nor

[1] Preliminary simulations where items are similar to one another (e.g., Murdock, 1989) suggest that one may get an appreciable LLE even with a scoring window. However, these simulations used short lists (4, 8, and 16 items), a token number of prior items (4), and only a single value of the similarity parameter ρ (.50), so we have no idea how general this result is. Also, we need a detailed theoretical analysis to understand it.

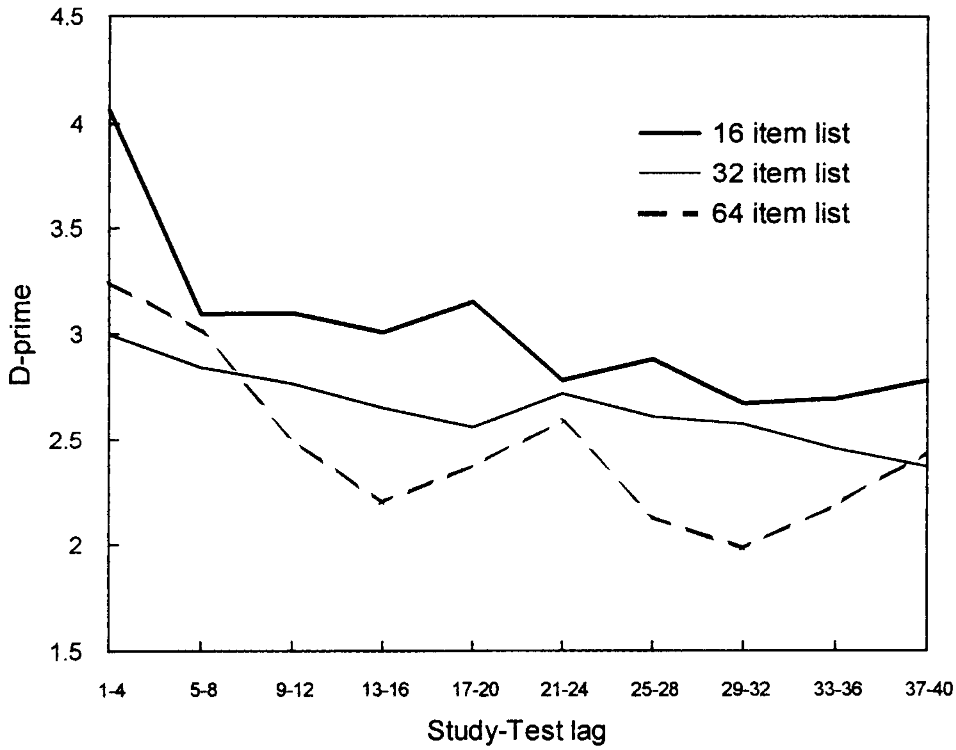

Figure 1. The d' value is shown as a function of study–test lag with separate false-alarm rates at each lag block. (Data are from Experiment 3 of Ratcliff & Murdock, 1976.)

variable encoding (decreased attention) is included, and there are three other major omissions as well. There is no representation of context and no representation of similarity; also, we disregarded the fact that most of the LSE experiments conducted by Shiffrin and his colleagues used a paired-associate presentation format to discourage rehearsal. Yet even this simplified version of TODAM can explain the major pattern of data (absent LSE with present LLE), which Shiffrin and his colleagues have repeatedly claimed is so problematic for global matching memory models.[2]

Consequently, our conclusion stands; a simplified version of TODAM can explain why there is little or no LSE when at the same time there is an LLE when the LLE is studied under the standard conditions. The reasons are very simple; one does not get an LSE (when list length is controlled) because the differential variance from the current list is swamped by the variance from prior lists and from preexperimental memories. However, one does get an LLE under standard conditions because mean lag covaries with list length and, in TODAM, the old-item mean resulting from the dot product of the probe item with the memory vector decreases with lag. As Shiffrin et al. (1993) pointed out, this simplified version of TODAM cannot handle the finding of an LLE when a scoring window is used or when lag is controlled (if such is the case). Whether an augmented version of TODAM could handle such findings remains to be seen.

[2] Shiffrin et al. (1990) also stated that, because recall acts like forced-choice associative recognition in TODAM, assumptions that tend to equalize the variance would then predict that the LSE should also disappear for cued recall. Although this would be true for associative recognition tests (and there does not seem to be an LSE in associative recognition; S. E. Clark, personal communication, April 2, 1993), it would not necessarily be true for cued recall because in TODAM an additional process (deblurring) is also necessary.

References

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

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review, 96,* 208–233.

Metcalfe-Eich, J. (1982). A composite holographic associative recall model. *Psychological Review, 89,* 627–661.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89,* 609–626.

Murdock, B. B. (1989). Learning in a distributed memory model. In C. Izawa (Ed.), *Current issues in cognitive processes: The Tulane Floweree Symposium on Cognition* (pp. 69–106). Hillsdale, NJ: Erlbaum.

Murdock, B. B., & Hockley, W. E. (1989). Short-term memory for associations. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 24, pp. 71–108). San Diego, CA: Academic Press.

Murdock, B. B., & Kahana, M. J. (1993). An analysis of the list-strength effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 689–697.

Murdock, B. B., & Lamon, M. (1988). The replacement effect: Repeating some items while replacing others. *Memory & Cognition, 16,* 91–101.

Murnane, K., & Shiffrin, R. M. (1991). Interference and the representation of events in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 855–874.

Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). The list-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 163–178.

Ratcliff, R., & Murdock, B. B. (1976). Retrieval processes in recognition memory. *Psychological Review, 83,* 190–214.

Shiffrin, R. M., Ratcliff, R., & Clark, S. (1990). The list-strength effect: II. Theoretical mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 179–195.

Shiffrin, R., Ratcliff, R., Murnane, K., & Nobel, P. (1993). TODAM and the list-strength and list-length effects: Comment on Murdock and Kahana (1993a). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1445–1449.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72,* 89–104.

Received April 7, 1993
Revision received June 2, 1993
Accepted June 3, 1993 ■

Instructions to Authors

The editorial policy of the *Journal* favors publication of integrative articles containing multiple experiments and extensive theoretical development. However, articles reporting single experiments are also appropriate, so long as they report compelling data with a clear theoretical message. Other appropriate forms of submission include brief comments on articles published in the *Journal* and replies to such comments. In addition, occasional literature reviews advancing our knowledge in the field will also be considered. For further information on content, authors should refer to the editorial in the January 1990 issue of the *Journal* (Vol. 16, No. 1, pp. 3–4). For information on the other three *JEP* journals, authors should refer to editorials in those journals.

Authors should prepare manuscripts according to the *Publication Manual of the American Psychological Association* (3rd ed.). All manuscripts must include an abstract of 960 characters and spaces (which is approximately 120 words) typed on a separate sheet of paper. Typing instructions (all copy must be double-spaced) and instructions on the preparation of tables, figures, references, metrics, and abstracts appear in the *Manual*. Also, all manuscripts are subject to editing for sexist language.

APA policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications. In addition, it is a violation of APA Ethical Principles to publish "as original data, data that have been previously published" (Standard 6.24). As this journal is a primary journal that publishes original material only, APA policy prohibits as well publication of any manuscript that has already been published in whole or substantial part elsewhere. Authors have an obligation to consult journal editors concerning prior publication of any data upon which their article depends. In addition, APA Ethical Principles specify that "after research results are published, psychologists do not withhold the data on which their conclusions are based from other competent professionals who seek to verify the substantive claims through reanalysis and who intend to use such data only for that purpose, provided that the confidentiality of the participants can be protected and unless legal rights concerning proprietary data preclude their release" (Standard 6.25). APA expects authors submitting to this journal to adhere to these standards. Specifically, authors of manuscripts submitted to APA journals are expected to have available their data throughout the editorial review process and for at least 5 years after the date of publication.

Authors will be required to state in writing that they have complied with APA ethical standards in the treatment of their sample, human or animal, or to describe the details of treatment. A copy of the APA Ethical Principles may be obtained by writing the APA Ethics Office, 750 First Street, NE, Washington, DC 20002-4242.

Masked reviews are optional, and authors who wish masked reviews must specifically request them when submitting their manuscripts. Each copy of a manuscript to be subjected to masked review should include a separate title page with authors' names and affiliations, and these should not appear anywhere else on the manuscript. Footnotes that identify the authors should be typed on a separate page. Authors should make every effort to see that the manuscript itself contains no clues to their identities.

Manuscripts should be submitted in quintuplicate. All copies should be clear, readable, and on paper of good quality. In addition to addresses and phone numbers, authors should supply electronic mail addresses and fax numbers, if available, for potential use by the editorial office and later by the production office. Authors should keep a copy of the manuscript to guard against loss.

Mail manuscripts and address all correspondence at all stages of handling to the Editor, Keith Rayner, Department of Psychology, Tobin Hall, University of Massachusetts, Amherst, Massachusetts 01003.

For the other *JEP* journals, authors should submit manuscripts (in quadruplicate) to one of the editors at the following addresses: *Journal of Experimental Psychology: General*, Earl Hunt, Department of Psychology NI-25, University of Washington, Seattle, Washington 98195; *Journal of Experimental Psychology: Human Perception and Performance*, Thomas H. Carr, Incoming Editor, Department of Psychology, Michigan State University, East Lansing, Michigan 48824; *Journal of Experimental Psychology: Animal Behavior Processes*, Stewart H. Hulse, Johns Hopkins University, Department of Psychology, Ames Hall, Baltimore, Maryland 21218. When one of the editors believes a manuscript is clearly more appropriate for an alternative journal of the American Psychological Association, the editor may redirect the manuscript with the approval of the author.