5 6

Reprinted from:
SHORT-TERM MEMORY
©1975
ACADEMIC PRESS, INC.
New York San Francisco London

CHAPTER 3

MORE ON THE LONG AND SHORT OF MEMORY[1]

WAYNE A. WICKELGREN

A few months after writing the previous article, I generated an alternative single-trace theory that explains even the three phenomena I had thought supported the distinction between short- and long-term memory. The theory was not generated specifically for the purpose of providing a

[1] This work was supported by Grant 3-0097 from the National Institute of Education and by Contract F 44620-73-C-0056 from the Advanced Research Projects Agency, Department of Defense.

single-trace explanation for these phenomena. Instead, I generated a new theory of the long trace to account for a phenomenon contradictory to an earlier version of the theory. Then I realized that this new theory of the long trace would probably eliminate the need to assume any dynamically distinct short trace at all. This appears to be true, and I now believe that the conclusion of the previous paper can be strengthened to say that there is *no* evidence favoring a distinction between short-term and long-term memory.

The purpose of this addendum is to present the single-trace theory briefly and explain how one can account for the form of the retention function, the effects of similarity on storage interference, and long-acquisition amnesia within a single trace formulation. The explanation of the first two of these phenomena follows from the particular single-trace theory described, but the single-trace explanation of long-acquisition amnesia could have been provided in the absence of this particular theory. I simply had a blind spot previously for the inadequacies of the amnesic-patient evidence.

I. SINGLE-TRACE, DUAL-DECAY THEORY

The theory of long-trace storage dynamics described in Wickelgren (1972) was a pure interference theory, that is, the decay of the long trace occurred solely due to storage interference (unlearning) caused by acquiring subsequent similar traces. The results of many studies summarized in Wickelgren (1972) support the existence of a storage interference process acting to reduce the strength of the long trace in proportion to the degree of similarity. However, according to the interference theory described in Wickelgren (1972), the long trace has a second important dynamic property, resistance (to interference). The resistance of the long trace to interference was assumed to increase with increasing trace age. For reasons irrelevant to this discussion, I now assume that the second property is trace fragility, which decreases with increasing trace age. The assumption of gradually decreasing trace fragility appears to be necessary to account for the continual decrease in the rate of forgetting and the temporally defined character of retrograde amnesia and recovery from it. The trace fragility property also accounts for the ability to make recency judgments under conditions that prevent the use of associations to time concepts. Thus, the assumption that the long trace has a second dynamic property that depends only on trace age is well motivated.

However, the combination of the decreasing trace resistance on the one hand with a pure interference theory of forgetting on the other hand leads to the prediction that the storage interference effect of similar learning (AB–AC versus AB–CD) learning will be greater the earlier the similar interfering learning is interpolated during a retention interval. Previous studies using a recall measure (Archer & Underwood, 1951; Houston, 1967; Newton & Wickens, 1956) provide no support for this prediction, but one could argue that the recall studies are confounded by competition effects and other retrieval interference problems that mask the expected storage interference effects. A previous study by Howe (1969) that used a recognition measure also yielded results contradictory to this prediction. However, I was not convinced without doing the experiment myself, so I did two experiments manipulating the delay between original and similar interpolated learning. I found a negative storage interference effect of similar interpolated learning, but one that appeared to be independent of the delay between original and interpolated learning. Thus, a central prediction of the previous theory appeared to be convincingly contradicted.

The most obvious solution to this problem is probably to assume that two factors produce decay of the long trace: (*a*) storage interference from subsequent similar learning and (*b*) an interference-free time-decay process. Now one can assume that the decrease in the fragility of the trace affects the time-decay process but has no effect on the interference process (the latter being affected only by the similarity of subsequent learning to original learning).

I resisted formal consideration of this dual-process theory for some time since it seemed more complex than a single-process theory. However, after many fruitless attempts to formulate a theory consistent with all of the previously mentioned phenomena, I was forced to formulate explicitly a differential equation model of the dual-process theory. When I did this, I immediately noticed that the resulting form of the retention function for the long trace was a product of a power function decay term and an exponential decay term. Immediately, I saw that such a more complex form of retention function could provide both the power function decay characteristic of long-term retention and the exponential decay characteristic of certain types of short-term retention. This meant that a single trace with two decay processes (time and interference) might permit one to eliminate the assumption of a separate short-term memory trace altogether. If so, there was nothing *ad hoc* and inelegant about the two-process theory.[2]

[2] From a philosophy of science point of view, I think it might be well to note in passing how foolish it would have been not to have considered the two-process

According to the single-trace, two-process theory, the form of the retention function is as follows:

$$d_m = \lambda(1 + \beta t)^{-\psi}(e^{-\pi t})$$

where

$$\alpha, \beta, \psi, \text{ and } \pi > 0$$

In this equation, d_m represents the interval scale d' measure of memory strength, λ represents the degree of learning, β and ψ represent rate parameters characterizing the time decay process, and π represents the rate parameter (degree of similarity) for the interference process.

II. FORM OF THE RETENTION FUNCTION

In the above equation for the long-term retention function, if π is close to zero, the form of the retention function will approximate a power function, as has been obtained for long-term retention (Wickelgren, 1972). On the other hand, if π is large, as will occur when the interpolated material is highly similar in its encoding to the originally learned material, then the form of the retention function will approximate exponential decay, as is characteristic of certain types of short-term retention (Wickelgren & Norman, 1966; Wickelgren, 1970). In fact, the conditions under which power function decay is obtained are precisely those in which encoding is known to be largely semantic in nature and therefore characterized by low similarity between original and interpolated material. The conditions under which exponential decay is obtained are precisely those in which encoding is known to be largely phonetic, characterized by high similarity

---

theory because it seemed like a purely *ad hoc* explanation for a particular phenomena. All new ideas are probably initially *ad hoc* in that they are generated to account for one or at most two phenomena. The important question is not whether a theory is initially *ad hoc,* but whether it remains *ad hoc,* never accounting for additional phenomena. To rule out any consideration of a theory because it is initially *ad hoc* is quite wrong in that it will prevent one from investigating "interesting" theoretical reformulations of data. Thus, one must be quite careful in the uncritical use of the attribution *ad hoc* to a component of a theory. I think we have no better way to judge whether a theoretical idea is worth pursuing than our own intuition that it is "interesting," whatever that may mean. There is nothing more absurd than to reject a useful criterion, such as our intuitive judgment that an idea is interesting, simply because we do not understand that intuition, in favor of a criteria, such as "*ad hocness,*" which we may understand somewhat better, but which is a much less relevant criterion.

between original and interpolated material. In studies of short-term retention in which the intervening material is very dissimilar to the originally learned material, one does not obtain rapid exponential decay. Thus, there appears to be an extremely plausible single-trace explanation for obtaining two different forms of retention functions under these different conditions, and this evidence does not support a two-trace theory.

III. SIMILARITY AND STORAGE INTERFERENCE

As discussed in the preceding paper, the similarity of interpolated learning appears to have no effect on short-term retention in most tasks where a pure exponential decay has been obtained, while greater similarity produces greater interference in the longer-term retention studies characterized by power function decay. This important qualitative difference between the nature of interference can be given a single-trace explanation as well.

According to the single-trace theory, greater similarity of interpolated learning should cause greater storage interference no matter what the conditions. However, the magnitude of this effect will depend on the relative magnitude of the disparity between the high similarity and low similarity conditions of interpolated learning being compared.

In the long-term retention studies characterized by a large degree of semantic encoding, the difference in similarity between an AB–AC condition, on the one hand, and an AB–CD condition, on the other hand, is enormous, because the semantic similarity (π) of any two words (AC) is very low. In fact preliminary testing of the single-trace theory indicates that the π parameter is close to zero for any two words drawn from a population of 10,000 common English words. Thus, one is comparing zero similarity with perhaps 50 per cent similarity in comparing the storage interference effects of AB–CD with AB–AC.

By contrast, in the cases where a rapid exponential decay has been obtained, the encoding is thought to be predominately phonetic and each item has a rather high degree of phonetic similarity on the average to any other item. Thus, the discrepancy between the high similarity characteristic of AB–AC and the unknown but still fairly high similarity characteristic of AB–CD will be much lower in this case. So, one expects to find a much smaller effect of similarity in this case than in the former case. Therefore, the negative results I obtained in these situations cannot be taken to provide any convincing evidence that interference operates differently in this type of memory task than in the other type of memory task. Only the magnitude of the effect may be different, and that may be grossly different. Thus, a truly enormous number of trials might be necessary to detect the effect of similarity in these situations whereas only a modest number of trials is necessary in the former class of situations.

Furthermore, when subjects receive an AC interpolated learning trial only a few seconds after receiving an AB learning trial, there may be a greater tendency to go back and rehearse the AB pair (even contrary to instructions) than when one receives a CD interpolated learning trial. Such differential uncontrolled rehearsal could mask a small negative storage interference effect. Thus, these phenomena are totally unconvincing as evidence for the dual-trace theory.

In fact, the single-trace explanation for when a large effect of similarity will be found and when it will not actually accounts for the one discrepant finding (Bower & Bostrom, 1968) mentioned in regard to the effects of similar interpolated material in long-term retention. Although Bower and Bostrum described their findings as applying to short-term memory, this was because the design was similar to that used in short-term memory studies and the materials were letter-digit pairs (likely to have a substantial degree of phonetic encoding). The retention intervals are actually long enough so most experimental psychologists would consider the study to be tapping long-term memory. Thus, the Bower and Bostrom finding was actually contradictory to the dual-trace hypothesis that similarity affected storage interference in long-term memory but not short-term memory. Since the type of materials and design used by Bower and Bostrom are of the type likely to produce phonetic encoding, with high similarity of interpolated and original learning even in the control condition, the single-trace explanation is actually consistent with the Bower and Bostrom result, while the dual-trace explanation is inconsistent. Thus, what little evidence exists actually favors the single-trace explanation over the dual-trace explanation regarding the effects of similarity on amount of retroactive interference in recognition memory.

IV. LONG-ACQUISITION AMNESIA

It is now quite clear that amnesic subjects have relatively normal motor skill and perceptual long-term learning and memory despite their severe deficits in the establishment of new cognitive long-term memory (Baddeley & Warrington, 1970; Corkin, 1968; Warrington & Weiskrantz, 1970). Such findings suggest that amnesic subjects have a modality deficit in the formation of cognitive (semantic) memory traces and do not provide evidence for the reality of dynamically different traces (short-term and long-term memory) either within or across modalities. The fact that amnesic patients often have relatively small impairments on such short-term memory tasks as the memory span test, while having much larger deficits on tests tapping long-term verbal retention, may be simply due to the patients having a relatively unimpaired capacity for phonetic encoding, while having a severely impaired capacity for establishing new semantic memory traces. Under the conditions characteristic of most of the longer-term verbal learning studies, a phonetic trace will be rapidly interfered with and thus rapidly lost, for reasons discussed in the previous paper. Hence, amnesic patients will show severe deficiencies at long-retention intervals. However, it should be clear that the explanation need not be that they lack the capability of forming long-term memory traces, in general, while retaining the capacity to acquire short-term traces. Rather, the most parsimonious explanation appears to be that amnesic patients lack the capability of forming "higher-level" cognitive (including semantic) memory traces, while often retaining substantial capacity for acquiring "lower-level" sensory and motor memories. The modality explanation clearly seems more attractive at the present time than the dual dynamic trace explanation.

V. CONCLUSION

The arguments presented in this and the preceding paper indicate that there is no psychological support whatsoever for distinguishing two dynamically different memory traces, and that appears to be the long and the short of it.

REFERENCES

Archer, E. J., & Underwood, B. J. Retroactive inhibition of verbal associations as a multiple function of temporal point of interpolation and degree of interpolated learning. *Journal of Experimental Psychology,* 1951, **42,** 283–290.

Baddeley, A. D., & Warrington, E. K. Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior,* 1970, **9,** 176–189.

Bower, G. H., & Bostrom, A. Absence of within-list PI and RI in short-term recognition memory. *Psychonomic Science,* 1968, **10,** 211–212.

Corkin, S. Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia,* 1968, **6,** 255–265.

Houston, J. P. Retroactive inhibition and point of interpolation. *Journal of Verbal Learning and Verbal Behavior,* 1967, **6,** 84–88.

Howe, T. S. Effects of delayed interference on list 1 recall. *Journal of Experimental Psychology,* 1969, **80,** 120–124.

Newton, J. M., & Wickens, D. D. Retroactive inhibition as a function of the temporal position of the interpolated learning. *Journal of Experimental Psychology*, 1956, **70**, 237–245.

Warrington, E. K., & Weiskrantz, L. Amnesic syndrome: Consolidation or retrieval? *Nature*, 1970, **228**, 628–630.

Wickelgren, W. A. Time, interference, and rate of presentation in short-term recognition memory for items. *Journal of Mathematical Psychology*, 1970, **7**, 219–235.

Wickelgren, W. A. Trace resistance and the decay of long-term memory. *Journal of Mathematical Psychology*, 1972, **9**, 418–455.

Wickelgren, W. A., & Norman, D. Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology*, 1966, **3**, 316–347.