Altmann, E. M. & Schunn, C. D. (in press). To appear in *Proceedings of the 24th annual meeting of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Integrating Decay and Interference: A New Look at an Old Interaction

Erik M. Altmann **(ema@msu.edu)**
Department of Psychology
Michigan State University
East Lansing, MI 48824

Christian D. Schunn **(schunn@pitt.edu)**
Learning Research & Development Center
University of Pittsburgh
Pittsburgh, PA 15260

Abstract

An old but important debate about human memory concerns whether decay (indexed by time) or interference (indexed by amount of distracting information) is the cause of forgetting. We argue, based on a simple functional analysis, that this is a false dichotomy. Both processes must be at work, in that distracting information must decay to allow the cognitive system to have any hope of retrieving target information amidst the unavoidable clutter of a well-stocked memory. This analysis predicts that subtle decay effects should be pervasive, even in data produced by interference theorists to show that decay was impossible. A re-analysis of data from Waugh and Norman (1965) does indeed reveal decay effects that were dismissed by the authors as inconsequential and have been ignored by most investigators since. We fit a formal model integrating decay and interference to the Waugh and Norman data, and to the decay data of Peterson and Peterson (1959) to show that one model provides an improved account of two ostensibly divergent data sets.

Introduction

> "Decay must be one of the most discredited theories in psychology, amongst many distinguished competitors."
> — Memory researcher Robert Bjork, Michigan State University, Sept. 27, 2000.

How is information lost from human memory? Of the many potential metaphors, the two main competitors have historically been decay (a process indexed by time) and interference (a process indexed by the amount of distracting information "cluttering up" the mental desktop).

Of these two metaphors, decay has been the less popular, as the quotation above suggests. Memory researchers have often simply not wanted to credit the idea that memory could deteriorate by a time-indexed biological process (e.g., Keppel & Underwood, 1962; McGeoch & Irion, 1952; Postman, 1971; Waugh & Norman, 1965). Evidence often cited against decay includes the slowdown of forgetting during sleep (e.g., Ekstrand, 1972), though to interpret this slowdown as evidence against decay one must assume that the decay rate is the same during sleep as during wakefulness. Given the controversial nature of what little we do understand about brain activity during sleep, it seems equally likely that the decay rate is different in different states of consciousness. Another argument against decay is based on the observation that time by itself cannot be causal. As famously put by McGeogh, "In time, iron may rust and men grow old, but the rusting and the aging are understood in terms of the chemical and other events which occur in time, not in terms of time itself" (McGeogh & Irion, 1952, p. 402). Today, many important memory theories exclude decay (e.g., Gillund & Shiffrin, 1984; Hintzman, 1988; Murdock, 1992), and cognitive textbooks often present decay theory as a historical footnote rather than as an active hypothesis (e.g., Ashcraft, 2002; Galotti, 1999; Reed, 2000.)

But decay is far from a footnote. Since the original studies of Brown (1958) and Peterson and Peterson (1959), various approaches have been taken to try to isolate decay from interference (Reitman, 1974; Baddeley & Scott, 1971; Turvey, Brick, & Osborne, 1970). Interference theorists themselves discovered that retention interval moderates proactive interference (e.g., Loess & Waugh, 1967). Decay is represented in select memory theories (Anderson & Lebiere, 1998; Richman, Staszewski, & Simon, 1995) and interpretations of the literature (Anderson, 2000; Baddeley, 1990; Wickelgren, 1977). Finally, it is increasingly clear that McGeoch's polemic (quoted above) misses the mark, given converging evidence that decay has neural correlates. For example, Fuster (1995, p. 247) observes that firing rates of particular pyramidal cells in the monkey show decay "reminiscent of the well-known decay of human short-term memory." And decay in the form of "leak currents" is an integral part of neural network simulation of the hippocampus (O'Reilly & Munakata, 2000).

The current study aims to bolster the case for a general and functional decay process and, more specifically, to show that evidence for decay exists behind "enemy lines," in the very data that purportedly showed that decay was impossible. The paper is organized as follows. We begin with a functional analysis that argues that decay (or some similar, non-interference forgetting process) must function in memory if memory is to function at all. We then develop predictions of this analysis for a classic study in the literature on interference, the Waugh and Norman (1965) probe digit experiment. A re-analysis of the data from that experiment supports the prediction of subtle decay effects. To make the prediction quantitative, we fit the data with a model based on a memory theory that offers the building blocks to integrate decay and interference. Finally, we turn to the original Peterson and Peterson (1959) data set on decay, and fit the same model to it. The goal is to show that one model, with decay and interference represented as functionally interacting processes, parsimoniously accounts for two ostensibly divergent data sets.

A Functional Perspective on Decay

The functional argument for decay is perhaps best illustrated with an example. As one drives an automobile through various speed zones, it is important to mentally register each change in the speed limit and update memory accordingly. However, if each change in the speed limit contributed

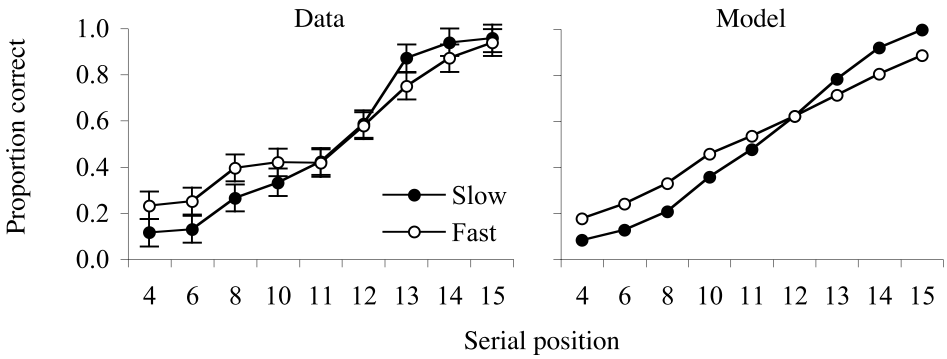

Figure 1. Left: Data from Waugh and Norman (1965), Exp. 1. Error bars are within-participants confidence intervals (Loftus & Masson, 1994) for the serial position x presentation rate interaction. Right: Fit of the functional decay model.

monotonically to interference in memory for speed limits, it would quickly become impossible to remember the current speed limit, whatever it might be. Although interference is clearly a potent source of forgetting, even at long delays (Keppel, Postman, & Zavortink, 1968), there must be some mitigating process if memory is to support our everyday needs (c.f., Luria, 1968).

Interference theorists themselves have been forced to acknowledge this functional need, and the literature consequently has known decay by other names. One example of decay being reinvented is the notion of forgetting through "stimulus fluctuation" (Estes, 1955; Gillund & Shiffrin, 1984; Landauer, 1975). The hypothesis is that changes in the environment gradually reduce the set of cues available to activate distracting information. However, this gradual environmental change in how distractors are cued is typically only specified at an abstract level, in terms that are functionally equivalent to a simple time-indexed process. A second example is the mystical sounding "spontaneous recovery" of previously extinguished distractors. Inherited from the behaviorists, this construct was applied by early interference theorists to explain, among other findings, the effect of retention interval in the Brown-Peterson paradigm (Keppel & Underwood, 1962).[1] Again, however, spontaneous recovery and decay seem functionally equivalent: In one case distractors gain strength relative to the target, and in the other case the target loses strength relative to distractors.

Thus, we argue that any non-interference forgetting indexed by time may as well be known as decay. What we propose to do is play out the functional implications and search for decay in time-based experimental manipulations where one might not otherwise expect to find it.

Revisiting Waugh and Norman (1965)

The classic data set of Waugh and Norman (1965) is discussed in many modern cognitive textbooks (including those cited above) usually to illustrate the importance of interference as a forgetting mechanism. In the *probe-digit* paradigm used in this experiment, the participant is presented with a list of digits and then asked to report one of them. The target digit (the one to report) is indicated by a probe digit given immediately after the list is presented. The probe also occurred exactly once during the list proper, and the target is the digit that followed the probe in the list proper. The serial position of the probe in the list proper changes randomly across trials, so accuracy across trials measures item retention as a function of serial position.

The experiment included a within-subjects manipulation of presentation rate to test whether the chronological age of the target item affected its retention. In the Fast condition, digits were presented once every 250 msec, and in the Slow condition they were presented once per second. The logic was that if decay caused forgetting, then Fast items should be more accurately remembered in response to the probe. If interference caused forgetting, then only serial position should affect retention. That is, late items in the list should be recalled better than earlier items, in both the Fast and Slow conditions, because late items suffered fewer intervening items before the probe, and hence should suffer less retroactive interference.

The Waugh and Norman data are plotted in Figure 1 (left panel), with serial position of the target along the x-axis. The curves represent the two presentation rates. The effect of serial position is readily apparent in both conditions, with later items recalled much more accurately than earlier items. This effect, and its similarity across the two conditions, was enough for Waugh and Norman to reject decay as cause of forgetting. They do note "a slight interaction" of serial position and presentation rate, but dismiss its importance — "the effect of rate is relatively small compared to the effect of serial position" — and report no statistics. Following their lead, no contemporary textbook that we have examined (including those cited above) even acknowledges the interaction, despite the fact that it fairly leaps off the page. The occasional investigator has observed the interaction and speculated about causes (Massaro, 1970; Hintzman, 1978; Wickelgren, 1977), but the theoretical significance of this interaction for decay theory has not been fully pressed.

To confirm the interaction statistically, we conducted a 9 x 2 repeated measures analysis of variance on the data.[2] The interaction of serial position and presentation rate is highly reliable, $F(8, 24) = 5.1$, $p < 0.001$, $MSE = .0033$.

[1] "The increase in [proactive interference] with increase in length of the retention interval may be accounted for by the recovering of extinguished interference associations" (Keppel & Underwood, 1962).

[2] We reconstructed the Waugh and Norman data by digitally scanning the individual participant data graphs in that report and overlaying a grid to estimate the actual numbers.

The Functional Decay Model

To explain the Waugh and Norman interaction, we modeled it using the central memory constructs of the ACT-R cognitive theory (Anderson & Lebiere, 1998). The model's fit is presented in the right panel of Figure 1. In addition to qualitatively capturing the interaction, quantitative measures of fit are quite close, with $RMSD = .054$ and $R^2 = .965$.

This *functional decay* model is based on the activation mechanism illustrated in Figure 2. The two curves in that figure plot the activation of each item in a list, just before the probe is presented. That is, the curves represent a snapshot of every potential target's activation immediately after all the potential targets have been presented. The curves are produced by the following activation function, adapted from ACT-R's Base Level Learning equation.

$$A = \ln(\frac{n}{\sqrt{T}}) \qquad \text{Equation 1: Base-level activation}$$

A is the activation of an item, n is the number of times the item has been retrieved from memory since it was encoded, and T is the age of the item (time from encoding to present). Equation 1 thus computes activation as a function of frequency of use. The premise behind this function is that historical need for information is a predictor of future need and thus should affect item availability (Anderson & Milson, 1989). Activation decays in this function as a power function of age of the item, T. The exponent of this function (-0.5) is, within ACT-R, a relatively constant parameter of the cognitive architecture governing the decay rate of a memory trace. However, this is not the only factor governing an item's base-level activation. For example, a rehearsal process could increase activation by increasing the value of the usage parameter, n. In the Waugh and Norman model, we fix n at 1, on the assumption that items were not differentially rehearsed. Waugh and Norman anticipated the possibility that differential rehearsal could confound their results, and instructed their participants to rehearse only the most recent item of the list, if they rehearsed at all.

In terms of the activation curves in Figure 2, Equation 1 (with n constant across items) predicts that the latest (most recent) item is the most active, the next most recent item the next most active, and so on. Items in the Slow condition are on average less active than items in the Fast condition, because a Slow item at a given serial position is older than a Fast item at the same serial position, so has decayed more.

In addition to the base-level activation governed by Equation 1, the second source of activation for a memory trace under ACT-R theory is priming through associative links. In other words, an item is activated associatively when cues to that item are themselves activated. Associative priming is how ACT-R must explain that fact that elements other than the latest (in the probe digit paradigm) can be retrieved at all. The background theoretical assumption is that, in response to a retrieval request from central cognition, the memory system delivers the trace that is the most active at that instant. In Figure 2, the item with the highest base-level activation is the last one in the list, so if base-level activation were the only activation a memory trace could have, then only the last item would ever be retrieved. This is clearly not how the cognitive system operates, in the probe digit paradigm or in general; people are perfectly able to retrieve thoughts other than the most recent. ACT-R implies that such retrieval depends heavily on retrieval cues delivering activation through associative links. In Bayesian terms, base-level activation reflects the influence of history (retrieval history, in particular, as captured by base-level activation), whereas associative priming from retrieval cues reflects the influence of the current context.

In the probe-digit paradigm, we assume that the role of associative priming plays out through an associative link between the probe and the target. That is, when the probe and the target co-occur in the list proper, this co-occurrence causes an associative link to be encoded between the two traces in memory. When the probe is re-activated at the end of the list, activation spreads from the probe to the target through this link, priming the target. This assumption is grounded in associatonism generally and various memory theories in particular (e.g., Gillund & Shiffrin, 1984), and in specific evidence that such associations are formed between neighbors in a list of random items (Nairne, 1983).

In the functional decay model, associative priming from the probe is implemented simply as a constant amount of activation added to the target. The effect of this priming is also illustrated in Figure 2, by the arrow labeled *priming*. Whereas the curves represent item activations immediately before the probe is presented, the elevated point at the head of the arrow represents the activation of a target when the probe is presented. (We have arbitrarily chosen the target to be the item at serial position 7, in the Fast condition.) The target item is shown to have much more activation than its neighbors, and also more activation than most recent list items (which would otherwise be the most active).

One other necessary model parameter, not represented in Figure 2, captures the differential effect (across the two presentation rates) of proactive interference due to previous trials. According to Equation 1, old items decay but

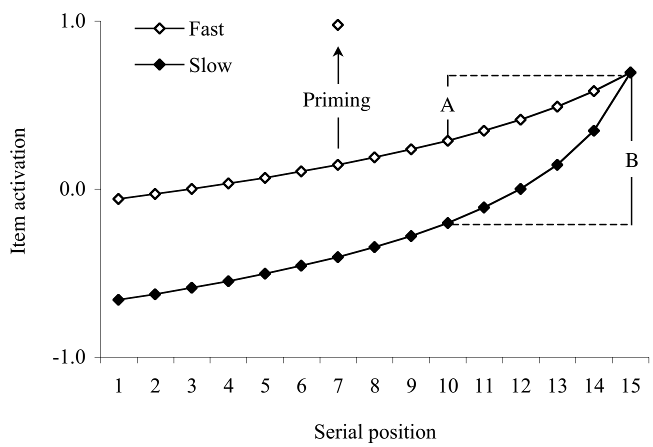

Figure 2: A snapshot of item activations from Equation (1) at the instant the full list of items has been presented. The arrow labeled *priming* indicates the magnitude of associative activation passing from probe to target (item 7, in this scenario) when the probe is presented. A and B are activation differences used to illustrate relative activation (see text).

nonetheless retain some activation well into the future, meaning that items from previous trials ("prior items") will continue to exist in memory as a source of proactive interference. However, the age of these prior items will differ across the Fast and Slow conditions, in that prior items in the Fast condition will retain more of their activation and hence cause greater proactive interference. This difference across the two conditions is captured in a prior-items parameter in the model. The parameter is implemented in terms of a single distracting element with an activation value that is estimated from the data.

The final step is to map item activations to the probability of retrieving a given target item. In ACT-R this mapping is the "soft-max" rule below, which predicts that the most active item has the highest probability of being retrieved (as we discussed above), but that other items can intrude from time to time. This rule determines the probability of retrieving a given item as a function of its activation relative to the activation of its distractors, and thus specifies the extent to which distractors interfere with the target.[3]

$$P_i = \frac{e^{A_i/s}}{\sum_j e^{A_j/s}} \qquad \text{Equation 2: Retrieval probability}$$

P_i is the probability of retrieving i, the target, and A_i is the target's activation (from Equation 1). The quantity s represents the assumption that memory is susceptible to noise (i.e., transient fluctuations in activation levels).

Fitting Waugh and Norman (1965)

We can now describe how the model captures the interaction in Figure 1. One of the two basic patterns in the data is that late items are recalled better in the Slow condition than in the Fast condition. In the model, this effect results from late items having more *relative* activation in the Slow condition than in the Fast condition, as compared with distracting elements. By relative activation, we mean the difference in activation between the target and its distractors. Equation 2, which defines activation-based interference in ACT-R, predicts that the greater the difference in activation between the target and its distractors, the greater the probability of retrieving the target. In Figure 2, relative activation is represented by the differences A and B, which are differences in activation levels of items at two arbitrarily chosen serial positions (10 and 15). The difference A is between items 10 and 15 in the Fast condition, and the difference B is between items 10 and 15 in the Slow condition. When the target is a late item (i.e., 15), then the probability of retrieving it depends on the activation difference between it and earlier items (i.e., 10). This difference is larger in the Slow condition (distance B) than in the Fast condition (distance A). Thus, late items will be recalled more accurately in the Slow condition than in the Fast condition.

The second pattern in the data is that earlier items are recalled better in the Fast condition than in the Slow condition. In the model, this effect again results from relative activation, with target and distractor now reversing roles. With respect to the scenario in Figure 2, when the target is the earlier item (i.e., 10), then the distractor is the late item (i.e., 15). (Recall that although item 10 has less base-level activation than item 15, the associative priming illustrated in Figure 2 will compensate for this deficit and improve the probability of the target being retrieved.) The activation difference between target and distractor now favors the Fast condition, because item 10 in that condition faces a smaller activation deficit relative to its primary distractors (the later items in the list). Thus, earlier items will be recalled more accurately in the Fast condition than in the Slow condition.

Model parameters

The fit in Figure 1 depends on estimating three parameter values from 18 data points. Activation noise (s in Equation 2), was estimated at 0.19, a value in line with other applications of this equation (e.g., Anderson, Bothell, Lebiere, & Matessa, 1998). The priming (associative activation) contributed by the probe digit was estimated at .83 units of activation. Finally, the prior-item activation (more specifically, the difference in prior item activation across the Fast and Slow conditions) was estimated at 1.1 units of activation. An Excel spreadsheet implementing the model and the two fits presented in this paper is at http://www.msu.edu/~ema/functionaldecay.

Fitting Peterson and Peterson (1959)

So far we have searched for and found evidence for decay in data on interference. Given our aim to integrate decay and interference functionally, we now turn to data on decay (Peterson & Peterson, 1959), and ask what role interference might play. The Brown-Peterson paradigm involves presenting a verbal item (e.g., a consonant trigram) and testing retention as a function of time. During the retention interval, verbal rehearsal is suppressed by a task like counting backwards. Figure 3 shows the data on recall accuracy from Peterson and Peterson (1959), Experiment 1. The x-axis shows retention interval and the y-axis shows proportion correct. The data show an even, negatively accelerating decline in accuracy with retention interval. This decline was interpreted to mean that maintenance of information in STM depended on active rehearsal, such that preventing rehearsal caused loss of information (Peterson & Peterson, 1959). Formal models of these data are not new (e.g., Baddeley, 1976, p. 130), but what is a novel integration is to explain these data on decay with the same processes that account for data on interference.

Our interpretation of the data in Figure 3 is that the current item (trigram) is represented in memory against a background of interfering items from previous trials. This kind of proactive interference has been demonstrated in a number Brown-Peterson studies (e.g., Dillon & Reid, 1969;

---

[3] Equation 2, termed the Chunk Choice Equation in Anderson and Lebiere (1998), plays a broader role in defining interference in our model than in ACT-R proper. In ACT-R, retrieval probability is a function both of Equation 2 and of a threshold parameter τ that specifies a minimum activation below which an item is invisible to the system. We assume no such threshold; the probability of retrieving an item is solely a function of that item's activation relative to the activation of distractors. We thus place greater emphasis on the role of interference from distracting information.

Keppel & Underwood, 1962; Wickens, Born, & Allen, 1963). In our model, this proactive interference plays the same role here as prior-item interference did in fitting the Waugh and Norman data. An element of this mental clutter can intrude when the system attempts to retrieve the target — and is more likely to, the more the target has decayed. Thus, relative activation is again the factor determining retrieval accuracy. Here, however, relative activation is a factor between trials only, whereas in the probe-digit model it was a factor between and within trials. The only other change to the model was to remove associative priming as a source of activation for the target, reflecting the absence of a specific probe in the Brown-Peterson paradigm.

As shown in Figure 3, the model again fits closely, with $RMSD$ = .027 and R^2 = .977. Fitting the model required estimating two parameters from six data points. Activation noise s (Equation 2) was estimated to be .34, which is again in the range used in other ACT-R models (Anderson et al., 1998). Prior item activation was estimated at -2.31 units.

Note that in fitting the Peterson and Peterson data we carried over the decay rate from the Waugh and Norman model (-.5). This illustrates the value of incorporating interference in a model of decay. A simple power-law decay model, without interference as a factor, is $P_i = a{+}bT_i^{\ d}$, with P_i the probability of retrieving item i, T_i the age of i, d the decay rate, and a and b free parameters. Fitting this model to the Peterson and Peterson data produces measures of fit $RMSD$ = .029 and R^2 = .973, and parameter values a = -19.8, b = 20.7, and d = -.14. Of particular interest is the decay rate, -.14. This deviates substantially from the value of -.5 that we carried over from the Waugh and Norman model, and from many ACT-R models before that (Anderson & Lebiere, 1998). Thus, although the *apparent* decay rate may differ from situation to situation, we propose that what varies is not the *architectural* decay rate but the background level of interference, which is situation-dependent and thus a more plausible source of variation. Importantly, this variable can also be estimated quantitatively, for example by counting the number of trials preceding the trial of interest (Keppel & Underwood, 1962).

In sum, if interference is indeed a primary mechanism of forgetting, then it would be odd if it played no role in forgetting in the Brown-Peterson paradigm. Our analysis suggests that decay by itself cannot cause forgetting — forgetting arises because decay takes place *relative to* background interference in memory.

Discussion

We propose that decay and interference are functionally related processes — decay of distractors mitigates the extent to which they interfere with the target. Playing out the consequences of this proposal makes functional sense of an empirical result that has lain largely dormant for a generation, absent the right theoretical framework in which to interpret it. We have also formalized the integration of decay and interference in another sense, by fitting the Waugh and Norman data set and its "opposite," the Peterson and Peterson data, with the same model. These model-fitting successes converge with our functional logic to argue that decay and interference must both operate in memory.

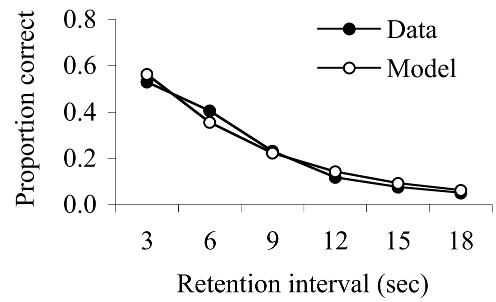

Figure 3: Data from Peterson and Peterson (1959), Experiment 1, and the fit of the functional decay model.

We began by hinting that if decay is important to the functionality of memory, then its effects should be found pervasively in behavioral data. Indeed, functional decay models similar to the one presented here have been applied to diverse domains. These include task switching (Altmann & Gray, 2002) and the time course of Stoop interference (Altmann & Davidson, 2001). Thus, the claim that decay is pervasive has some backing beyond our excursion here to the headwaters of the debate over forgetting mechanisms.

There are a number of caveats on the current work that will be important to address in the future. First, the Waugh and Norman interaction, though it seems visible in other probe digit data (Norman, 1966), needs to be replicated before we invest more in interpreting it. Second, our spreadsheet model needs to be implemented as a running simulation, to test whether we have missed important interactions among processes. Third, the model makes specific predictions about which distractor items should intrude in what proportions; later items should intrude more often, because they are more active. These predictions clearly need to be tested.

We should also note that the construct of relative activation underlying our model has other expressions in memory theory, such as the discrimination ratio (Baddeley & Hitch, 1993) and temporal distinctiveness (Neath, 1993). We would argue, however, that the grounding of the current model in ACT-R anchors it more directly to observable environmental processes. ACT-R is premised on the notion that memory is a mirror reflecting patterns of information need imposed by the environment. Thus, for example, retrieval frequency is a predictor of activation because it is also a predictor of impending need for that item. Consistent with the functional interpretation of decay, we favor a functional interpretation of memory generally, in which quantities like activation reflect the tasks that the memory system accomplishes for us.

Finally, we should emphasize that our claims about the importance of decay are not meant to conflict with the idea that interference is the dominant cause of loss of retrievable information in memory. Wherever they have been isolated, including in the Waugh and Norman data, the effects of decay are quite small (c.f., Reitman, 1974) compared to the effects of interference. Indeed, a small effect of decay is all that is functionally necessary to tilt retrieval probability toward the target, particularly when strategic memory processes like rehearsal are available for the system to manipulate target activation.

Acknowledgements

We thank the conference reviewers for valuable comments. CDS received support from ONR grant N00014-01-1-0321.

References

Altmann, E. M., & Davidson, D. J. (2001). An integrative approach to Stroop: Combining a language model and a unified cognitive theory. In *Proceedings of the twenty-third annual meeting of the Cognitive Science Society* (pp. 21-26). Hillsdale, NJ: Erlbaum.

Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science*, *13*, 27-33.

Anderson, J. R. (2000). *Learning and memory: An integrated approach* (2nd ed.). New York: Wiley.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory & Language, 38*, 341-380.

Anderson, J. R., & Lebiere, C. (Eds.). (1998). *The atomic components of thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, *96*, 703-719.

Ashcraft, M. H. (2002). *Cognition* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Baddeley, A. D. (1976). *The psychology of memory*. New York: Basic Books.

Baddeley, A. D. (1990). *Human memory: Theory and practice*. Boston: Allyn & Bacon.

Baddeley, A. D., & Hitch, G. (1993). The recency effect: Implicit learning with explicit retrieval? *Memory & Cognition*, *21*, 146-155.

Baddeley, A. D., & Scott, D. (1971). Short term forgetting in the absence of proactive interference. *Quarterly Journal of Experimental Psychology, 23*, 275-283.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10*, 12-21.

Dillon, R. F., & Reid, L. S. (1969). Short-term memory as a function of information processing during the retention interval. *Journal of Experimental Psychology,81*, 261-269.

Ekstrand, B. R. (1972). To sleep, perchance to dream. In C. P. Duncan & L. Sechrest & A. W. Melton (Eds.*), Human memory: Festschrift in honor of Benton J. Underwood* (pp. 59-82). New York: Appleton-Century-Crofts.

Estes,W.K.(1955).Statistical theory of spontaneous recovery and regression. *Psychological Review*, *62*, 145-154.

Fuster, J. M. (1995). *Memory in the cerebral cortex*. Cambridge, MA: MIT Press.

Galotti, K. M. (1999*). Cognitive psychology in and out of the laboratory* (2nd ed.). New York: Wadsworth.

Gillund, G., & Shiffrin, R (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*, 1-67.

Hintzman, D. L. (1978). *The psychology of learning and memory*. San Francisco: W. H. Freeman.

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

Keppel, G., Postman, L., & Zavortink, B. (1968). Studies of learning to learn: VIII. The influence of massive amounts of training upon the learning and retention of paired-associate lists. *Journal of Verbal Learning and Verbal Behavior*, *7*, 790-796.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, *1*, 153-161.

Landauer, T. K. (1975). Memory without organization: Properties of a model with random storage. *Cognitive Psychology*, *7*, 495-531.

Loess, H. & Waugh, N. C. (1967). Short-term memory and intertrial interval. *Journal of Verbal Learning and Verbal Behavior, 6*, 455-460.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476-490.

Luria, A. R. (1968). *The mind of a mnemonist*. New York: Basic Books.

Massaro, D. W. (1970). Perceptual processes and forgetting in memory tasks. *Psychological Review*, 77, 557-567.

McGeogh, J. A., & Irion, A. L. (1952). *The psychology of human learning*. New York: MacKay.

Murdock, B. B. (1992). Serial organization in a distributed memory model. In A. F. Healy (Ed.), *Essays in honor of William K. Estes* (1; pp. 201-225). Hillsdale, NJ: Erlbaum.

Nairne, J. S. (1983). Associative processing during rote rehearsal. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 3-20*.

Neath, I. (1993). Distinctiveness and serial position effects in recognition. *Memory & Cognition, 21*, 689-698.

Norman, D. A. (1966). Acquisition and retention in short-term memory. *Journal of Exp. Psychology, 71,* 369-381.

O'Reilly, R. C., & Munakata, Y. (2000*). Computational explorations in cognitive neuroscience*. Cambridge, MA: MIT Press.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*, 193-198.

Postman, L. (1971). Transfer, interference, and forgetting. In J. W. Kling & L. Riggs (Eds.), Psychology (3rd ed., pp. 1019-1132). New York: Holt, Rinehart, and Winston, Inc.

Reed, S. K. (2000). *Cognition: Theory and applications* (5th ed.). Belmont, CA: Wadsworth.

Reitman, J. S. (1974). Without surreptitious rehearsal, information in short-term memory decays. *Journal of Verbal Learning and Verbal Behavior*, *13*, 365-377.

Richman, H. B., Staszewski, J. J., & Simon, H. A. (1995). Simulation of expert memory using EPAM IV. *Psychological Review*, *102*, 305-330.

Turvey, M. T., Brick, P., & Osborn, J. Proactive interference in short-term memory as a function of prior-item retention interval. *Quarterly Journal of Experimental Psychology, 22,* 142-147.

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

Wickelgren, W. A. (1977). *Learning and memory*. Englewood Cliffs, NJ: Prentice Hall.

Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *2*, 440-445.