# How to Kill a Computer Model of Short Term Memory in Psychological Review

## Part I: Separate Fittings of Experimental Data

Eugen Tarnow, Ph.D.

18-11 Radburn Road

Fair Lawn, NJ 07410

etarnow@avabiz.com (e-mail)

## Abstract:

The short term memory field is inundated with computer models purporting to describe the workings of our mind. I show that several of them are not based in science: they use parameters that do not have intrinsic values but are simply separately fitted to each experimental situation, in effect treating each memory experiment as having different-species subjects.

Keywords: Free recall; short term memory; memory search, computer model of short term memory

## Introduction

Models of short term memory are created all the time but, perhaps due to excessive politeness and the inability of model creators to admit defeat, wrong ones are rarely killed off. At least not in public. A prominent example is the Atkinson-Shiffrin theory which is considered invalid by the contemporary research psychologists, including Dr. Shiffrin himself (see Tarnow, 2010). Nevertheless every year this theory gets more citations every year (about 160 citations per year) and few authors seem to realize that the theory is dead and buried.

Zombie theories are bad things to have around. They waste student time and new research keeps being created based off of, or relating to, incorrect theories. Zombie theories also lower the barrier for the creation of new theories – the original authors never have to worry about a public discourse that may be embarrassing to them. They also do the opposite: zombie theories can be used by referees to prevent newer theories from being published on account of the zombie theories already describing the same data.

The fight against zombie theories might seem particularly difficult when they are not easily identifiable statements but computer models. If model authors present plausible concepts together with lots of complicated math, who can argue with that? Who even has the time or inclination to check the work?

I will avoid the difficulty of arguing the conceptual building blocks and also about the overly complex math that sometimes go into them because it becomes an unwinnable argument with each side joining an obtuse arms race (for example, Estes, 1975, remarks that theories can become infinite sequences of theories, each with yet another twist to the concepts because the authors do not have a stopping rule defining when the authors will give up on the model. The authors act like gamblers who keep betting against the bank. Only eruditeness of their product are keeping their losses from becoming public).

Rather I am going to focus on one of the core concepts of many computer models of short term memory which are not falsifiable and thus not science: they consist of parameters without intrinsic values that are independently fit to each new experimental situation or even each variation of an experimental situation. In other words, if a model is fitting the 10-2 free recall data of Murdock (1962), where 10 indicates the number of items presented and 2 is the interval between item presentations, a completely new fit will be done for the 15-2 data in which an additional 5 items were presented. A scientific model would have one parameter change to accommodate the five additional items, not have all the parameters change at the same time. After all, experiment provides an average over subjects, whose brains the models purport to describe, and this average should not change appreciably between experiments – that is why journals insist on sample sizes, after all. The parameter-changing models imply that if the experimental situation changes ever so little, it is fair to consider that the subjects are another biological species. This issue is not an old issue: a 2013 issue of Psychological Review features an "improvement" of a model that fits the Murdock data with different parameters for each experimental situation (Lee and Pooley, 2013).

Strangely, or perhaps not so strangely, this problem has not been previously publicized. Rather the criticism of computer modeling has assumed that the modeling is done correctly but that there are higher level theoretical issues that need to be considered (Hanna, 1969) another example in which eruditeness trumps science. Estes got close but did not nail the issue when he wrote "on the evaluation of theories and models" that "intensive efforts should be directed toward finding additional basis [other than goodness of fit] for evaluating quantitative theories." He wrote that "the purpose of constructing models is not to describe data, which must be described before models can be applied to them, but rather to generate new classifications or categorizations of data… What we hope for primarily from models is that they will bring out relationship between experiments or sets of data that we would not otherwise have perceived… models will not force us to new categorizations of phenomena if they are so flexible that they can always be made to fit each new experiment by suitable choices of parameter values." Thus Estes was on the same track as the current author. But instead of questioning whether these theories are indeed scientific theories which can be falsifiable, he concluded that we need a complex "information-theoretic approach" and continued "we lack any methodology for statistical testing of the descriptive adequacy of alternative computer simulation models", and again eruditeness trumps science (that has since been fixed by Pitt et al, 2002). That a good fit is not enough was also pointed out by Roberts & Pashler (2000).

That this issue has not been publicized before does not mean authors do not know about it though they do not seem to appreciate the gravity. For example, X, upon prompting as to why her/his paper uses different parameters for each of the Murdock (1962) variations, writes:

> "Good question. A better model would either have them the same, or have a psychological theory (implemented formally in the model) of how they change as the memory task characteristics change. But,

> it's not our model, and our goal was to correct a more basic error. To make that point, we needed to replicate what Y assumed."

Z, when asked the same question about his coauthored publication writes

> "It would be quite desirable to have a model that fit data from all memory experiments with a single set of parameters, but we aren't there yet, as a field. There are a number of arguments why one might not expect the parameters to be the same across experiments. One might imagine that different individuals have different characteristic parameter settings, and these experiments were run on different sets of people. Probably more importantly, there is variability in the stimulus materials and methodological details that is not yet captured by the model framework, and so the behavioral variability engendered by these methodological differences must be accounted for by shifts in the best-fit parameters."

Another senior author in the field, Laming (2010) writes that in his model "these sets of data are modeled with, again, common parameter values applied to the different conditions within each experiment (with just one exception), so that the effects of the experimental manipulations are represented in the model, not in the choice of parameter values" and "but, of course, those different experiments are modeled with different parameter values." Thus while Laming fixes the parameters within experimental manipulations, he allows them "of course" to vary between different experimental situations. Laming (2010) notes that "parameters take different values for different lengths of lists" in Brown et al (2007) but fails to note that the same is true for Howard & Kahana (2002). A similar ambivalence to fitting parameters separately to each experimental situation is express by Sederberg et al (2008): "Our goal was to produce a reasonable qualitative description of dissociations between short and long-term recency. This approach has several advantages over attempting to find excellent quantitative fits by separately estimating parameters for each experimental condition being studied. First, it avoids the tendency to overfit the model to certain phenomena, which would would prevent the model from simultaneously capturing the trends observed across multiple experiments. In addition, it ensures that the model and not the parameters are doing the work in producing the pattern of dissociations observed across multiple experimental conditions."

And here is an editor commenting on this manuscript. Notice the last sentence, it seems to indicate a general lack of logical thinking in the field:

"a) We know it.  At the discussion of models at any conference these issues come up.  I'm surprised that you can't find more to cite in print.

b) You submitted a similar paper (though not hugely overlapping) to us a while back on Atkinson-Shiffrin.  And then I said -- we know it.  We know that models aren't theories -- but they help us build them."

Contrast this with Davelaar et al (2005) who states "even though methodological differences across experiments may warrant some variation in parameter values (which may be needed for precise quantitative fits), all simulations reported in this article were conducted with the same set of parameters to provide a stringent test of the model's ability to account for the qualitative patterns of the critical data."

## Computer Models of Short Term Memory that Can Be Dismissed

The computational model called SIMPLE by Brown, Neath and Chater, (2007) and improved by Lee & Pooley (2013) can be killed off since the parameters are different for each experiment they fit in both papers.

The Context Maintenance & Retrieval (CMR) model (Polyn, Norman, & Kahana, 2009 can be dismissed as well. The authors state that they use 11-14 parameters that are fitted to 93 data points of either the Murdock (1962) dataset or the Murdock & Okada (1970) dataset. Thus instead of fitting the two experiments with the same parameters the authors choose to believe the experimental subjects are different species.

The creators of the OSCillator-based Associative Recall (OSCAR) model (Brown, Preece & Hulme, 2000) show how their model can simulate various experimental situations. In each of these experimental situations the parameters are varied, thus the theory can be dismissed.

Laming (2010) uses different parameters for different experiments and can thus also be dismissed. Note that he is a little more careful than the others – he uses the same parameters within the same set of experimental variations.

The Start-End Model (SEM) (Henson, 1998) is more carefully constructed. Some of the parameters are kept fixed between different experiments and others that are allowed to vary are necessary to "accommodate differences between experiments beyond present concerns (e.g, the stimuli, presentation modality, or method of recall employed)." Thus at this point we do not dismiss this theory.

## References


Brown GDA, Neath I, Chater N (2007) A temporal ratio model of memory. Psychol Rev 114(3):539–576

Brown, G. D. Preece, T. Hulme, C. (2000), Oscillator-Based Memory for Serial Order, Psychological Review 107 (1), 127-181.

WK Estes (1975) "Some Targets for Mathematical Psychology" JOURNAL OF MATHEMATICAL PSYCHOLOGY 12,263-282

Hanna JF (1969) "Some information measures for testing stochastic models" Journal of Mathematical Psychology Volume 6, Issue 2, Pages 294–311

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. Journal of Mathematical Psychology, 46(3), 269-299.

Laming, D. (2010). Serial position curves in free recall. Psychological review, 117(1), 93.

Lee, M. D., & Pooley, J. P. (2013). Correcting the SIMPLE model of free recall. Psychological review, 120(1), 293.

Murdock, B. B., & Okada, R. (1970). Interresponse times in single-trial free recall. *Journal of Experimental Psychology*, *86*(2), 263.

Pitt, M.A., Myung, I.J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. Psychological Review, 109, 427–491.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological review*, *116*(1), 129.

Roberts S, Pashler H (2000) How Persuasive is a Good Fit? A Comment on Theory Testing Psychological Review Vol. 107, No. 2, 358-367

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. Psychological review, 115(4), 893.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: empirical and computational investigations of recency effects. Psychological Review, 112(1), 3.

Henson RNA (1998) Short-Term Memory for Serial Order: The Start-End Model, Cognitive Psychology 36 (2) 73-137.

Murdock Jr., Bennet B. (1962). The serial position effect of free recall. Journal of Experimental Psychology. Vol 64(5): 482-488.

Tarnow, E. (2010). Why The Atkinson-Shiffrin Model Was Wrong From The Beginning. WebmedCentral NEUROLOGY 2010;1(10):WMC001021