Why the Atkinson-Shiffrin model is wrong

Eugen Tarnow, Ph.D.
18-11 Radburn Road
Fair Lawn, NJ 07410
etarnow@avabiz.com (e-mail)

Abstract:

The Atkinson-Shiffrin (1968) model, a standard model of short term memory cited over three thousand times, mimics the characteristic shape of the free recall curves from Murdock (1962). However, I note that it is not a theoretically coherent explanation and that it does not fit any other relationships present in the same Murdock data. As a result, future theorists are challenged with defining the buffer concept properly, with defining the long term store properly, and with correctly predicting new relationships found in the Murdock data that directly probe various theoretical concepts.

Introduction

The free recall experiments of Murdock (1962) exhibit a characteristic bowing effect: both initial and final items are better recalled than items in the middle. These experimental data were modeled by Atkinson & Shiffrin (1968), (this paper is from now on referred to as the A&S model). Their theory is widely cited (more than 3000 citations according to Google Scholar) and hailed as “the next decisive advance in human learning and memory [after Ebbinghaus]” (see p.1, Izawa, 1999. Izawa was a former graduate student in the same Stanford department as Atkinson and Shiffrin). Despite the citation count, this paper is apparently rarely read (Cowan, Rouder and Stader, 2000) and this presents an opportunity to revisit the model and reevaluate it.

In this paper we will see that the A&S model succeeds in mimicking the characteristic primacy effect seen experimentally by using four concepts: the Rehearsal Buffer (RB) “control process”, (A&S, p. 113), the RB Starting Empty, the Randomly Emptied RB and the Long Term Store (LTS). In this paper I will point out that each of these concepts are problematic and together they do not form a coherent theory. As a result, A & S does not quantitatively fit the free recall curves and does not fit any of the other relationships present in the same Murdock (1962) data that it can be fitted to.

I ask that the reader for a few minutes sets aside the authority lent to the A&S paper by its overwhelming citation index and allows her or his mind to be open to a new interpretation of the same.

Theoretical problems with the four A&S concepts needed to fit the primacy effect

When constructing a theory, all elements of this theory has to be well defined. The first problem with the A&S model is the Rehearsal Buffer (RB): it is not well defined. A&S define the RB in two incompatible ways. In definition 1 (my term), RB is process controlled by the subject. In definition 2 RB is a black box not apparent to the subjects (contradicting definition 1) but defined by bootstrapping; <u>because it works</u> when fitting the experimental data.

Let me give you examples of definition 1. RB is defined as a "control process" of the Short Term Store (STS) (A&S, p. 113), it is a control "process entirely under the control of the subject" (p. 133) and when a stimulus enters STS "the subject must decide whether or not to place the new item in the rehearsal buffer. There are a number of reasons why every incoming item may not be placed in the buffer" (A&S, p. 128). RB is to be distinguished from STS which is a permanent feature of memory and cannot be controlled or changed (p. 90). "The hypothesis of an ordered fixed-size buffer is given support by the subjects' reports and the authors' observations while acting as subjects." (p. 127).

Definition 2 occurs to assuage the readers of A&S: "The reader is not asked, however, to take our word on this matter; the analysis of the results will provide the strongest support for the hypothesis" (p. 127). In other words, there is no reason to keep definition 1 but move to definition 2 in which RB is a black box that exists because <u>fitting the experimental data</u> tells us that RB exists. A&S continues "while this paper consistently considers a fixed-size short-term buffer as a rehearsal strategy of the subject [definition 1], it is possible to apply a fixed-size model of a similar kind to the structure of the short term system as a whole, that is, to consider a short-term buffer as a permanent feature of memory [definition 2]" (p. 114).

Here I stress that if a theory is not logically well defined it is not a theory, even if it fits a subset of experimental data. An ill-defined non-theory can survive criticism longer (critics can always to be told they misunderstood the theory- definition 2 rules, not definition 1, or vice versa) as long as this is not discovered. Indeed, this can be applied to Izawa who refers to A&S in glowing terms, but writes that "I did not completely agree with the rehearsal mechanisms/processes postulated in the model (they did not quite correspond to my personal experiences in the learning of several languages over a number of years [she did not agree with definition 1]" (p. xiii, Izawa, 1999). She must have misunderstood the definition of RB – she should have used definition 2! Oberauer and Lewandowsky (2008) criticize other theories with rehearsal buffers along similar lines: "Any decay model can explain the absence of forgetting, should it arise in the data, by appealing to compensatory rehearsal, and at the same time, it can explain the presence of forgetting, should it occur, by suggesting that rehearsal was withheld or impossible."

The Starting Empty RB concept, the second problem, is one in which memory items in the beginning of a list gets special access to RB. I will show below that A&S is forced to assume that RB is empty to start with so that the first list items are able to spend more time in RB than subsequent items which leads to the primacy effect. If RB is a control process that we are in control over this might make sense (we could empty RB at will). But if RB is considered to be an important part of STS, why should we start with an empty RB? If RB is an important part of memory would we not expect it to always be filled with something?

Just how the items in RB are replaced is another problem. The content of RB cannot be replaced using what would be the most obvious: first in – first out – that leads to a constant recall probability of 1 for the last four items of the Murdock free recall data. Rather as I will show below A&S is forced to assume that the item that leaves RB is randomly chosen from the items in RB in order to fit the recency effect. Since RB is supposed to be what allows for purposeful quick calculations, this seems aberrant. Knowing what we know about the nervous system, there might not be that much space for randomness. Fluctuations that occur because of neurons that either fire or not become smaller as the purposeful firing rate increases as a subject is reading or hearing an item.

The last problem is the presence of LTS, the long term store, in the <u>short term</u> free recall experiments. To make the appropriate number of initial items stay around long enough to remain in memory with a higher probability than subsequent items, the Starting Empty RB has to be combined with the LTS. Between each item presentation there is a given probability of items in RB of being transferred into LTS. Once in LTS the items can be recalled by the subject after they leave RB presumably indefinitely. Biochemical evidence strongly suggest that long term memory is created on the time scale of one hour (Kandel, 2001), not on the less than a minute time scale of typical free recall experiments. The free recall models sometimes introduces a decay time for LTS (A&S) but then LTS is not really LTS anymore and the A&S theory is yet again ill-defined. In fact, A&S states: “it is important not to confuse our theoretical constructs STS and LTS … with the terms Short Term Memory … and Long Term Memory … These latter have come to take on an operational definition in the literature.” Operational definitions is exactly what a theory should should consist of.

There is also another fundamental problem with the A&S theory: the large number of concepts introduced in order to fit the bowing effect: four concepts to fit two parts of a curve (all four are needed as I will show below). A&S does not state that those four concepts were essentially akin to “fitting constants” but they state: “a large number of results may be handled parsimoniously [with their theory]” (p. 91).

When a correct theory is fit to some part of the data it is then able to predict other data that were not fitted. Bending concepts into a pretzel as is done by A&S typically does not work, a certain elegance seems to be needed in order to describe nature properly. Thus we will find that after A&S is fitted to the Murdock (1962) recall curves it incorrectly predict other relationships in the same Murdock (1962) data.

A One Parameter A&S Model – Four Concepts Needed for Fit

I constructed an A&S model to illustrate the various concepts in the model. In this model an item can be either in RB, in the LTS or lost. RB can handle up to four items. As a new item is presented the probability that it goes into RB is 1 and the probability that each of the items in working memory leave working memory is an equal ¼ once RB is full and 0 if RB is not full. While in RB the probability per unit time that an item is transferred into LTS is α for small times. Before the list is presented, RB is unfilled.

This model is a simplification of A&S in which four parameters were used: the buffer size (set to 4 by the data in Cowan (2000), assuming that the capacity limitation of RB is the same as of STS), the probability of entering RB (effectively set to 1 in my model), the decay rate of information from LTS (removed in my model). This simplification will make the failures of the A&S theory more transparent without losing any essential ingredient.

The Murdock data consists of six experiments labeled 10-2, 15-2, 20-2, 20-1, 30-1 and 40-1. The first number refers to the number of list words, the second is the number of seconds allowed to pass between list items (presentation rates of 0.5/s and 1/s, respectively). The fitted model result for the 10-2 data is shown in Fig. 1A (the black diamonds are the experimental data, the model prediction are the white squares). The shape of the recall curve is modeled nicely. The same model can fit the other five sets of data if one allows for a change in the probability of entering RB (it becomes smaller as the number of words grow).

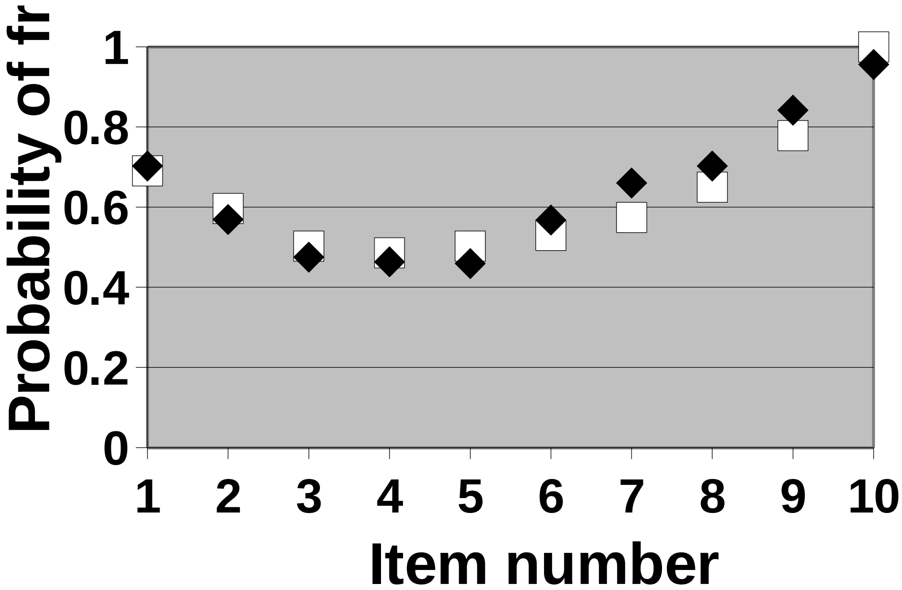

Fig. 1a. The probability of free recall from the Murdock 10-2 experiment (black diamonds) compared with the model fit using the parameter α=0.057/second (white squares). R squared is 0.994.

LTS, Randomly Emptied RB, Starting Empty RB, and RB are all needed for the model to fit the data. I will now remove one concept at a time to illustrate this. First, we remove LTS and obtain the results in Fig. 1b. The primacy effect is gone. Second, we allow RB to be emptied using a different algorithm than random, say first-in first-out and we obtain the results in Fig. 1c. There is no bowing effect and the final four items are all in RB and all are recalled with probability of 1. Third we remove the Starting Empty RB concept and obtain the results in Fig. 1d in which there is no bowing effect. Fourth, changing the size of RB to, say, 1, the results in Fig. 1e show that there is no bowing and the only item that is different from the others is the last one.

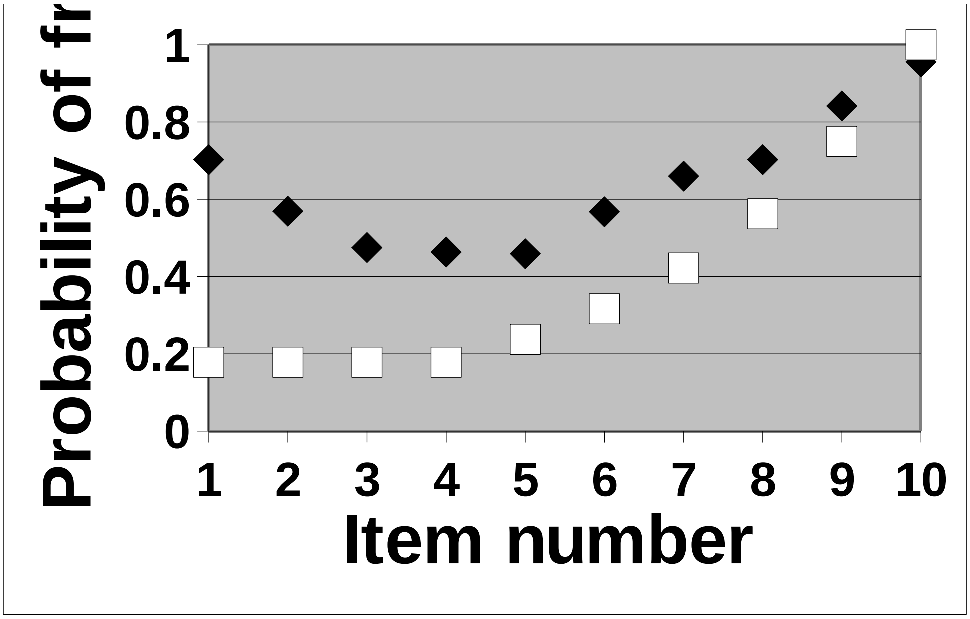

Fig. 1b. The probability of free recall from the Murdock 10-2 experiment compared with the model without the long term memory store (α=0). There is no bowing effect and the four first items have the same recall probability (the probability that they remain in RB).

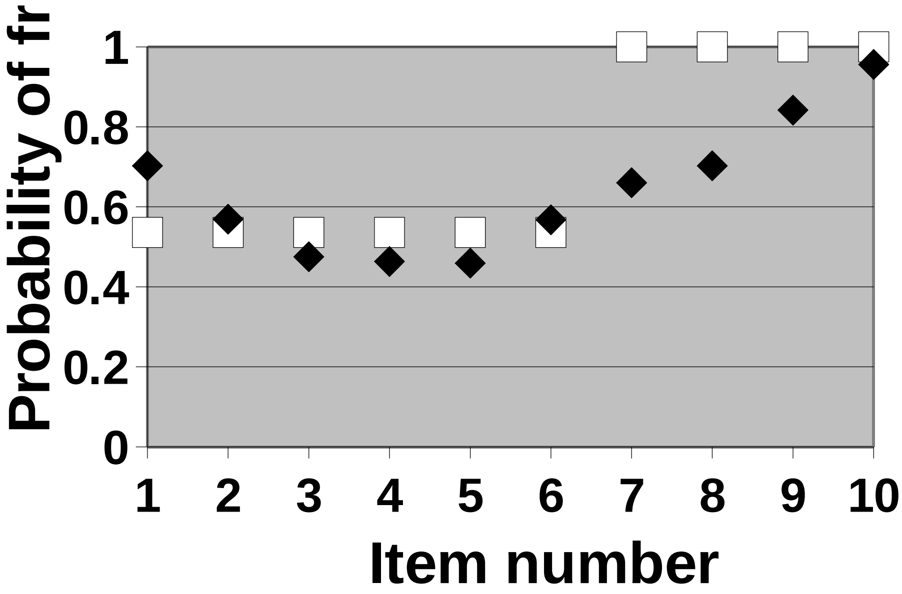

Fig. 1c. The probability of free recall from the Murdock 10-2 experiment compared with the model fit using first in first out and the parameter α=0.067/second. There is no bowing effect and the final items are all to be found in RB.

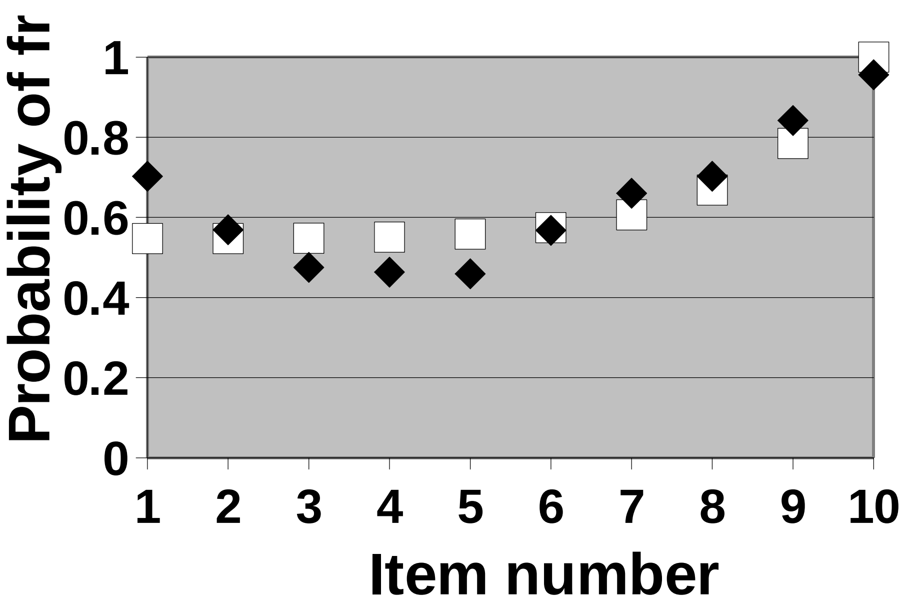

Fig. 1d. The probability of free recall from the Murdock 10-2 experiment compared with the model fit using an initially filled RB and the parameter α=0.065/second. There is no discernible bowing effect because the time the first four items spends in RB is roughly equal if the number of items is larger than the size of RB.

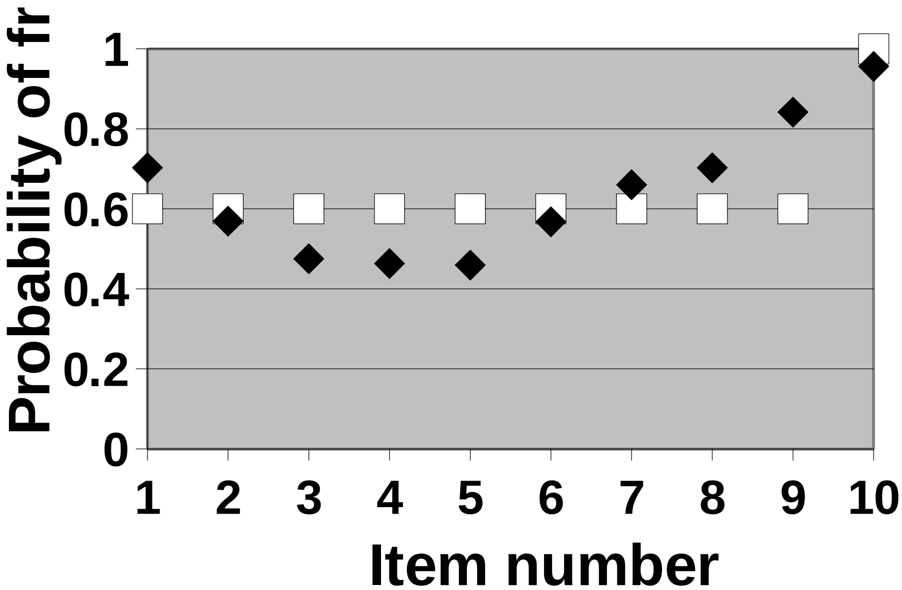

Fig. 1e. The probability of free recall from the Murdock 10-2 experiment compared with the model fit using RB size=1 the parameter α=0.3/second. There is no bowing. All items but the last one has the same recall probability.

The simplified A&S model presented here seemingly has only one fitting constant (which depends on the presentation and the number of items). But I believe that the number of fitting constants in some sense also have to include the four relatively unappealing concepts, LTS, Randomly Emptied RB, Starting Empty RB, and RB. As we have seen they are all needed for the bowing effect. Whether the number of fitting constants is one or five I will now explore – can the theory predict the other relationships in the Murdock data?

Incorrect Predictions of Other Relationships In The Same Murdock (1962) Data

Within the narrow confines of the A&S model fitting the free recall data of Murdock we can go a step further: beyond the fit of the free recall data, what are the other predictions of the model?

I define the "primacy strength" to be the ratio of the probability of recall of the first list item divided it by the lowest probability of recall of an intermediate item. As we can see from Table 1 the experiment shows a sizable variation with presentation rate, theory does not.

	“10-2” data	“15-2” data	“20-2” data	“20-1” data	“30-1” data	“40-1” data	0.5 second presentation rate average	1 second presentation rate average
Experiment	1.53	1.71	2.4	3.32	2.94	3.49	1.9	3.3
A&S theory	1.65	1.78	1.81	1.8	1.83	1.5	1.7	1.7

Table 1. Primacy strength.

To test the Randomly Emptied RB concept, I concentrate on the last items. The probability that the last item is in RB is 1. If the subject recalls the last item, since the last item is already in RB this recall should not interfere with any of the other items in RB or LTS and the total sum of recall probabilities should remain the highest. If, on the other hand, the subject recalls the second to last list item, there will be some interference since the probability that the second to last item was in RB was not 1 but 0.75 thus changing the content of RB. Thus the total sum of recall probabilities should be lower and if the item recalled first is the third to last item the sum should be even smaller. The experimental result is shown in Fig. 2 and shows the opposite trend: the total number of recalled items is worst if the first item recalled is the last list item (others have noted a similar trend in an experiment by Craik, 1969, see also Cowan et al, 2002). Thus the total recall probability is incorrectly predicted by A&S.

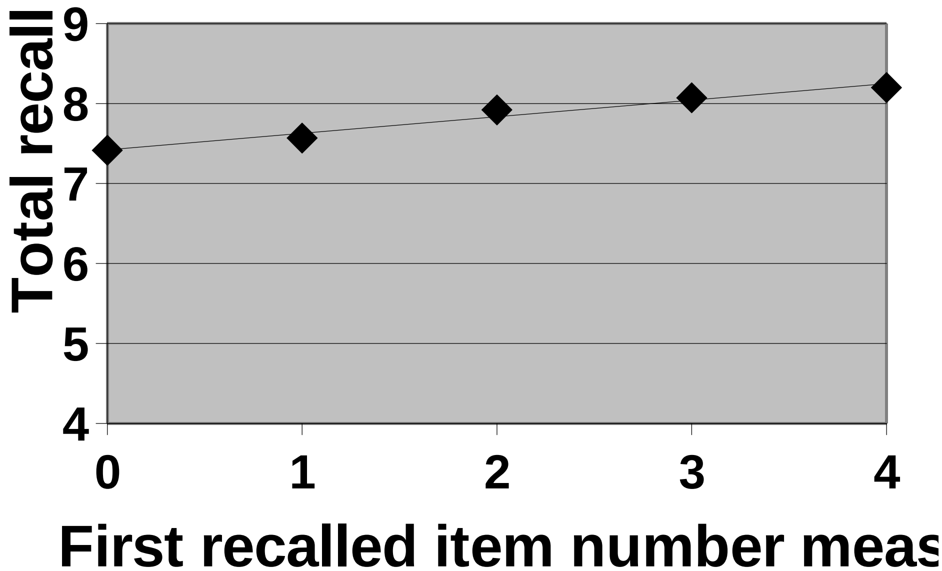

Fig. 2. Total number of items recalled as a function of the item number of the first recalled item measured from the last item (at 0) averaged over all Murdock data. Randomly emptied RB predicts incorrectly that the total recall from recall starting with the last item (0 in this case) should be the highest. It is the lowest.

The Starting Empty RB concepts predicts that the three differences in recall probabilities: the first and second, second and third and third and forth items are proportional to α with proportionality constants $1-(3/4)^{(N-4)}$, $1-(3/4)^{(N-4)}$ and $(3/4)^{(N-4)}(1-(3/4)^{(N-4)})$ where N is the number of items in the list. Various combinations of these differences have similar properties. These proportionality constants are <u>ratio rules</u> that experimental measurements should fulfill if the theory is correct. In particular, their ratios are independent of α and only dependent on the Starting Empty RB concept.

To make a complete list of A&S fitting concepts I also include the "channel" possibility i.e. whether a single item in RB has a probability α to go into LTS or whether it has a probability 4α and whether two items in RB have the probability α or 2α to go into LTS (these assumptions were contemplated by A&S and the latter assumption was used in the case of free recall by Raaijmakers & Shiffrin (1981)).

The "channel" concept gives rise to different ratio rules as shown in Table 2. The two cases are distinguished by the ratio of $\frac{(P1-P2)}{(P2-P3)}$. If each item gets a single channel in an otherwise empty RB the ratio should be 1, otherwise it should be 1.5, <u>independent of the number of items in the list</u>.

	One item, one channel	One item, 4 channels, two items 2 channels each, 3 items 4/3 channels each
$\frac{(P1-P2)}{\alpha}$	$1-(3/4)^{(N-4)}$	$2(1-(3/4)^{(N-4)})$
$\frac{(P2-P3)}{\alpha}$	$1-(3/4)^{(N-4)}$	$4/3(1-(3/4)^{(N-4)})$
$\frac{(P3-P4)}{\alpha}$	$(3/4)^{(N-4)}(1-(3/4)^{(N-4)})$	$(3/4)^{(N-4)}(1-(3/4)^{(N-4)})$
$\frac{(P1-P2)}{(P2-P3)}$	1	1.5

Table 2. Theoretical values of difference constants in two types of Starting Empty RB: the second column shows the case for which each item is treated as a single item even if RB is partially filled, the third column shows the case for which each item fills up as much of RB as is possible. The last row shows the ratio that experiment should use to distinguish between the two cases (see below). It is independent of the total number of items in the list.

In Table 3 is shown the constants calculated from the Murdock (1962) data. I have eliminated the α value dependencies by only reporting ratios in which the α values cancel out. It seems that the experimental data on the first ratio rules out the A&S theory.

	"10-2" data	"15-2" data	"20-1" data	"20-2" data	"30-1" data	"40-1" data	Average
$\frac{(P1-P2)}{(P2-P3)}$	1.42 (1 or 1.5)	3.38 (1 or 1.5)	2.39 (1 or 1.5)	3.15 (1 or 1.5)	2.16 (1 or 1.5)	1.52 (1 or 1.5)	2.34 (1 or 1.5)
$\frac{(P1-P3)}{(P1-P4)}$	0.95 (0.92)	0.83 (0.99)					

Table 3. Values of difference constants calculated from Murdock (1962) data. The theoretical values are shown in parenthesis (single channel or multichannel model).

Let's consider another experimental relationship that we know: the differences between same values for the Murdock 20-1 and 20-2 experiments. The experimental differences are shown in Fig. 3 (a) and the theoretical prediction is shown in Fig. 3(b). While there is good overall agreement the theoretical curve shows a discontinuity at low item numbers when RB is partially filled which does not seem to be present in the experimental data. This continuity becomes more pronounced the larger the total number of items in the experiment (Fix. 3(c) shows the corresponding theoretical prediction for a total of 10 items). A lower noise experiment would be able to confirm or rule out the theory and shine further light on the concept of the Starting Empty RB.

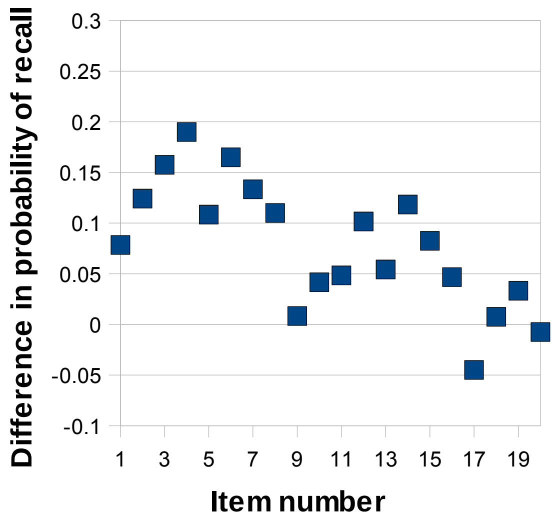

Fig. 3(a). Difference in recall probabilities for same item numbers between the 20-2 and 20-1 experiments.

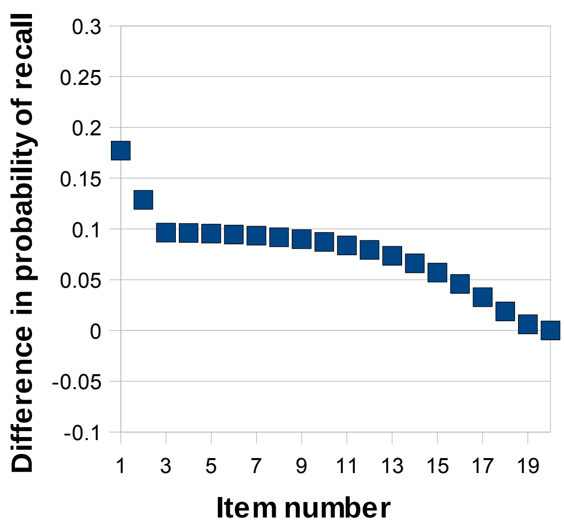

Fig. 3(b). Theoretical differences in recall probabilities for same item numbers between 20-2 and 20-1 data.

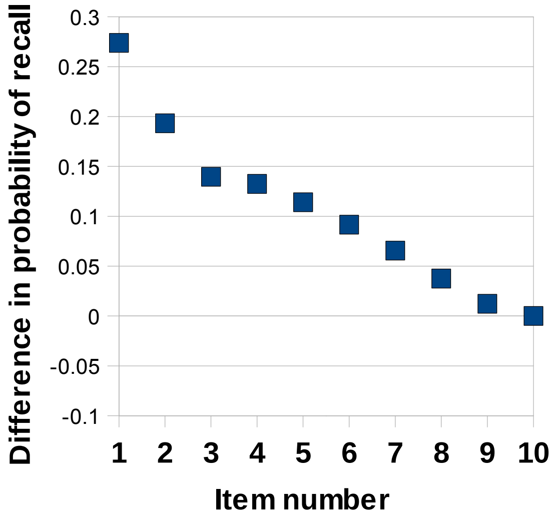

Fig. 3 c. Theoretical differences in recall probabilities for same item numbers between 10-2 and 10-1 data using α =0.098/second for both sets of data. Note that the discontinuity for low item numbers is less pronounced than for the 20-2/20-1 data.

What about RB? It is thought that there is a built-in storage limitation of four items for RB (Cowan, 2000). It is a mystery as to why there are four items and a mystery why these four items should be treated as equals. I am not sure how strong the case is for the existence of a four item RB in the free recall data of Murdock (1962). It is hard to "see" the number four in the experimental recall data. In particular, if we plot the recall data as a function of the logarithm of the time of recall there is a straight line and the recent items do not stand out above the "horizontal asymptote" (see Tarnow, 2009b) as they do in the conventional plot of probability versus item number. Second, if we calculate the sum of all items remembered we get a result between 6 and 8.5, not 4 (see Table 4). The only way four shows up is with the questionable concepts of the A&S model.

SUMMARY

I have shown that the A&S theory is problematic in many ways . First, some of the concepts such as the RB and the LTS are not well defined and therefore A&S does not form a theory. Second, four concepts which are all somewhat problematic are needed to fit the shapes of two parts of the free recall curves. As one might expect, when too many concepts are needed in a fit, the theory only interpolates the data for which it was fitted for, it cannot correctly predict other properties. The latter include the failure to predict the strong presentation rate dependency of the primacy strength, the failure to predict the correct dependency of the total recall probability as a function of the initially remembered item, the apparent failure of the experimental data to adhere to the theory's ratio rules and the apparent failure of the experimental data to show a discontinuity for the probability differences of the first items in experiments with different presentation rates.

Future theorists are thus challenged with defining the buffer concept properly – is it a control process or is it a black box and does it even exist; with defining the long term store properly (how can it be present in a short term memory experiment), and with correctly predicting new relationships found in the Murdock data that probe various theoretical concepts.

Experiment	Total items remembered	Assuming Partially Filled RB / LTS: in LTS
Murdock 10-2	6.4	2.4
Murdock 20-1	6.83	2.83
Murdock 15-2	8.19	4.19
Murdock 30-1	8.51	4.51
Murdock 20-2	8.38	4.38
Murdock 40-1	8.12	4.12
Rubin, Hinton, Wenzel		

Table 4. Total number of items stored in memory in the various Murdock experiments.

References

Atkinson, R.C. and Shiffrin, R.M. (1968) Human memory: a proposed system and its control processes. In: Spence K.W. and Spence J.T. (Eds.), The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 2. Academic Press, New York, pp. 89–195.

Cowan N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity.- Behavioral and Brain Sciences 24(1):87-114.

Cowan N, Rouder JN, Stadler MA, (2000) “Conjuring a Work From the Dream Time of Cognitive Psychology” The American Journal of Psychology, Vol. 113, No. 4 (Winter, 2000), pp. 639-648

Cowan N. Saults JS, Elliott EM, Moreno MV (2002). Deconfounding serial recall. Journal of Memory and Language 46, 153-177.

Craik, F. I. M. (1969). Modality effects in short-term storage. Journal of Verbal Learning and Verbal Behavior, 8, 658–664.

Izawa C (1999) On Human Memory: Evolution, Progress, and Reflections on the 30th anniversary of the Atkinson-Shiffrin Model. Lawrence Erlbaum Associates, Mahwah, New Jersey, USA.

Kandel, E.R. (2001). The Molecular Biology of Memory Storage: A Dialogue Between Genes and Synapses. Science, pp. 1030-1038.

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

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? Psychological Review, 115(3), 544-576.

Raaijmakers, J.G.W. & Shiffrin, R.M. (1981). Order effects in recall. In J.B. Long and A.D. Baddeley (Eds), Attention and Performance IX (Pp. 403-415). Hillsdale, N.J.: Erlbaum.

Rubin DC, Hinton S, Wenzel A.(1999). The Precise Time Course of Retention. Journal of Experimental Psychology: Learning, Memory and Cognition, Vols. 25:1161-1176.

Shiffrin, R,M. (1968). Search and retrieval process in long-term memory. Technical Report No. 137, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Shiffrin, R.M. (1970). Memory search. In D.A. Norman (Ed.), Models of human memory. (Pp. 375-447). New

York: Academic Press.

E. Tarnow. (2008). Response probability and response time: a straight line, the Tagging/Retagging interpretation of short term memory, an operational definition of meaningfulness and short term memory time decay and search time. Cognitive Neurodynamics, 2 (4) p. 347-353.

E Tarnow. (2009) Short term memory may be the depletion of the readily releasable pool of presynaptic neurotransmitter vesicles of a metastable long term memory trace pattern. Cognitive Neurodynamics 3:3 263-269.

E Tarnow. (2009b) Short term memory decays and high presentation rates hurry this decay: The Murdock free recall experiments interpreted in the Tagging/Retagging model. Accepted to Cognitive Neurodynamics.