Locating Working Memory in Word Free Recall and Estimating its Capacity at Less Than 3.1 +/- 0.2 Words

Eugen Tarnow

18-11 Radburn Road

Fair Lawn, NJ 07410

etarnow@avabiz.com

Abstract:

The evidence for limited capacity is strong yet there is a controversy as to whether this limited capacity can be seen in word free recall data. After five decades since the first word free recall experiments, I present clear direct evidence for a dual store model in word free recall. I show that there is a discontinuity in the item probability during the first recall, the recall that is the most likely to originate in working memory. This discontinuity indicates a separation between the items that are initially in working memory and the items which are retrieved from elsewhere at a later time, presumably via an association with one or more working memory items. The items that are not initially in working memory are those that are the least likely to be recalled overall. Assuming that the first recall is an accurate statistical representation of working memory, and that the probability of an item in working memory cannot exceed the overall probability of recalling that item, I arrive at an upper limit on the working memory capacity of 3.1 +/- 0.2 words.

Keywords: Free recall; working memory; short term memory; memory search

Introduction

NOTE TO REFEREES: ANY SPECIFIC SUGGESTIONS FOR FURTHER REFERENCES WILL BE TAKEN SERIOUSLY, AS WILL ANY SPECIFIC ARGUMENTS AS TO WHY MY ARGUMENTS ARE OLD OR WRONG OR INAPPROPRIATE IN ANY OTHER WAY.

Free recall stands out as one of the great unsolved mysteries of modern psychology. Items in a list are displayed or read to subjects who are then asked to retrieve the items. It is one of the simplest ways to probe short term memory but the results (Murdock, 1960; Murdock, 1962; Murdock, 1975) have defied explanation. Theories of short term memory (for example, Atkinson and Shiffrin, 1968) have come and gone because they did not reflect the details of the free recall data (Tarnow, 2010). And in particular, if one may speak about a consensus in memory psychology, that short-term memory has a limited capacity of 3-6 items (Cowan, 2001), there is no consensus as to just where this limit can be seen in free recall data. ("I think it's because you have to get the conditions just right to see fixed capacity." N. Cowan, personal communication, September 4, 2012).

Existing theories of short term memory are multi-store models or multi-state models of a single store of which working memory is either a store or a state. If there are truly several stores or states, convincing supporting evidence would involve one or more discontinuities in the experimental data. If a working memory store or state is emptied first during free recall, then there should be some kind of discontinuity when the next short term memory store is emptied and another discontinuity when the following store is emptied and so on.

But the search for discontinuities has been elusive. For example, there is no discontinuity in word free recall response times (Tarnow, 2013), nor in errors (Tarnow, 2014) and latency distributions are the same for all but the first recall (Laming, 1999). The lack of discontinuities extend beyond word free recall to counting of blocks (Balakrishnanl and Ashby (1992) did not find any discontinuity in reaction time distributions in an experiment asking for enumeration of colored blocks when the number of blocks increases from 1 to 8) and recognition time (McElree (2006) showed that there is no discontinuity in item recognition time beyond the first item).

The lack of discontinuities in word free recall is a significant issue that needs to be understood in the light of the evidence for a limited capacity model being very strong (Cowan, 2001). Hintzman (2011) considers free recall just too complex to be useful and, similarly, the Baddeley (2012) review of working memory hardly mentions free recall and does not mention that there are no discovered discontinuities as a result of a limited capacity working memory.

Method

This article makes use of the Murdock (1962) data set (downloaded from the Computational Memory Lab at the University of Pennsylvania (http://memory.psych.upenn.edu/DataArchive). In Table 1 is summarized the experimental processes which generated the data sets used in this paper.

Work	*Item types*	*List length and presentation interval*	*Recall interval*	*Subjects*	*Item presentation mode*
Murdock (1962)	*Selection from 4000 most common English words*	*10,15,20 words in a list each word presented every 2 seconds* *20 ,30, and 40 words in a list, each word presented once a second*	1.5 minutes	103 undergraduates	Verbal

Table 1. Information about experiments included in the study

Results

In Figure 1 left panel is shown the item recall probability versus serial position from the 10-2 experiment of Murdock (1962). In the right panel is shown the same data for the initial recall only. Both curves show the famous bowing effect - intermediate words are remembered less well than initial or final words. But the curves are also different. In fact, the initial recall distribution has some word recall probabilities amplified compared to the overall recall distribution and some recall probabilities eliminated (three of the words in the initial recall distribution have close to 0 recalls). Thus, instead of stochastically independent events which would mean a proportional relation:

(initial recall of item i) =const* (overall recall of item i)

we have:

(initial recall of item i)= -discontinuity + (overall recall of item i) * amplifier.

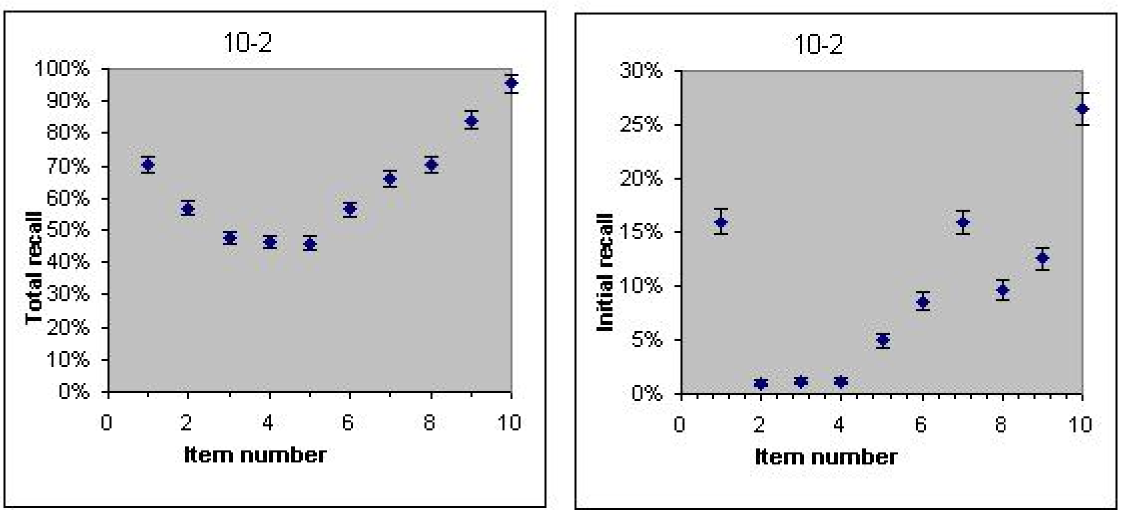

Fig. 1. The left panel shows the famous bowed curve of total recall versus word number. The right panel shows the bowed curve of initial recall versus word number. The error bars in each direction are the

standard deviation of a Poisson distribution (no systematic errors were considered). Experimental data from Murdock (1962).

This relationship is shown in the upper left corner graph of Figure 2. Instead of a straight line continuing through the origin, it is displaced away from the origin by a discontinuity. The other graphs of Figure 2 displays the remainder of the Murdock data (which covers two presentation rates - one per two seconds and one per second - and five word list lengths – 10, 15, 20 for the two second presentation rate and 20, 30 and 40 for the one second presentation rate). All of the relationships are displaced away from the origin by a discontinuity.

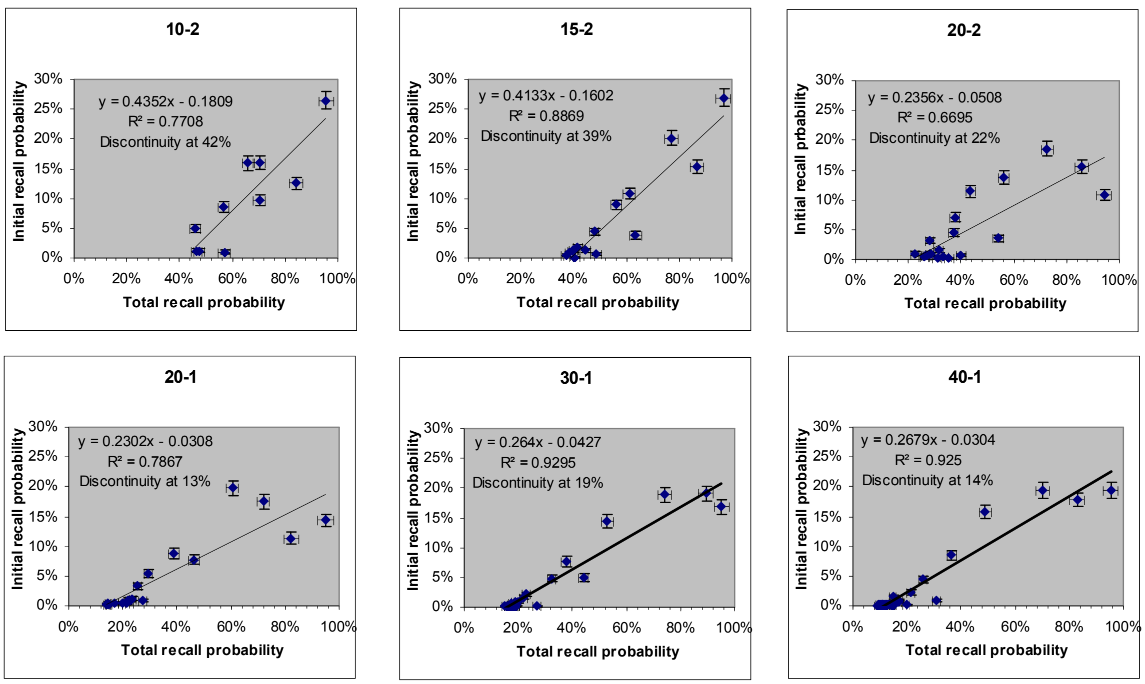

Figure 2. Number of initial recalls for each word versus the total number of recalls for that word. The six panels correspond to the six Murdock (1962) experiments labeled on top as M-N. M is the number of words in the list and N is the number of seconds between word presentations. The error bars in each direction are the standard deviations assuming a Poisson distribution (no systematic errors were considered).

The discontinuities eliminate low probability words from the initial recall. Figure 3 shows that the discontinuity in each case is the total recall probability of the word least recalled.

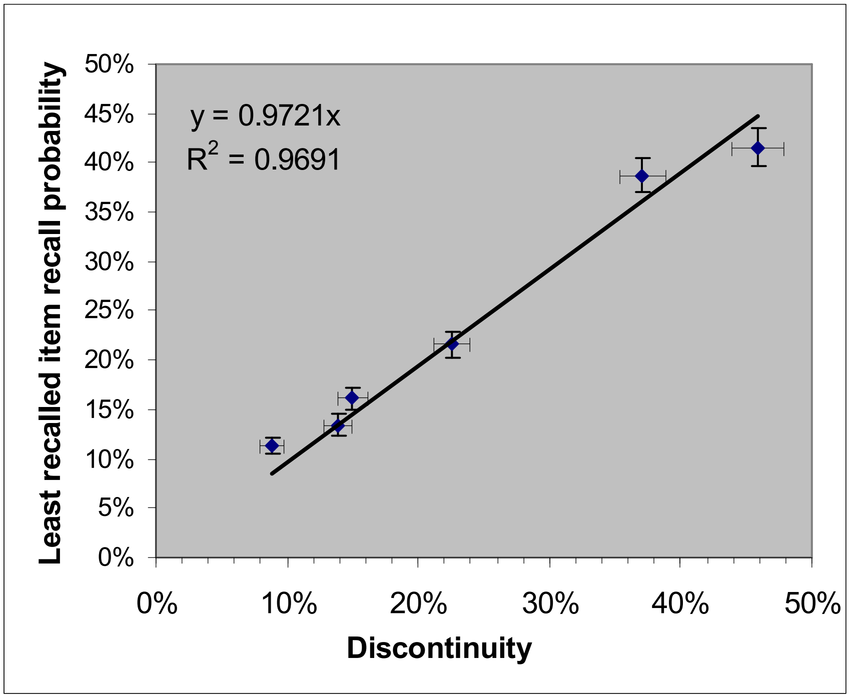

Figure 3. *Recall probability of the least recalled word versus discontinuity from the curves in Figure 2. Note the relationship* ***Least recalled item recall probability=Discontinuity.*** *The error bars in each direction are the standard deviation of a Poisson distribution (no systematic errors were considered)*

Discussion

I have identified the first discontinuity in free recall and this discontinuity indicates that items with the lowest overall probability of recall are not present in the initial recall. Since the initial recall is generally thought to arise from working memory, the missing items are items that are collected from a different store at a later time. This is the first confirmation of a dual store in word free recall experimental data.

From the requirement that the total recall cannot exceed what is in working memory, item by item, I can set an upper limit on the capacity of working memory. This limit is

Min(overall recall of item i / initial recall of item i)

This is shown in Figures 4 and 5. The upper limit is 3.61, 3.60, 3.81, 3.06, 3.65, and 3.11 for the 10-2, 15-2, 20-2, 20-1, 30-1 and 40-1 data, respectively (for the 40-1 data the item numbers corresponding to zero initial recalls were omitted from this calculation). The overall minimum is 3.06 which is then the upper limit of the capacity of working memory (the statistical error is less than 7% estimating it as sqrt(1/total recall+1/initial recall) even though the two measures are not independent and the errors would tend to cancel out). This is in some agreement with recent serial recall experiments (Chen and Cowan, 2009, find a limit of 3 chunks) and recent still controversial experimental findings for the capacity of vision working memory (see, for example, Anderson, Vogel and Awh, 2011; for an opposing view see Bays and Husain, 2008; and Bays, Catalao and Husain, 2009).

I should note that my limit of 3.06 refers to words, not chunks, and, strictly speaking, it refers to words from the Toronto Word Pool. A word on chunking is also appropriate: Since my limit is the smallest upper limit, it could be that it signifies a set of measurements that happened to include the least amount of chunking in all of the experiments.

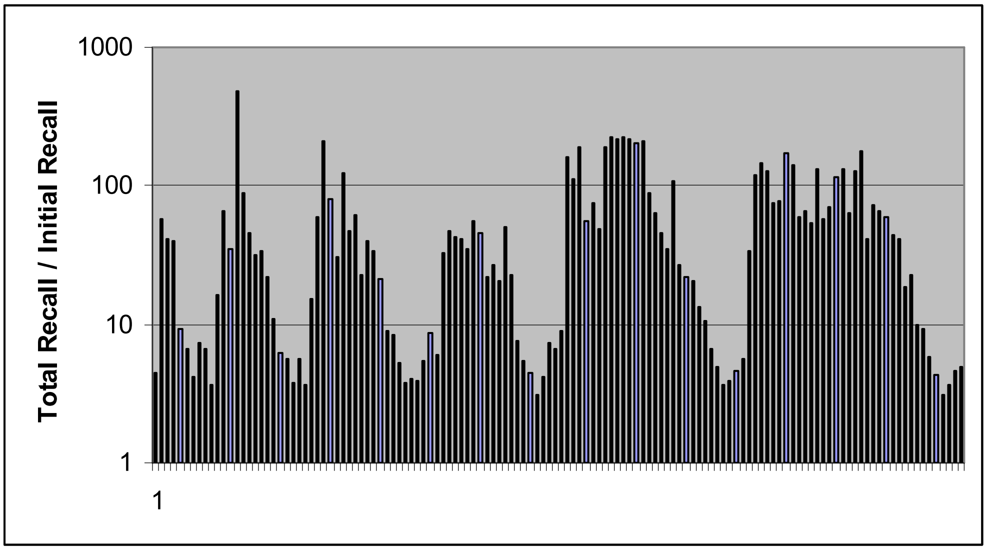

Figure 4. Upper limits on working memory capacity using the ratio of total recall to initial recall. The points are, in item order, 10-2, 15-2, 20-2, 20-1, 30-1, and 40-1 (with a few items missing from 40-1 in which the total recall is zero). The lowest values, which are my estimates for the upper limit of the capacity of working memory, tend to come from one of the last items in each series.

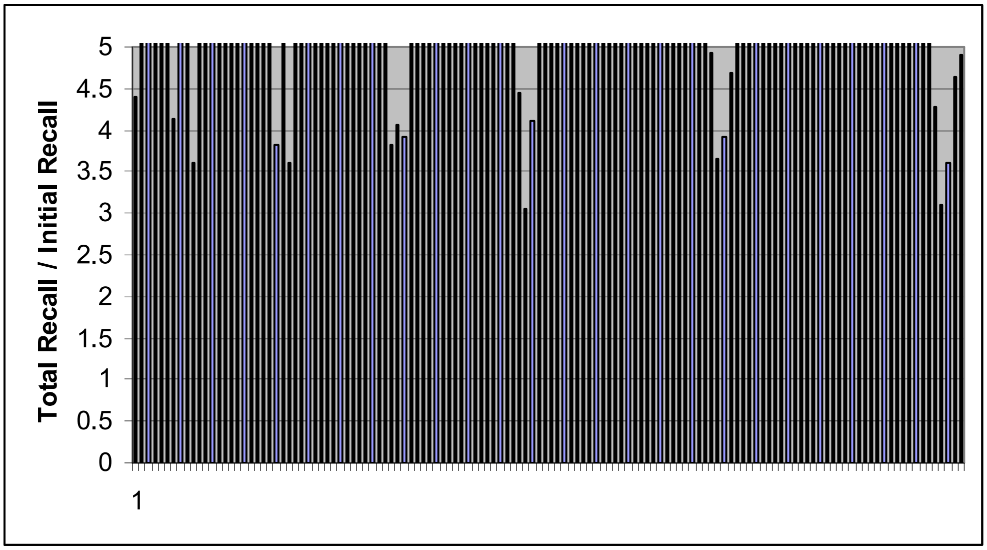

Figure 5. Upper limits on working memory capacity from Figure 5 with a different scale. The smallest upper limit is 3.06 and comes from the 17th item in the 20-2 series.

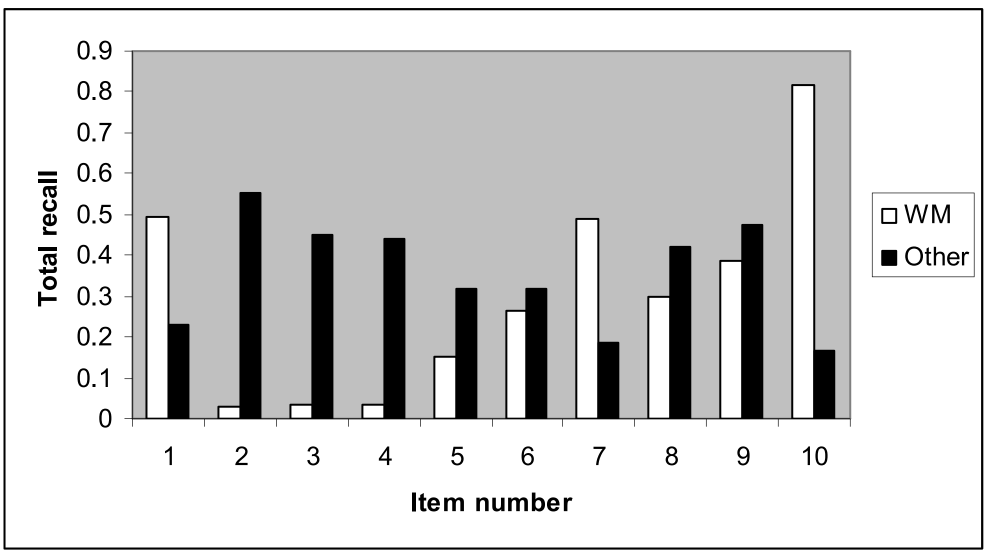

Figure 6. Recall separated into working memory (white columns) and other store (black columns).

In Figure 6 I have plotted what might be the separate contributions from working memory (white columns) and from the other store (black columns). The other store presumably contains words associated with items in working memory as well as the working memory items. I assumed that there are three items in working memory and that the initial recall is an accurate statistical representation of the content in working memory. I note that the variation of recall probability with item number is larger in working memory than in the other store.

This finding provides yet another supporting piece of evidence in favor of the existence of a capacity limited working memory. It also shows that word free recall, while seemingly complex, continuous to elucidate how our memory works, fifty years after the experimental method was introduced.

References

Anderson, D. E., Vogel, E. K., & Awh, E. (2011). Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. *The Journal of neuroscience, 31*(3), 1128-1138.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *The psychology of learning and motivation*, *2*, 89-195.

Baddeley, A. (2012). Working memory: theories, models, and controversies. Annual review of psychology, 63, 1-29.

Balakrishnanl, J. D., & Ashby, F. G. (1992). Subitizing: Magical numbers or mere superstition?. *Psychological research*, *54*(2), 80-90.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*(5890), 851-854.

Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*,*9*(10), 7.

Chen, Z., & Cowan, N. (2009). Core verbal working-memory capacity: the limit in words retained without covert articulation. *Quarterly journal of experimental psychology (2006)*, *62*(7), 1420-1429.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences,* 24(1), 87-114.

Hintzman D, (2011), Research Strategy in the Study of Memory: Fads, Fallacies, and the retrieval for the "‘Coordinates of Truth", *Perspectives on Psychological Science* 6(3) 253–271.

Laming D (1999) Testing the Idea of Distinct Storage Mechanisms in Memory, *International Journal of Psychology* 34, 5-6, 419-426.

McElree, B. (2006). Accessing recent events. *Psychology of learning and motivation*, 46, 155-200.

Murdock Jr, B. B. (1960). The immediate retention of unrelated words. *Journal of Experimental Psychology*, *60*(4), 222.

Murdock Jr, B. B. (1962). The serial position effect of free recall. *Journal of experimental psychology*, *64*(5), 482.

Murdock, B. B. (1974). *Human memory: Theory and data*. Lawrence Erlbaum.

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

Tarnow E (2013) Have we been looking at the wrong curve? Submitted for publication.

Tarnow E (2014) Free Recall Retrieval Process Creates Errors in Short Term Memory but Not in Long Term Memory; Indicates Breaking Down of Store Model. Submitted for publication.