Response probability and latency: a straight line, an operational definition of meaning and the structure of short term memory

Eugen Tarnow, Ph.D.

18-11 Radburn Road

Fair Lawn, NJ 07410

etarnow@avabiz.com (e-mail)

Abstract:

The functional relationship between response probability and time is investigated in data from Rubin, Hinton and Wenzel (1999) and Anderson (1981). Recall/recognition probabilities and search times are linearly related through stimulus presentation lags from 6 seconds to 600 seconds in the former experiment and for repeated learning of words in the latter. The slope of the response time vs. probability function is related to the meaningfulness of the items used. The Rubin et al data suggest that only one memory structure is present or that all memory structures probed show the same linear relation of response probability and time. Both sets of data also suggest that the memory items, presumably in the neocortex, have a finite effective size that shrinks in a logarithmic fashion as the time since stimulus presentation increases or the overlearning decreases, away from the start of the search. According to the logarithmic decay, the size of the memory items decreases to a couple of neurons at about 1500 seconds for recall and 1100 seconds for recognition – this could be the time scale for a short term memory being converted to a long term memory. The incorrect recall time saturates in the Rubin et al data (it is not linear throughout the experiments), suggesting a limited size of the short term memory structure: the time to search through the structure for recall is 1.7 seconds. For recognition the corresponding time is about 0.4 seconds, to compare with the 0.243 seconds given by the data analysis of Cavanagh of Sternberg-like experiments (1972).

1. Introduction

In this paper we will examine the functional relationship between recall/recognition probability and response times. This relationship appears to have been neglected in the literature in part because the two measurements developed separately. In their review of research on recall/recognition probabilities and responses times, Kahana and Loftus (1999) wrote that before the 1970s typically only probabilities were measured because of the difficulty involved in response time measurements. After the proliferation of personal computers in the labs response time measurements became easier to perform. However, if response times were studied they were thought of as separate from response probabilities and were not studied together. Kahana and Loftus wrote that "it is probably fair to say that almost all RT research is concerned with tasks where error rates are negligible" and that "rarely are both investigated simultaneously in a given experimental design." Indeed, even in the Kahana and Loftus paper recall/recognition probabilities and response times are drawn in separate graphs, and with one exception there is no graph showing how the response time varies with response probability. The exception is speed-accuracy trade-off curves for which the manipulated variable is the response time. Similarly, there are no recall/recognition probability versus response time graph in reviews on memory research by Neath (1998) and by Anderson (1995).

The neglect of a simultaneous study of response probability and time also appears in the modeling of experimental data. Global memory models are typically static models (Gronlund & Ratcliff, 1989) and do not involve the element of time needed to account for response times (for a review, see Clark & Gronlund, 1996). There is at least one exception, REM-ARC (Nobel & Shiffrin, 2001) which is an extension of REM (Shiffrin & Steyvers) which replaced SAM (Gillund and Shiffrin (1984)). However, the times considered were those of episodic memory data with lag times between 0.1 and 4.5 seconds, substantially shorter than the lag times we will use in this paper. Since global memory models are not directly derived from the underlying neuronal mechanisms, the predictions of such are probably limited to the experimental results they were fitted to or interpolations thereof. Since they have not been fitted to the data will consider below and since they do not cover the 6 – 600 second time scale global memory models will not be further considered in this paper.

John Anderson (1981) studied recognition and recall probabilities and response times with and without interference. He focused his attention on the fact that interference shifts the curves of response time and probability but also noted that when he plotted response probabilities and times for probabilities from 0.8 to 1.0 he found a straight line. We will use his experimental results later and show that there is more to be found in the data.

I use one additional set of data: in 1999 Rubin, Hinton and Wenzel created a set of data on word recall/recognition probabilities and times ranging from 6 to 600 seconds time lags with very small statistical error bars. The accuracy makes this experimental data a center piece for memory researchers interested in recall and recognition probabilities and response times.

2. Experimental information

In the Rubin et al (1999) experiment, the items used for recall and recognition were different. For recall they used words chosen from Kucera and Francis (1967) to have frequencies between 10 and 100 per million. Proper names, plurals, words with apostrophes, and highly emotional words were excluded. For recognition, they used digit-letter-digit trigrams of the form used in Canadian postal codes. Their data was reported in “lags”. Each trial took six seconds which means that lag of 0 corresponds to 6 seconds after the stimulus presentation started and N lag corresponds to 6*N+6 seconds after stimulus presentation. The data I will use is restated here from the original paper with the additional time component (Tables 1 and 2).

Lag	Seconds after end of stimulus presentation (calculated)	Probability of recall (all 3 measures)	Cued recall response times in seconds for correct responses – (all three measures)	Response times in seconds for incorrect responses – (all three measures)
0	0	.944	1.356	2.292
1	6	.646	1.822	2.722
2	12	.434	2.017	2.938
4	24	.379	2.086	2.872
7	42	.335	2.111	2.960
12	72	.301	2.238	3.001
21	126	.231	2.279	2.970
35	210	.183	2.402	2.978
59	354	.133	2.540	2.969
99	594	.112	2.427	2.927

Table 1 – Recall data (corresponding to tables A1, A4 and A5 in Rubin et al (1999):

Lag	Seconds after end of stimulus presentation (calculated)	Probability of recognition (all 3 measures)	Reaction time in seconds for correct recognition	Reaction time in seconds for incorrect recognition
0	0	0.81	1.128	1.324
1	6	0.642	1.214	1.456
2	12	0.503	1.227	1.509
4	24	0.475	1.247	1.481
7	42	0.401	1.261	1.505
12	72	0.358	1.282	1.517
21	126	0.278	1.254	1.463
35	210	0.195	1.292	1.485
59	354	0.141	1.278	1.472
99	594	0.134	1.287	1.472

Table 2 – Recognition data (corresponding to tables A2, A4 and A5 in Rubin et al (1999))

In the Anderson (1981) experiments, the word items used were similar for recall and recognition and they were selected from Paivio, Yuille, and Madigan (1968) to be high in imagery, concreteness, and meaningfulness.

3. Results and Discussion:

3.1. Correct recall (recognition): Response time is linearly related to probability of correct answer with R^2 of 98% (83%).

Let us begin by plotting the response time against the probability of correct recall in Rubin et al (1999) (Figure 1(a)). The response time is linearly related to the probability of recall with R squared being 98% over a probability range of 0.11 to 0.95 and over a time range of 6 seconds to 600 seconds. A recent item (6 seconds after start of stimulus presentation) requires a total response time of about 1.3 seconds while an item that is typically no longer to be found for most participants (600 seconds after stimulus presentation) requires 2.6 seconds.

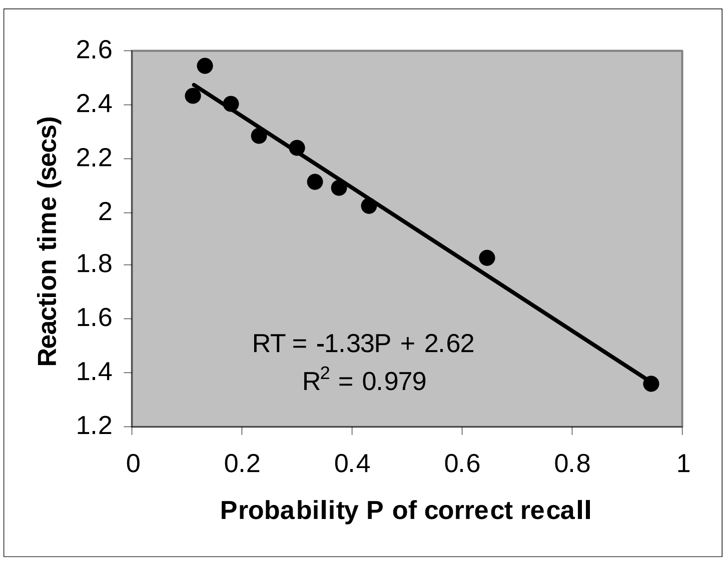

Figure 1 (a): Response time as a function of the probability of correct recall from Rubin et al (1999). The time after stimulus presentation is not shown but short times correspond to high probability of recall and long times correspond to low probability of recall. Data from Table 1.

In Figure 1(b) is shown the corresponding data for recognition. It seems to obey a linear relationship as well over roughly the same range of probabilities (0.13-0.81) and the same time range of 6 seconds to 600 seconds. A recent item requires a total response time of about 1.13 seconds while an item that is old and typically no longer to be found requires 1.33 seconds. The scale of the time differences is much smaller than for recall and the level of statistical noise present in the experiment lowers the R^2 but it is still an impressive 83%.

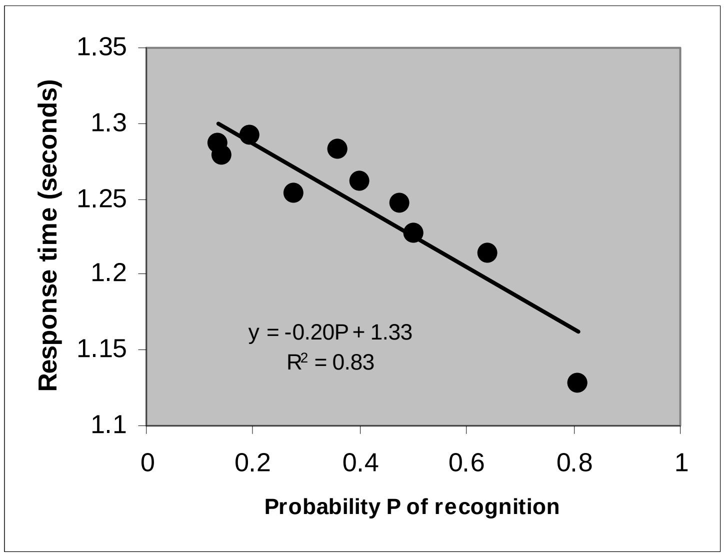

Figure 1 (b): Response time as a function of the probability of recognition from Rubin et al (1999). The time after stimulus presentation is not shown but short times correspond to high probability of recognition and long times correspond to low probability of recognition. Note that the time scale is much smaller than the time scale in Figure 1 (a) so the experimental noise accounts for a larger amount of R^2. Data from Table 2.

The data from Anderson (1981) is shown in Figure 1(c) where I have gone beyond Anderson by plotting all experimental data in the same graph, i.e. recall and recognition with and without interference and included are the data points below the 0.7 cutoff imposed by Anderson. Note that, as Anderson did, the data looks linear (even below Anderson's 0.7 cutoff). Just like the data in the Rubin et al (1999) experiment were linear over a surprisingly large time range, the Anderson data is linear even though it contains points corresponding to new learning and "improved learning" as the subjects studied a second list with similar words and were more adept at the task.

The linear functional curves found should be useful for memory modeling researchers because it presents a simple test for models (Kahana and Loftus, 1999, though the particular experimental circumstances have to be remembered, see, for example, MacLeod and Nelson (1984)).

Surprisingly, the slope of the lines is roughly the same for both recall and recognition (it varies from -1.69 to -1.84). The slopes for recall and recognition in Rubin et al (1999) were very different presumably because the memory items in the Rubin et al (1999) experiment were different for recall and recognition. In the Anderson experiment they were the same. In other words, the slope seems to be related to presented memory items, not to the particular experimental conditions and not due to differences between recall and recognition.

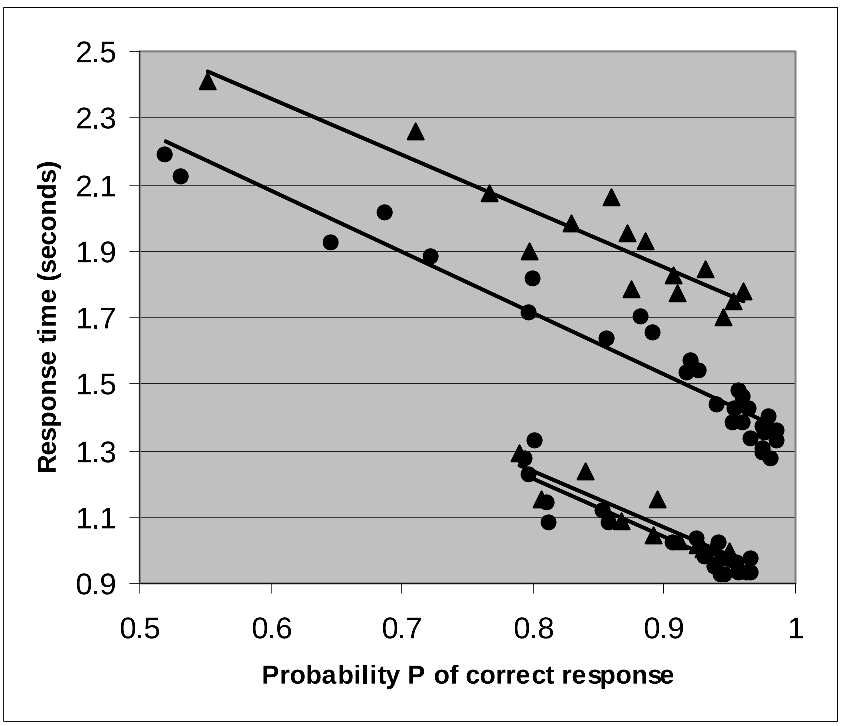

Figure 1 (c): Response time as a function of the probability of recognition (lower curves) and recall (upper curves) from Anderson (1981). The triangles represent the data with interference, the circles data without interference.

The slopes of the response time curves are summarized in Table 3 where I selected the order to be the meaningfulness of the memory items (low to high). The data is plotted in Fig. 2. The more meaningful the memory item, the more negative is the slope. Conversely, if the item has no meaning, the slope seems to be close to 0. Thus, the slope of the response time vs. response probability curve is an operational definition of meaning.

This is not unreasonable. If a stored memory item is recalled at the same speed no matter the probability of recall/recognition, it suggests that the memory item has one and only one structure which does not change either with forgetting (as in the Rubin experiment) or with learning (as in the Anderson experiment). If a stored memory item has multiple meanings, increased learning can add the memory item in different locations, thereby changing the structure of the stored memory item and potentially put some of it in a place that is quicker to find with a resulting quicker response time.

	Slope (seconds)	Memory items
Rubin recognition	-0.2	Digit-letter-digit trigrams (“nonsense”)
Rubin recall	-1.3	Unusual words (“unusual”)
Anderson recall	-1.84 and -1.69	Words high in imagery, concreteness, and meaningfulness (“meaningful”).
Anderson recognition	-1.73 and -1,72	Words high in imagery, concreteness, and meaningfulness (“meaningful”).

Table 3. Slopes for the Rubin and Anderson data.

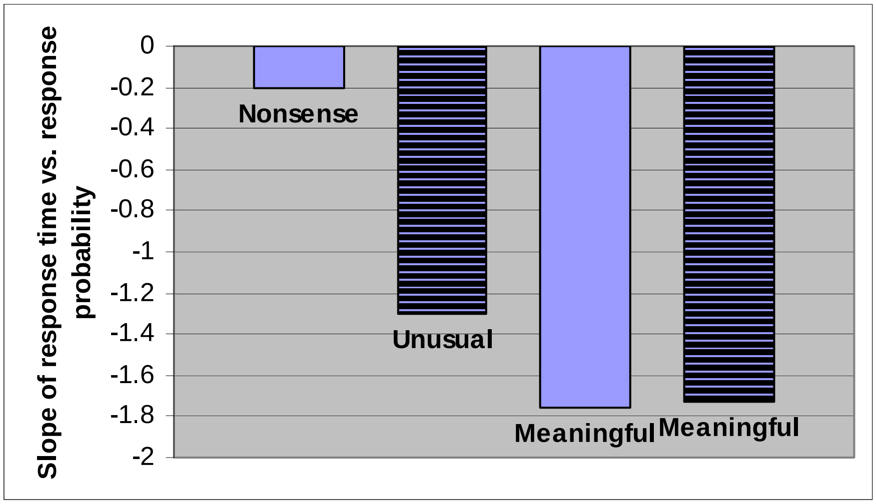

Figure 2: Slopes of the response time vs. response probability curves from Table 1. The solid bars correspond to recognition and the textured bars correspond to recall..

3.2. Correct recall/recognition: The linear relationships point to a single short term memory structure

The established linear relationship between response time and probability of recall and recognition between 6 and 600 seconds in the Rubin et al (1999) experiment reasonably suggests that only one structure is responsible for recall, and, potentially, recognition, during that time period. If there were several structures, it is unlikely that they would all be displaying the same linear relationship between response probability and time.

3.3. Correct recall/recognition: The short term memory structures seems to be shrinking

The linear relationships between response probability and time tell us something about the geometry of the short term memory structure probed. Let us consider three scenarios.

Scenario 1. A non-redundant randomly decaying memory structure fixed in space. This structure should have a search time for correctly identified items which is independent of the probability of finding the item. The items are either there or not and if they are, they are in the same spot whether the probability of finding them is high or low and take the same time to find. From Figure 3 we see that it is not a fit to the experimental data unless the memory items have no meaning.

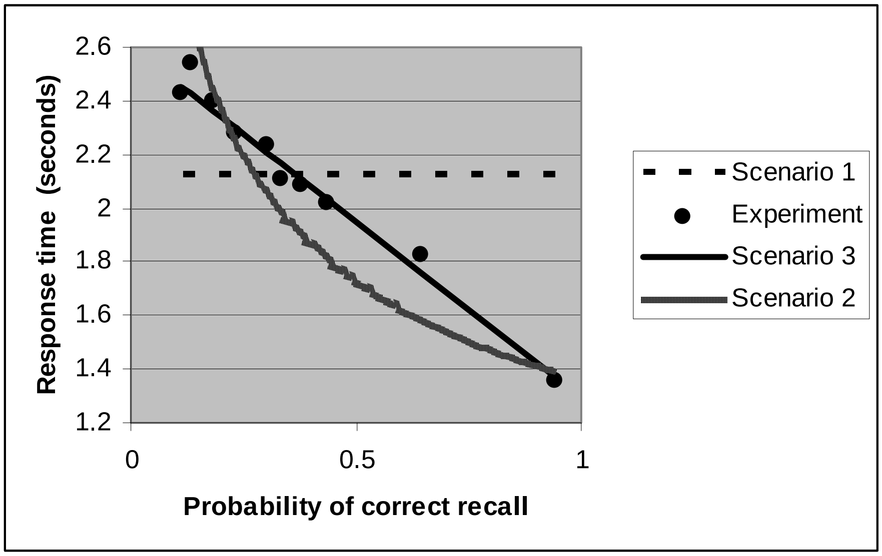

Figure 3. Experimental data from Figure 1(a) shown with best fits of the three scenarios described.

Scenario 2. A multiple redundant randomly decaying memory structure fixed in space. The response time would not be linearly related to the probability of recall but rather the response time is related to 1/P where P is the probability of recall. For example, if there are two copies of an item randomly positioned, it would on the average take half the time to find the item as compared to if there were only one item and so forth. From Figure 3 we see that it is not a fit to the experimental data. A similar scenario was also considered in Anderson (1981), p. 334 and similarly rejected. Ratcliff (1978) proposes that it takes a certain number of features to reach a criterion for detecting the memory item, presumably his theory would fit in this second scenario.

Scenario 3. The memory item has an effective size that shrinks with time after stimulus presentation and increases with repeated learning. The smaller the memory size is, the smaller the probability to find it and the longer away from the starting point it is (Fig. 4). It can be a fit with the experimental data in Figure 3. The size of the memory item is related to the "meaning" of the memory items which presumably is related to the excitation level of the neuron system surrounding the "core" of the memory: if the system is excited it will be quicker to set up the appropriate firing rates which presumably constitute a memory item. The location of the memory item may be in the cerebral cortex (Eichenbaum, 2000) with the hippocampus being the starting point of the search. I label this scenario the Effective Size Model.

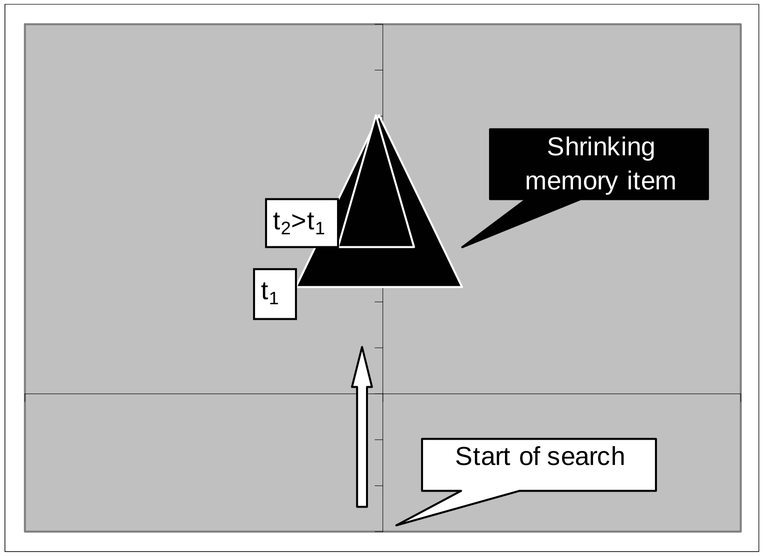

Figure 4: Shrinking/growing memory item. As time passes since the stimulus presentation, the effective size of the memory item shrinks and with it the probability of finding the memory. More learning increases the effective size of the memory item.

Other scenarios are possible. For examples, one could consider a model in which synchronized neuron oscillations are set up and that synchronization defines the memory (see, for example, Gray et al. (1989), Rodriguez et al. (1999) and Jensen and Lisman (1998)). If recall/recognition involves the setting up of such oscillations, it is conceivable that the time to set up such oscillations would increase with time induced changes in synaptic connections: i.e. that the older a memory becomes, the longer it would take to set up such an oscillation to identify a memory item. It would also seem reasonable that large changes in the synaptic connections would result in lower probability of setting up the oscillation and therefore a lower probability of recall.

Eimas and Zeaman (1963) showed that correct response times decreased as overlearning increased: in my model the overlearning stimulates the core memory and increases its effective size if the memory items are meaningful. Single neuron recordings show stimulus-specific sensitivities which decay with time. It appears that the aerial extent of such stimulus specific activity has not been measured.

3.4. Correct response: The short term memory items shrink logarithmically with time and suggested times for conversion of short term memory into long term memory

The effective size of the memory shrinks quickly at first and slower later on. I have defined the “size” to mean the distance in response time from the center of the memory core (response time when the probability P of recalling the item is close to 0) to its periphery (the reader can convert the size into units of neurons by dividing the time by, say, 0.02 seconds, a reasonable time to pass through a neuron). At the time scales measured, the shrinking can be described as a logarithmic relationship of t (Figure 5 (a) for recall and Figure 5(b) for recognition). So, for example, the size of the memory item for recognition is 0 seconds when the probability of finding an item is 0 and 1.29 seconds (2.62 seconds-1.33 seconds) when the probability of finding the item is 1. Notice the remarkably good fits with R^2 at 97% and 94% for recall and recognition.

The logarithmic curve breaks down at large times because the size becomes negative. A reasonable lower limit on the size (which is an upper limit on the time after stimulus presentation) is a couple of neurons. If each of them takes about 10 milliseconds to traverse, then the upper limit on the logarithmic formula for recall (recognition) is about 1500 seconds (1100 seconds). This can be interpreted as a time limit of the short term memory structure before the information is totally gone or converted into long term memory. The time for the probability of recall (recognition) to drop by 50% is about 11 seconds (10 seconds).

The shrinking is quick in the beginning, lowering the probability of correct answers by 50% in ten seconds (similar to the finding of Peterson, Peterson (1959)). The shrinking continues to follow the same logarithmic curve until perhaps 1500 seconds for recall and 1100 seconds for recognition at which point either memory item is the size of a few neurons. The nature of the connection of short term to long term memory is still unknown (Cowan (1993)). Cowan writes that “at present, the basis for believing that there is a time limit to STM is controversial and unsettled … any putative effect of the passage of time on memory for a particular stimulus could instead be explained by a combination of various types of proactive and retroactive interference from other stimuli” (2001). The limits found in this paper suggest that 1100-1500 seconds is a potential time scale to look for a conversion from short term to long term memory. There are other estimates in the literature of the duration of short term memory and non-permanent changes to motor memory appear to last a full 5 hours (R. Shadmehr, and T. Brashers-Krug (1997)).

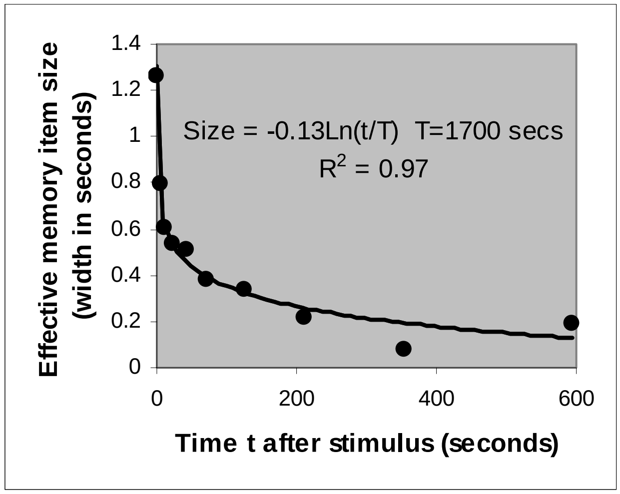

Figure 5(a). Shrinking of the effective size of the Rubin et al (1999) recall memory item where "size" is measured as distance in search time from the center of the memory core to its periphery. The curve represents a two parameter logarithmic fit, moving t=0 seconds to t=0.05 seconds to avoid a divergence. Data from Table 1.

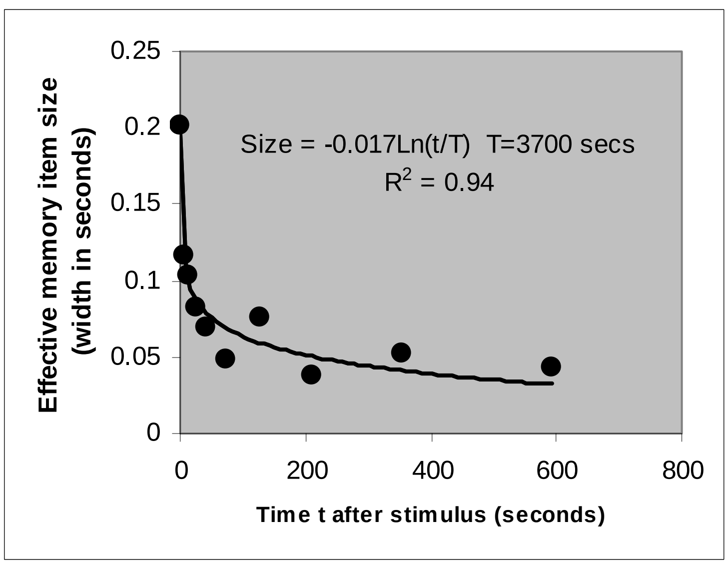

Figure 5(b). Shrinking of the effective size of the Rubin et al (1999) recognition memory item where "size" is measured as distance in search time from the center of the memory core to its periphery. The curve represents a two parameter logarithmic fit, moving t=0 seconds to t=0.05 seconds to avoid a divergence. Data from Table 2.

3.5. Incorrect recall/recognition: Saturation of the response time and the total time to search short term memory during recall.

Let us consider the relationship between response time and "incorrect" recall (recognition) in the Rubin et al (1999) experiment as shown in Figure 6 (a) (6 (b)). When the correct recall and recognition probabilities are large, the response times for incorrect recall and recognition changes linearly just like for correct recall and recognition. However, when the correct recall (recognition) probability decreases they saturate and become constant.

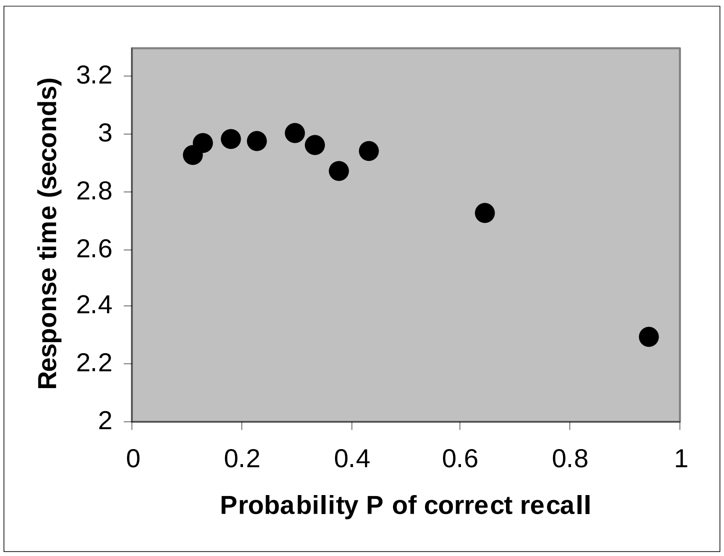

Figure 6(a): Response time for incorrect recall as a function of the probability of <u>*correct*</u> *recall (to keep the scales the same throughout the paper). Data from Table 1.*

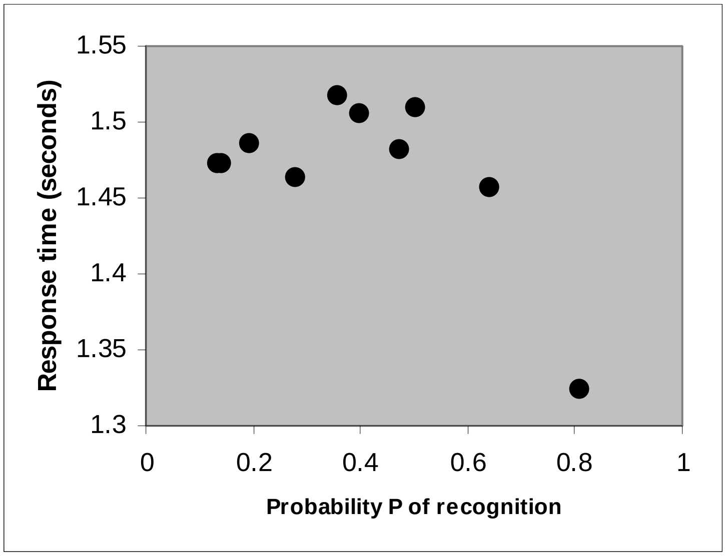

Figure 6(b): Response time for incorrect recognition as a function of the probability of correct recognition (to keep the scales the same throughout the paper). Data from Table 2.

The response times are always larger for incorrect recall or recognition than for correct recall or recognition (the differences in response time between the incorrect and correct searches are shown in Figures 7 (a) and (b) below). The data with the lowest level of noise is the recall data. It is possible to infer the maximal time to search the brain for recall, if we assume that the search yielding the correct result is not exhaustive but the search yielding the incorrect result is. The time it takes to finish an exhaustive search of the particular brain structure involved is the difference between the total response time for incorrect recall of 3 seconds at low correct recall probability (Figure 7 (a)) minus the shortest response time recorded, the response time for correct recall at P=1 (Figure 1 (a)), 1.3 seconds which yields 1.7 seconds. The noise in the data for recognition makes it more difficult to assess the corresponding time – a rough estimate is 1.5-1.13=0.4 seconds. This latter estimate appears to be the first non-Sternberg task result that can be compared to the Cavanagh (1972) time estimate to fully search short term recognition memory of 0.243 seconds.

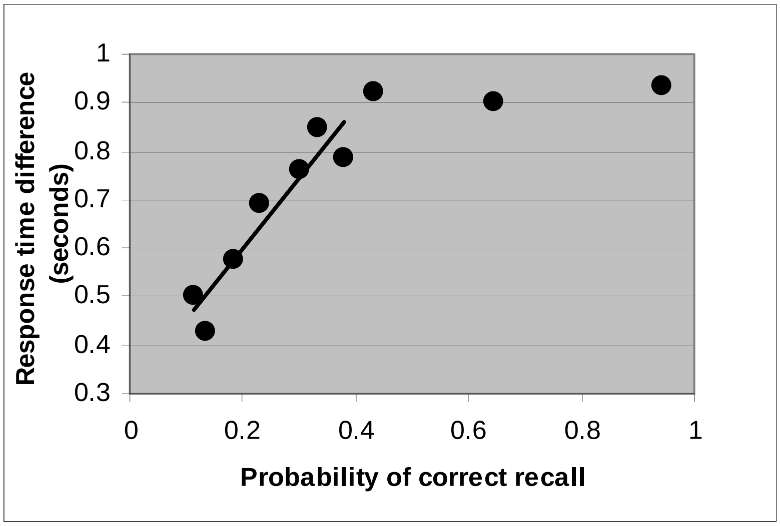

Figure 7(a): Difference in response times between incorrect and correct recall as a function of the probability of correct recall. Data from Table 1.

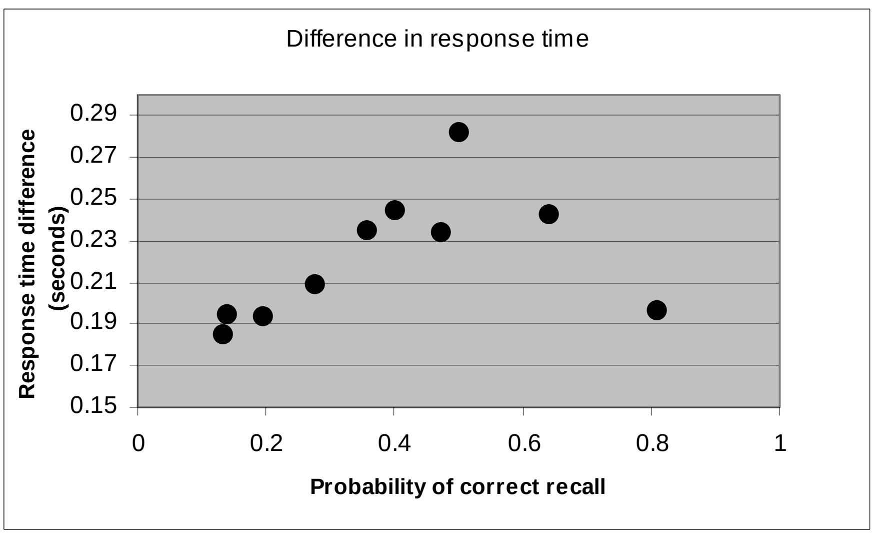

Figure 7(b): Difference in response times between incorrect and correct recognition as a function of the probability of correct recognition. Data from Table 2.

In eye-witness line-up identifications studies there is a current controversy as to whether there is a time boundary beyond which it is likely that the eyewitness identification is incorrect (see, for example, Brewer, N., Caon, Alita, Todd, Chelsea, Weber, Nathan (2006)). In the Rubin et al (1999), it would seem that the time boundary would be close to the shortest possible time for incorrect response.

4. ACKNOWLEDGEMENT

Thanks to Nelson Cowan and Stephen Schmidt for comments on the manuscript.

5. REFERENCES

Anderson, J. R. (1981). Interference: The relationship between response latency and response accuracy. Journal of Experimental Psychology: Human Learning and Memory, 7, 326-343.

Anderson, J.R. (1995). *Learning and Memory.* New York: Wiley.

Brewer, N., Caon, Alita, Todd, Chelsea, Weber, Nathan (2006). “Eyewitness Identification Accuracy and Response Latency”, Law and Human Behavior 30, p. 1.

Cavanagh, J. Patrick (1972). “Relation between the Immediate Memory Span and the Memory Search Rate”, Psychological Review, 79, 525- 530.

Clark, S. E., Gronlund, S.D. (1996) “Global matching models of recognition memory: How the models match the data”, Psychonomic Bulletin & Review 3(1), 37-60.

Cowan, N (1993). “Activation, attention and short term memory”, Memory & Cognition 21(2):162-7.

Eichenbaum, H, (2000). A cortical-hippocampal system for declarative memory. Nature Reviews: Neuroscience (1), 41-50.

Eimas. P.D. and D. Zeaman, 1963. Response speed changes in an Estes’ paired-associate ‘miniature’ experiment. Journal of Verbal Learning and Verbal Behavior 1, 38`4-388.

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

Gray, C. M., König, P., Engel, A. K. & Singer, W. (1989) “Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization, which reflects global stimulus properties. “Nature 338:334–37.

Jensen, O., Lisman, J.E. (1998). “An Oscillatory Short-Term Memory Buffer Model Can Account for Data on the Sternberg Task”. J. Neuroscience, 18, p. 10688-10699.

Kahana, M, Loftus, G. (1999). “Response Time versus Accuracy in Human Memory”.

MacLeod, C., Nelson, T. (1984). Response latency and response accuracy as measures of memory. Acta Psychologica 57, 215-235.

Neath, I (1998). “Human Memory”. Brooks/Cole. Pacific Grove. P. 253.

Nobel, P. A., & Shiffrin, R. M. (2001). Retrieval processes in recognition and cued recall. Journal of Experimental Psychology: Learning, Memory, and Cognition, 27, 384–413.

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

Ratcliff, R. (1978). “A theory of memory retrieval.” Psychological Review, 85, 59-108.

Gronlund, S.D., Ratcliff, R. (1989) “Time Course of Item and Associative Information: Implications for Global Memory Models”, Journal of Experimental Psychology 15, 846-858.

Rodriguez, E., George, N., Lachaux, J.-P., Martinerie, J., Renault, B. & Varela, F. J. (1999) “Perception’s shadow: long-distance synchronization of human brain activity.” Nature 397:430–33.

Rubin, D.C., Hinton, S., Wenzel, A., (1999), “The Precise Time Course of Retention”, Journal of Experimental Psychology: Learning, Memory and Cognition, Vol 25, No. 5, 1161-1176.

R. Shadmehr, and T. Brashers-Krug (1997). ”Functional Stages in the Formation of Human Long-Term Motor Memory.” *J. Neurosci.* 17: 409-419.