Journal of Experimental Psychology:
Learning, Memory, and Cognition
1982, Vol. 8, No. 4, 343-352

0278-7393/82/0804-0343$00.75

A Note on the Measurement of Primary Memory Capacity

Jeroen G. W. Raaijmakers
University of Nijmegen, Nijmegen, The Netherlands

It is argued that the modification of the traditional Waugh and Norman (1965) method for the estimation of primary-memory capacity proposed by Watkins (1979) is not consistent with the assumptions of the modal two-store model. An alternative method is described that does not incorporate the assumption, implicit in the method proposed by Watkins (1979), of independence between the probabilities of recall from primary and secondary memory. It is shown that the most general version of the new method is prone to error because the number of parameters to be estimated equals the number of data points. Two simplified solutions are described, one based on a geometric decay function for the probability of recall from primary memory and one based on the general buffer model described by Phillips, Shiffrin, and Atkinson (1967). It is shown that this method will give a valid estimate of primary-memory capacity; that is, when applied to simulated data with a known buffer capacity, it will generate an estimate of primary-memory capacity that corresponds quite closely to the known value.

In recent years the problem of the measurement of primary-memory capacity, first discussed by Waugh and Norman (1965), has received renewed attention (Martin & Jones, 1979; Watkins, 1974, 1979). In this article I confine myself to the free-recall paradigm, since that is the context in which those measurement procedures have been most frequently used. The goal of these techniques is to derive from the recency and asymptotic parts of the serial position curve an estimate of primary-memory capacity.

Waugh and Norman (1965) assumed that an item presented in serial position i could be recalled from primary memory or from secondary memory (or both). Specifically, they assumed that the probability of recalling item i from secondary memory (SM_i) was *independent* of the probability of recalling that item from primary memory (PM_i). Furthermore, they assumed that items still in primary memory at the time of recall were always recalled. The probability of recalling the item presented in serial position i (R_i) is then given by

$$R_i = PM_i + (1 - PM_i)SM_i. \qquad (1)$$

Hence,

$$PM_i = \frac{R_i - SM_i}{1 - SM_i}. \qquad (2)$$

Waugh and Norman assumed that SM_i was constant for all serial positions and could be estimated from the probability of recall for the middle, prerecency part of the serial position curve (R_c); that is,

$$SM_i = R_c \quad \text{for all } i.$$

Primary-memory capacity (r) is defined as the sum of all PM_i's from the recency and asymptotic parts of the serial position curve. Disregarding the primacy part, we may define

$$r = \sum_{i=1}^{L} PM_i,$$

where L refers to the item presented last.

One problem with the Waugh and Norman method is that it does not take into account the phenomenon of negative recency. Negative recency refers to the observation that the items at the final list positions are sometimes less well recalled than the items in the middle positions. This phenomenon was first observed by Craik (1970) in a delayed-recall procedure. In this procedure, a number of lists are presented and tested in an immediate-recall test. At the end of the session the subjects are unexpectedly asked to recall as many words as possible from all lists presented during the entire session. This negative-recency effect is usually explained by the assumption that the subject stops rehearsing the items when testing begins. Hence, the final list items are not rehearsed as much as the items in the prerecency part of the list. Given the assumption that amount of rehearsal and probability of recall are corre-

Requests for reprints should be sent to J. G. W. Raaijmakers, Psychologisch Laboratorium, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands.

lated, it follows that the final items are less well recalled. That negative recency is not usually observed when a distractor task is interpolated between presentation and testing of a list is probably due to continued rehearsal of the recency items during the course of the intervening task itself. In this procedure, negative recency only occurs when the distractor task is difficult and very long (see Gardiner, Thompson, & Maskarinec, 1974).

The occurrence of negative recency leads to an underestimation of primary memory capacity when the traditional Waugh and Norman method is used because SM_i is overestimated for items near the end of the list. Watkins (1974) proposed a modification of Equations 1 and 2 that was supposed to overcome this problem (in addition to correcting for nonindependence between PM_i and SM_i due to failure to encode the item). However, Martin and Jones (1979) correctly pointed out that this was not true and that in fact no adjustment had been made for negative recency. In response to this critique, Watkins (1979) presented a further modification that does take negative recency into account. The solution proposed by Watkins (1979) supposedly follows directly from the assumptions of the modal two-store model underlying the Waugh and Norman procedure. As described previously, the two-store model predicts negative recency because the final list items will not have been rehearsed as much as the other items. Watkins (1979) assumed that this would lead to the following conclusion:

> Put somewhat more formally, the average amount of rehearsal—and therefore SM_i—for items presented in a given recency position is proportional to the length of stay in primary memory by the time recall is signalled, which in turn is proportional to the sum of PM_i at its own and each subsequent serial position. (p. 446)

In the example given by Watkins (1979, Table 1) the following equation for SM_i was used:

$$SM_i = R_c \sum_{j=i}^{L} PM_j / \sum_{j=1}^{L} PM_j . \tag{3}$$

In this equation, serial position L corresponds to the item presented last (i.e., the most recent one). Substitution of Equation 3 into Equation 1 gives the estimation equation used by Watkins (disregarding for the moment the correction proposed to account for possible failures of encoding). Although Watkins did not mention how this equation was derived from the assumptions of the two-store model, it is likely that it was based on the following assumptions:

1. The probability distribution for the number of trials that item i has been in primary memory from its presentation until the beginning of recall is given by

$$P(n_i \geqslant j) = PM_{L-j+1} \quad 1 \leqslant j \leqslant L - i + 1,$$

hence,

$$P(n_i = j) = \begin{cases} 1 - PM_L & j = 0 \\ PM_{L-j+1} - PM_{L-j} & 1 \leqslant j \leqslant L - i \\ PM_i & j = L - i + 1. \end{cases} \tag{4}$$

This equation is based on the assumption that an item in primary memory at the beginning of recall is always recalled and that once an item leaves primary memory it never returns. Hence, PM_{L-j+1} is equal to the probability that an item stays at least j trials in primary memory. Note that it is assumed that this distribution is the same for all items except that it is truncated at a different point (since the item presented in position i cannot have been in primary memory for more than $L - i + 1$ trials). The average number of trials in primary memory is then given by

$$E(n_i) = \sum_{j=1}^{L-i+1} jP(n_i = j) = \sum_{j=1}^{L-i} j(PM_{L-j+1} - PM_{L-j}) + (L - i + 1)PM_i = \sum_{j=i}^{L} PM_j . \tag{5}$$

2. The *average* probability of recall from secondary memory for a given item i is proportional to the *average* number of trials that item i has been in primary memory ($E[n_i]$). Thus, for some constant c,

$$SM_i = c \cdot E(n_i). \tag{6}$$

According to the two-store model, $PM_i = 0$ for items chosen from the middle, asymptotic part of the serial position curve. For those items, SM_i is equal to R_c, the probability of recall for the middle section of the serial position curve. It follows that for items in the asymptotic part, $E(n_i)$ is equal to

$$\sum_{j=1}^{L} PM_j$$

(since $PM_j = 0$ for all $j < i$). Hence,

$$c = R_c / \sum_{j=1}^{L} PM_j . \tag{7}$$

Substitution of Equations 5 and 7 into Equation 6 gives the formula proposed by Watkins (1979).

However, Equation 3 does not follow from the assumptions underlying the two-store model. To see this, it is helpful to take a closer look at Equation 1. In this equation SM_i may be interpreted as the probability of recalling item i from secondary memory, *given that it is not recalled from* *primary memory*. Let R_i be the event that item i is recalled, and P_i the event that item i is in primary memory at the time of recall. Note that, since items in primary memory are always recalled, $P(P_i) = PM_i$. Equation 1 may then be rewritten as follows:

$$\begin{aligned} P(R_i) &= P(R_i \cap P_i) + P(R_i \cap \bar{P}_i) \\ &= P(P_i) + P(R_i|\bar{P}_i)P(\bar{P}_i) \\ &= P(P_i) + [1 - P(P_i)]P(R_i|\bar{P}_i). \quad (8) \end{aligned}$$

Hence,

$$SM_i = P(R_i|\bar{P}_i).$$

By using Equation 3, Watkins (1979) implicitly assumed that the probability of recalling item i from secondary memory was *independent* of whether or not that item was recalled from primary memory. That is, the formula used by Watkins for SM_i should have been equal to the probability of recall from secondary memory given that item i is no longer in primary memory at the time of recall. However, $E(n_i)$ equals the expected number of trials in primary memory, regardless of whether or not item i was still in primary memory at the time of recall. Thus, it was assumed that

$$E(n_i|\bar{P}_i) = E(n_i),$$

and this implies that the number of trials in primary memory was assumed to be independent of whether or not item i was recalled from primary memory. It should be evident that this independence assumption cannot be true given the rehearsal assumptions described previously. For example, suppose that primary-memory capacity equals 4, a list of 20 items is presented, and at the time of recall the item in Position 10 is still in primary memory (and hence is recalled from primary memory). In this case the probability of recalling this item from secondary memory would be much higher than if it had remained only, say, 4 "trials" in primary memory. Conversely, if the penultimate item is no longer in primary memory at the time of recall, it must have been immediately kicked out of primary memory; hence, it will be registered only weakly in secondary memory.

A second point that should be noted about Equation 3 is that it is based on the assumption that the *average* probability of recall is proportional to the *average* length of stay in primary memory. In the framework of a two-store model a more natural assumption would be that the probability of recall from secondary memory is proportional to the number of trials that the item spent in primary memory. Superficially, this might seem to be equivalent to the assumption made by Watkins. This is not the case, however. Since the probability of recall cannot exceed 1.0, the proportionality assumption must take the following form (where $SM_{i,n}$ denotes the probability of recalling item i from secondary memory given that it has spent n trials in primary memory):

$$SM_{i,n} = \begin{cases} \alpha n & n < 1/\alpha \\ 1.0 & n \geq 1/\alpha. \end{cases} \quad (9)$$

From Equation 9 it follows that the average probability of recall from secondary memory is *not* proportional to the average number of trials in primary memory. In the next sections I show how one can derive a theoretically better founded procedure for the estimation of primary memory capacity based on the assumptions of the modal two-store model.

A New Method for the Estimation of Primary-Memory Capacity

From Equation 8 it is evident that in order to calculate the probability of recall from secondary memory one needs the distribution for the number of trials that an item has been in primary memory given that it is no longer in primary memory at the time of recall. Assuming Equation 4, this probability is given by

$$P(n_i = j|\bar{P}_i) = \begin{cases} (1 - PM_L)/(1 - PM_i) & j = 0 \\ (PM_{L-j+1} - PM_{L-j})/(1 - PM_i) & 1 \leq j \leq L - i. \end{cases}$$

Note that because an item that has left primary memory never returns to it, $P(n_i = j \text{ and } \bar{P}_i) = P(n_i = j)$. Under the assumption that $SM_{i,0} = 0$, Equation 8 may be rewritten as

$$\begin{aligned} P(R_i) &= PM_i + (1 - PM_i) \sum_{j=l}^{L-i} SM_{i,j}P(n_i = j|\bar{P}_i) \\ &= PM_i + \sum_{j=l}^{L-i} SM_{i,j}(PM_{L-j+1} - PM_{L-j}). \end{aligned} \quad (10)$$

Since Equation 10 contains too many unknown parameters, it is of little use unless some restrictions are made. Because our aim is the estimation of the PM_i's, these restrictions will have to involve the specification of $SM_{i,n}$ in terms of the PM_i's and R_c, the probability of recall at asymptote.

The simplest assumption is that $SM_{i,n}$ is a linear function of n:

$$SM_{i,n} = \begin{cases} \alpha n + \beta & n \leqslant (1-\beta)/\alpha \\ 1.0 & n > (1-\beta)/\alpha. \end{cases} \quad (11)$$

The parameters α and β are assumed to depend on R_c and hence on experimental parameters such as list length and presentation time per item. Such a linear assumption agrees quite well with the results of overt rehearsal experiments (see Brodie & Murdock, 1977; Rundus, 1971; Rundus & Atkinson, 1970). However, Equation 10 still contains one parameter too many. Therefore, some restriction on α or β is necessary. Consider the restriction $\beta = 0$. In this case, Equation 11 reduces to Equation 9. Substitution of this equation into Equation 10 gives the following result:

$$P(R_i) = \begin{cases} PM_i + \alpha \sum_{j=1}^{L-i} j(PM_{L-j+1} - PM_{L-j}) & i \geqslant L - m, \text{ where } m \text{ equals the integer part of } 1/\alpha \\ PM_i + \alpha \sum_{j=1}^{m} j(PM_{L-j+1} - PM_{L-j}) \\ \quad + \sum_{j=m+1}^{L-i} (PM_{L-j+1} - PM_{L-j}) & i < L - m. \end{cases} \quad (12)$$

Since

$$\sum_{j=m+1}^{L-i} (PM_{L-j+1} - PM_{L-j}) = PM_{L-m} - PM_i,$$

it follows that for all values of $i \leqslant L - m$, $P(R_i)$ is a constant equal to $P(R_{L-m})$. Hence, $P(R_{L-m})$ should equal the asymptotic probability of recall, R_c. Thus,

$$R_c = PM_{L-m} + \alpha \sum_{j=1}^{m} j(PM_{L-j+1} - PM_{L-j}). \quad (13)$$

Since

$$\sum_{j=1}^{L-k} j(PM_{L-j+1} - PM_{L-j}) = \sum_{j=k+1}^{L} PM_j - (L-k)PM_k, \quad (14)$$

Equation 13 reduces to

$$R_c = PM_{L-m} + \alpha \sum_{j=L-m+1}^{L} PM_j - \alpha m PM_{L-m}. \quad (15)$$

Note that when $PM_{L-m} = 0$ (and thus $PM_j = 0$ for all $j \leqslant L - m$), α is equal to R_c/r (Equation 7).

Before applying these equations to estimate primary-memory capacity, one additional constraint on the parameters PM_i must be considered. From Equation 4 it is evident that PM_i must be a monotonically increasing function of i (otherwise $P[n_i = j]$ might become negative). It can be easily derived from Equation 12 that PM_i is an increasing function of i if and only if $P(R_i)$ is an increasing function of i. Hence, if the observed probabilities of recall are not perfectly monotonic (due to error), the parameters PM_i cannot be estimated in such a way that the predicted probabilities of recall are all equal to the corresponding observed probabilities. An obvious solution is to estimate the PM_i in such a way that the predicted probabilities give a least squares fit to the data. Fortunately, there is a well-known solution to this problem. It involves applying Equation 12 not directly to the observed probabilities but to a set of monotonic probabilities that have a least squares fit to the observed probabilities. This set of monotonic probabilities is obtained by *monotone regression*, that is, by averaging those observed probabilities that violate monotonicity (see Kruskal, 1964). For example, if one had a sequence 1.0, .8, .6, .7, one would average .6 and .7, and the monotone approximation would be 1.0, .8, .65, .65. The general rule is that one averages the first number that violates monotonicity and its predecessor. A more complicated example is given in Table 1. In this example, $P(R_{L-11})$ and $P(R_{L-10})$ are averaged. However, since $P(R_{L-13})$ is larger than this average, all four probabilities—$P(R_{L-10})$, $P(R_{L-11})$, $P(R_{L-12})$, and $P(R_{L-13})$—must be averaged. It should also be noted that monotonicity implies that the probabilities are restricted to values greater than or equal to R_c.

We are now in a position to solve Equation 12 for the PM_i. The procedure is as follows:

1. Choose an initial estimate for r,

$$\sum_{j=1}^{L} PM_j,$$

and compute initial estimates for PM_i, for example, by assuming a geometric decay function for PM_i.

2. Compute α (and m) from Equation 15.

3. Solve Equation 12 in order to get a new set of estimates for the PM_i.

4. If these new estimates are different from the previous estimates, return to Step 2 for a new iteration; if they are not different, the solution has been obtained.

Although we do not have a formal proof, I have found that this procedure always converges. In this way, r can be estimated in an iterative manner, similar to the procedure described by Watkins (1979).

Unfortunately, this procedure does not work very well. In order to show this, several sets of serial position functions were generated by Monte Carlo simulation of the Search of Associative Memory (SAM) model of Raaijmakers and Shiffrin (1980), a particular quantitative model for free recall that incorporates a buffer model for primary memory similar to the one proposed by Atkinson and Shiffrin (1968). It should be emphasized that the detailed structure of the SAM model is irrelevant to our purposes. This model is only used as one handy device for generating serial position curves that arise from a combination of short-term and long-term recall. Any other model with this property would have done as well. What is relevant is that with such a model we are able to generate simulated serial position curves (that is, the "data" are not error free, due to chance fluctuations arising from the Monte Carlo simulations). Thus, we have sets of serial position functions with *known* values for r, the primary-memory capacity.

The SAM model was used to generate 18 serial position curves, each curve based on 5,000 quasi subjects (simulation runs). These serial position curves differed in list length (20, 30, or 40 items), presentation time per item (1 or 2 sec), and the buffer size, r, used in the simulations (r = 3, 4, or 5). In order to see whether the proposed estimation method is also applicable when there are individual differences in primary memory capacity, serial position curves with different buffer-size parameters but with the same list length and presentation time were combined and averaged pairwise. For example, the three curves with list length equal to 20 items and a presentation time of 1 sec per item were averaged pairwise to yield three new curves with mean buffer sizes equal to 3.5, 4.0, or 4.5. Thus, a total of 36 different serial position curves were generated.

The procedure described previously was applied to these generated data and was found to give grossly inaccurate estimates for r, the primary memory capacity. An example is given in Table 1. The second column of this table gives the generated values for $P(R_i)$. In this example, list length

Table 1
Simulated Data and Resulting Estimates for the Probability of Recall From Primary Memory (PM_i)

i	$P(R_i)$	$P^*(R_i)$	PM_i	True PM_i
L	1.000	1.000	1.000	1.000
$L - 1$	.715	.715	.693	.708
$L - 2$	.587	.587	.544	.503
$L - 3$	.492	.492	.424	.359
$L - 4$	.452	.452	.368	.257
$L - 5$	.425	.425	.326	.184
$L - 6$	.414	.414	.307	.133
$L - 7$	.397	.397	.274	.096
$L - 8$	.394	.394	.267	.070
$L - 9$	.391	.391	.258	.051
$L - 10$	.384	.385	.238	.037
$L - 11$	.390	.385	.238	.027
$L - 12$	.377	.385	.238	.020
$L - 13$	.390	.385	.238	.014
$L - 14$	.382	.382	.000	.011

Note. i = item; L = item presented last; $P(R_i)$ = probability that item is recalled; $P^*(R_i)$ = probability of recall after monotone regression (see text for explanation).

was set equal to 30 items and presentation time was set to 2 sec per item. For these data the true buffer size equals 3.5 (i.e., these data are the average of the serial position curves generated by the Monte Carlo simulations for r = 3 and r = 4). The next column, labeled $P^*(R_i)$, gives the monotonic approximation to these data obtained by monotone regression. The last two columns give the estimated and the true values for PM_i, the probability of recall from primary memory. In this case an estimate of $\hat{r}$ = 5.4 was obtained, although the data were generated with a mean buffer size of 3.5. Such a gross overestimation was found for many of the generated serial position curves. Moreover, the PM_i tended to remain stable at a high value rather than decline to zero. The reason for this discrepancy quickly became obvious. Equation 12 represents a system of n equations and n parameters. This means that the procedure is very sensitive to error, as was the case here. Note that the monotonic restriction on the parameters does not constrain the set of equations. Monotone regression makes some of the $P(R_i)$ equal, but this simply means that the corresponding PM_i will also be equal. Hence the number of equations and the number of parameters will be reduced to the same extent, still leaving k equations and k unknowns.

It must therefore be concluded that such a method for the estimation of primary-memory capacity is not very useful. It is necessary to make additional simplifying assumptions that will constrain the estimates for PM_i. In the next section I describe two such simplifications, both based on Atkinson and Shiffrin's buffer model for rehearsal processes in primary memory (Atkinson & Shiffrin, 1968).

More Constrained Versions of the Estimation Method

One of the simplest ways to constrain the parameters is to assume a geometric distribution for the number of trials that an item is rehearsed in primary memory; that is,

$$P(n = j) = (1 - 1/r)^{j-1} 1/r,$$
$$PM_{L-i} = (1 - 1/r)^i. \quad (16)$$

Such an assumption follows from a buffer model for primary memory in which each item currently in the buffer has the same probability $(1/r)$ of being replaced from the buffer by a new item (r equals the buffer size). Substitution of Equation 16 into Equation 12 gives the following result:

$$P(R_i) = \begin{cases} \left(1 - \frac{1}{r}\right)^{L-i} + \alpha\left[r - \left(1 - \frac{1}{r}\right)^{L-i}(r + L - \mathrm{i})\right] & i \geqslant L - m \\ P(R_{L-m}) = R_c & i < L - m. \end{cases} \quad (17)$$

Equation 17 may be used to estimate r, the primary-memory capacity, in a very simple way. For any trial estimate for r, α can be estimated from the identity $P(R_{L-m}) = R_c$. Primary-memory capacity may then be estimated by using Equation 17 to generate a least squares estimate for r (i.e., the value of r that minimizes the sum of the squared differences between the observed and the predicted probabilities of recall at the recency part of the serial position curve). Note that this procedure does not make any assumptions concerning long-term memory retrieval other than that the probability of recall is proportional to the length of stay in primary memory. Since this assumption is somewhat arbitrary, I also explored some alternative assumptions.

The first assumption considered is that $SM_{i,n}$ is a general linear function of n (i.e., Equation 11). I analyzed the predictions obtained with this assumption by considering the two extreme special cases obtained by setting either $\beta = 0$ (Model 1a) or $\alpha = 0$ (Model 1b). Note that Model 1a corresponds to Equation 9 and leads to the solution given in Equation 17. In addition to this linear function for $SM_{i,n}$, I also examined the assumption that $SM_{i,n}$ is a negative exponential function of n, the number of trials in primary memory (Model 2). Thus, the following three models were considered:

Model 1a,

$$SM_{i,n} = \begin{cases} \alpha n & n \leqslant 1/\alpha \\ 1.0 & n > 1/\alpha, \end{cases}$$

Model 1b,

$$SM_{i,n} = \beta = R_c,$$

Model 2,

$$SM_{i,n} = 1 - \alpha^n \quad (0 < \alpha < 1).$$

Model 2 led to the following prediction for the serial position curve:

$$P(R_i) = \left(1 - \frac{1}{r}\right)^{L-i} + \sum_{j=1}^{L-i} (1 - \alpha^j)\left(1 - \frac{1}{r}\right)^{j-1}\left(\frac{1}{r}\right)$$
$$= 1 - \frac{\alpha - \alpha\left[\alpha\left(1 - \frac{1}{r}\right)\right]^{L-i}}{\alpha + r - \alpha r}. \quad (18)$$

Since at asymptote, $P(R_i)$ should equal R_c and PM_i should equal 0 (and hence $L - i$ should be relatively large), α must be as follows:

$$\alpha = \frac{r(1 - R_c)}{R_c + r(1 - R_c)}.$$

These three models were applied to each of the 36 serial position curves generated by Monte Carlo simulation of the SAM model, as described previously. For each of these models the estimates for r (estimated from those serial position curves that had the same true buffer size) were averaged. Figure 1 plots these mean estimated values for r against the actual buffer capacity used in the simulation of the serial position functions.

From these results it is evident that Model 1a (Equation 17) gives a reasonably accurate estimate of the true buffer size, whereas the other two models seriously underestimate primary-memory capacity. Note that Model 1b is very similar to the traditional Waugh and Norman (1965) method for estimating primary-memory capacity.

The present method, based on Model 1a, gives much higher estimates for primary-memory capacity than the traditional method. The latter method frequently generates estimates around 2.5 (see Glanzer & Razel, 1974), whereas the present method gives values around 4. It should also be noted that the present method does not tend to generate lower estimates of primary-memory capacity with higher levels of recall at the asymptotic portion of the serial position curve (see Watkins, 1974, p. 704). This results from the deletion of the assumption of independence between the probabilities of recall from primary and secondary memory.

The simple solution described above is based on the assumption of a geometric "decay" function for primary memory. In some cases, however, a distinctively S-shaped recency curve is observed (e.g., Murdock, 1962). Such data are inconsistent with our simple model. As was already shown by Atkinson and Shiffrin (1968, pp. 154–155; see also Phillips, Shiffrin, & Atkinson, 1967), such S-shaped recency curves can be accomodated by a slightly different type of buffer model.

In this less restricted version of the buffer model, the item to be replaced from the buffer upon presentation of a new item is not selected at random from the r items that make up the buffer. Instead, the probability that a particular item is replaced (kicked out of the buffer) is a function of its recency of presentation: The "oldest" item has the highest probability of being replaced, and the "newest" item has the smallest probability of being replaced. Phillips, Shiffrin, and Atkinson (1967) proposed the following formula for the probability that the item in the ith oldest position is replaced (where $i = 1$ corresponds to the oldest item and $i = r$ to the most recent item):

$$k_i = \frac{\delta(1-\delta)^{i-1}}{1-(1-\delta)^r}.$$

As δ approaches zero, k_i approaches $1/r$, and when $\delta = 1$, $k_1 = 1$: The oldest item will always be the one that is replaced.

Now let us define $\beta_{i,j}$ as the probability that an item currently in the ith oldest position is kicked out of the buffer by the jth succeeding item. Since each item starts in the rth oldest position (i.e., as the newest item) the probability that an item resides for exactly j trials in the buffer equals $\beta_{r,j}$. As shown by Phillips, Shiffrin, and Atkinson (1967), the $\beta_{i,j}$'s can be calculated by solving the following set of difference equations:

$$\beta_{1,j} = (1-k_1)\beta_{1,j-1}$$

$$\beta_{i,j} = [\sum_{l=1}^{i-1} k_l]\beta_{i-1,j-1} + [\sum_{l=i+1}^{r} k_l]\beta_{i,j-1} \quad 1 < i < r$$

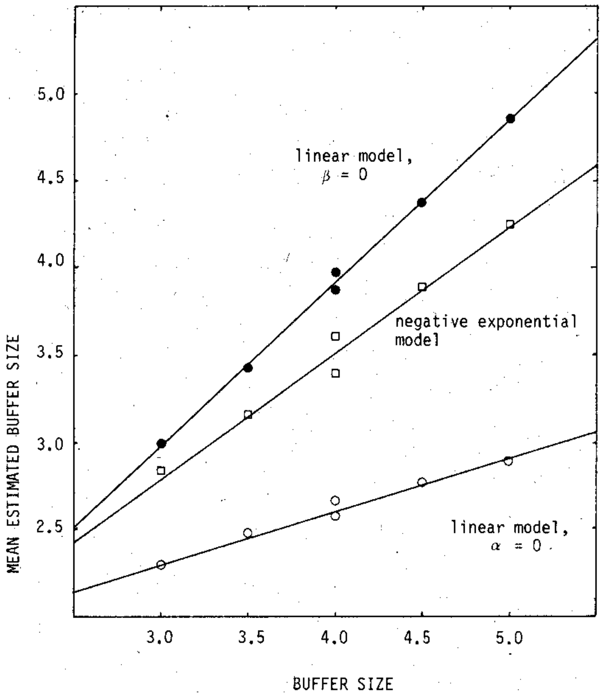

Figure 1. Plot of the mean estimated buffer size or primary memory capacity against the actual buffer size used in the simulations. (Filled circles indicate that probability of recall from secondary memory, $SM_{i,n}$, is proportional to length of stay in the buffer [Model 1a]; open circles indicate that $SM_{i,n}$ is constant [Model 1b]; squares indicate that $SM_{i,n}$ is a negative exponential function of n, the number of trials in primary memory [Model 2].)

$$\beta_{r,j} = (1-k_r)\beta_{r-1,j-1},$$

with the initial conditions $\beta_{i,1} = k_i$. The probability that the item currently in position i is knocked out by the jth succeeding item equals the probability that on the next trial one of the other items is replaced times the probability that the item stays an additional $j-1$ trials in the buffer (starting from the next trial). From the above definitions it follows that

$$P(n_i = j) = \beta_{r,j} \quad (1 \leqslant j \leqslant L-i)$$

and

$$PM_{L-i} = 1 - \sum_{j=1}^{i} \beta_{r,j}. \tag{19}$$

Hence, the probability of recall for the item in serial position i is given by

$$P(R_i) = \begin{cases} (1 - \sum_{j=1}^{L-i} \beta_{r,j}) + \alpha(\sum_{j=1}^{L-I} j\beta_{r,j}) & i \geqslant L-m \\ P(R_{L-m}) = R_c & i < L-m. \end{cases} \tag{20}$$

In order to estimate r, the primary-memory capacity, one has to estimate for each of a number of (integer) values of r the best-fitting value of δ by minimizing, for example, a least squares function; the estimate of r is then equal to that value of r resulting in the best (i.e., lowest) value for the least squares criterion.

Note that in this more general procedure the characteristic aspect of the present method is retained: No assumptions are made concerning the long-term storage (LTS) retrieval process other than that in any given condition the probability of recall from LTS is proportional (within limits) to the number of trials in the buffer (Equation 9).

This procedure was applied to the serial position curves obtained by Murdock (1962). Estimates of r and δ were obtained that fit simultaneously the curves from the 15–2, 20–2, 20–1, 30–1, and 40–1 conditions, where the first number stands for the list length and the second number for the presentation time, in seconds, per item (the 10–2 condition was not used, since R_c cannot be estimated reliably with short list lengths). In this estimation procedure the last 15 serial positions of each curve were used, except for the 15–2 condition in which the last 10 positions were used. The best fitting values of r and δ were $\hat{r} = 4$ and $\hat{\delta} = .301$. With these parameter estimates the oldest item is about three times as likely to be replaced as the newest item ($k_1 = .396$, $k_2 = .276$, $k_3 = .193$, and $k_4 = .135$). As can be seen in Figure 2, with these parameter estimates I was able to fit the recency parts of the serial position curves quite accurately.

Finally, Watkins (1974) proposed a correction of the traditional Waugh and Norman method in order to accomodate the observation that the final list item is not always recalled. Since it is assumed that all items that are in primary memory at the time of recall are recalled, it follows that not all items enter primary memory. Watkins's correction of Equations 1 and 2 is based on the assumption that the independence of primary and

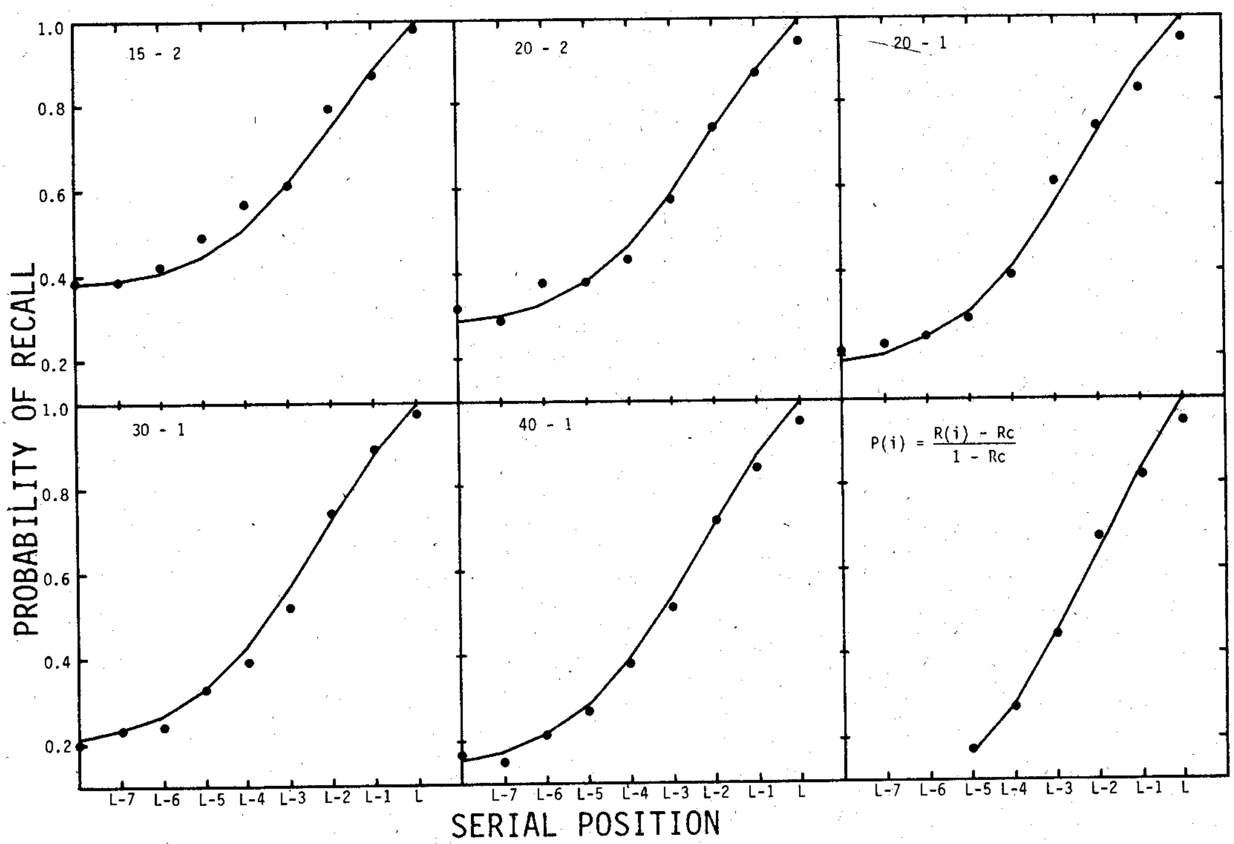

Figure 2. The recency parts of various serial position functions obtained by Murdock (1962). (The continuous lines give the predicted probabilities of recall resulting from the present method of estimating primary memory capacity. Numbers in upper right corners of first five panels indicate list lengths and presentation times of various conditions. The lower right figure gives the mean probability of recall from primary memory resulting from the application of the Waugh and Norman, 1965, method to both the observed and the predicted serial position functions. See text for discussion of equation. L = item presented last.)

secondary memory applies only to those items that gained access to primary memory. It should be noted that the most general version of our method is still applicable when some items do not enter primary memory (see Equation 10). In that case, $PM_L \neq 1.0$, hence $P(n_i = 0) \neq 0$. However, in the more constrained versions of the method, we do assume that all items enter primary memory. Hence, some correction may be desirable when the probability of recall for the final list item is less than 1.0.

There seem to be two ways to explain this observation within the framework of a two-store model. First, one might assume that some items do not enter primary memory (the correction proposed by Watkins, 1974). This leads to the following modification of Equation 16:

$$PM_{L-i} = v(1 - v/r)^i,$$

where v denotes the probability that an item enters the buffer upon presentation. A similar correction may be applied to Equation 19. Alternatively, one might assume that although all items enter primary memory, the probability of recall does not equal 1.0 for all items still in primary memory at the time of recall. Due to a number of factors (e.g., output interference) it might sometimes be the case that one of the r items still in primary memory at the beginning of recall is lost from primary memory before it is actually recalled. With this assumption Equation 8 has to be changed to

$$P(R_i) = vP(P_i) + (1 - v)P(P_i)P(R_i|P_i) + P(\bar{P}_i)P(R_i|\bar{P}_i), \quad (21)$$

where v denotes the probability of recall from primary memory for those items still in primary memory at the beginning of recall. Since item i will have been in primary memory for $L - i + 1$ trials if it was still in primary memory at the beginning of recall, it follows that $P(R_i|P_i)$ will be equal to $SM_{i,L-i+1}$. It is practically impossible to distinguish these two alternatives on the basis of the predicted serial position curves. Both predict similar curves and both lead to a slightly higher estimate for primary-memory capacity. However, the second alternative is supported by some results observed in overt rehearsal studies. Both Rundus (1971) and Rundus and Atkinson (1970) observed that the probability of recall was the same for all items in the final rehearsal set. However, this probability was not equal to 1.0 but was equal to .96 in the experiment of Rundus (1971) and .92 in the study of Rundus and Atkinson (1970). For this reason I prefer the second alternative. As an example, Equation 20, corrected according to Equation 21, was applied to Murdock's serial position data (in the same way as described previously). In this application v was estimated as .965, and $\hat{r}$ and $\hat{\delta}$ did not change very much ($\hat{r} = 4$ and $\hat{\delta} = .348$, giving $k_1 = .424$, $k_2 = .277$, $k_3 = .182$, and $k_4 = .120$).

To recapitulate, I have described a method based on the general two-store model that makes it possible to estimate primary-memory capacity without making any particular processing assumptions concerning retrieval from LTS other than the reasonable approximation that the probability of recall from LTS or secondary memory within a given condition is proportional (within limits) to the amount of time the item has spent in primary memory. I have shown that the most general version of this method does not work in practice because of its sensitivity to measurement error. Because it is based on a system of n equations and n unknowns, it tries to find a solution that fits the data exactly, including the measurement error, and this makes it necessary to introduce additional simplifying assumptions. I next showed that the buffer model proposed by Atkinson and Shiffrin (1968) generates a relatively simple solution to this problem. Moreover, this method gives a valid measure of primary-memory capacity (i.e., the estimate is close to the true value), a property lacking in both the original Waugh and Norman (1965) method and Watkins' modifications of that method (Watkins, 1974, 1979).

References

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2). New York: Academic Press, 1968.

Brodie, D. A., & Murdock, B. B., Jr. Effect of presentation time on nominal and functional serial-position curves of free recall. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 185–200.

Craik, F. I. M. The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 143–148.

Gardiner, J. M., Thompson, C. P., & Maskarinec, A. S. Negative recency in initial free recall. *Journal of Experimental Psychology*, 1974, *103*, 71–78.

Glanzer, M., & Razel, M. The size of the unit in short-term shortage. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 114–131.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964, *29*, 1–27.

Martin, M., & Jones, G. V. Negative recency and measurement of primary memory. *British Journal of Psychology*, 1979, *70*, 441–443.

Murdock, B. B., Jr. The serial position effect in free recall. *Journal of Experimental Psychology*, 1962, *64*, 482–488.

Phillips, J. L., Shiffrin, R. M., & Atkinson, R. C. Effects of list length on short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1967, *6*, 303–311.

Raaijmakers, J. G. W., & Shiffrin, R. M. SAM: A theory of probablistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 14). New York: Academic Press, 1980.

Rundus, D. Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, 1971, *89*, 63–77.

Rundus, D., & Atkinson, R. C. Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 99–105.

Watkins, M. J. Concept and measurement of primary memory. *Psychological Bulletin*, 1974, *81*, 695–711.

Watkins, M. J. Modifying Waugh & Norman's primary measurement procedure: An alternative solution. *British Journal of Psychology*, 1979, *70*, 445–447.

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

Received July 20, 1981
Revision received January 26, 1982 ■