# RECALLING THE LIST-BEFORE-LAST: A CAUTIONARY TALE

### Donald LAMING1

RÉSUMÉ – Mémorisation et restitution de l'avant-dernière liste : un conte moral

Une question non résolue dans le domaine de la recherche sur la mémoire est la suivante : comment se déclenche un souvenir libre ? On a cité une expérience de Shiffrin [1970], où les participants doivent se remémorer l'avant-dernière liste, comme preuve que la récupération dans les expériences de rappel libre est de quelque manière provoquée par un signal. L'autre possibilité serait que la récupération soit spontanée et que les mots rappelés soient choisis rétrospectivement après récupération. Cet article illustre la différence entre les deux hypothèses et compare les données de Shiffrin à celles de Murdock & Okada [1970] et à trois autres expériences, pour soutenir que le taux de souvenirs de l'avant-dernière liste n'est pas plus grand que ce que l'on attendrait d'une récupération spontanée qui prend potentiellement en compte toutes les listes précédentes. Le nombre des intrusions, quand le souvenir ne doit porter que sur la liste fondamentale, corrigé par l'estimation de la proportion des récupérations superflues supprimées, suffit à expliquer le nombre des mots retrouvés de l'avant-dernière liste.

MOTS CLÉS – Expérience de Shiffrin, Mémoire, Rappel libre, Restitution, Souvenir de l'avantdernière liste

ABSTRACT – An unresolved question in memory research is: What cues free recall? An experiment by Shiffrin [1970], in which participants were instructed to recall the list-before-last, has been cited as evidence that retrieval in free recall experiments is (somehow) cued. The alternative is that retrieval is spontaneous and the words recalled are selected only in retrospect, after they have been retrieved. This article illustrates the difference between these two hypotheses and then compares Shiffrin's published data with data from Murdock & Okada [1970], and with three other experiments, to argue that the rate of recall from the list-before-last is no greater than one should expect from a spontaneous retrieval that (potentially) addresses all preceding lists. The number of intrusions, when recall is requested from the ultimate list, corrected by an estimate of the proportion of unwanted retrievals that are suppressed, is sufficient to account for the number of words recalled from the list-before-last.

KEYWORDS - Free recall, Memory, Recalling the list-before-last, Retrieval, Shiffrin's experiment

### 1. INTRODUCTION

These days reviewers increasingly tend to sign their reviews. A signed review signifies a conscientious evaluation of the paper submitted and a reviewer who is prepared to justify his or her comments. But there is one circumstance in which a reviewer does well to preserve anonymity.

<sup>&</sup>lt;sup>1</sup> Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB (Great Britain), drjl@hermes.cam.ac.uk.

I submitted a paper on inter-recall times in free recall to the *Psychological Review*. My analysis of latencies was based on the assumption that retrieval from memory was spontaneous; that is, each memory trace has its individual probability of retrieval and pops up in memory in a Poisson process with precisely that density. A cue (if one may still call it a "cue") acts only in retrospect to identify the material required after it has been retrieved. This provided a simple basis for calculating the waiting time until the first of those words, not yet recalled, was retrieved.

However, one anonymous reviewer objected:

...the attempt is to do away with the notion that free recall is cued at all. I don't think that can reasonably work. For instance, in the Shiffrin [1970] recall list N-1 paradigm, something has to be focusing participants on the list before last. If it was purely about accessibility relative [sic] based on temporal order then intrusions of the Nth list would have been catastrophic.

In the summary by the action editor this became:

... Shiffrin [1970] showed that subjects given successive lists can selectively recall items from the next-to-last list if cued to do so (without having to first take the time to recall items from the most recent list).

The argument below first recounts Shiffrin's [1970] experiment and explains the issue at stake. A subsequent section, then compares Shiffrin's experiment with four others of similar design, to demonstrate just how catastrophic intrusions from the  $N^{\text{th}}$  list would be.

## 2. TWO HYPOTHESES

Shiffrin [1970] showed his participants a series of lists of either 5 or 20 words. The words were presented visually for 1 s each, with one minute allowed for recall. After each list (except the first) participants were asked to recall words from the last list but one (that is, from list N-1). The schedule of list lengths was arranged to provide equal numbers of lists of lengths 5 and 20 (as list N-1) followed by equal numbers of lists of lengths 5 and 20 (as list N-1) followed by equal numbers of lists of lengths 5 and 20 (as list N-1) followed by equal numbers of lists of lengths 5 and 20 (as list N-1) followed by equal numbers of lists of lengths 5 and 20 (as list N-1) followed by equal numbers of lists of length 5 and 20 (as list N-1) followed by equal numbers of lists of lengths 5 and 20 (as list N-1) followed by equal numbers of lists of lengths 5 and 20 (as list N-1) followed by equal numbers of lists of lengths 5 and 20 (as list N-1) followed by equal numbers of lists of lengths 5 and 20 (as list N-1) followed by equal numbers of lists of lengths 5 and 20 (as list N-1). This invokes the unspoken assumption that the loss of memory with lapse of time is due to degradation of that which has been stored and that 'memory search' is independent of time since presentation.

The two alternative hypotheses of interest here are unrelated to Shiffrin's conclusion, but closely related to his data. The first hypothesis says that memory search is actuated by some cue (unidentified) that specifically retrieves words from the listbefore-last; the second says that words are retrieved spontaneously and output only if they are judged to come from the list-before-last, 'memory search' acting in retrospect after candidate words have been retrieved.

To illustrate these two opposed hypotheses, Figure 1 presents the data from two successive trials (by the same participant) in an experiment by Brodie and Murdock (1977, Expt 1; there is no comparable data available from Shiffrin's experiment). Eighteen words were presented on a memory drum at 1.25 s/word. Each number in the figure is the serial position of a word in the list. The numbers in inverse contrast denote the presentation of each successive word. The other numbers denote the overt rehearsal or recall of a word already presented. 'Int' denotes an intrusion, a word that was not in the list just presented.

#### LIST 2

Rehearsal sequence coded by serial position in the stimulus list:

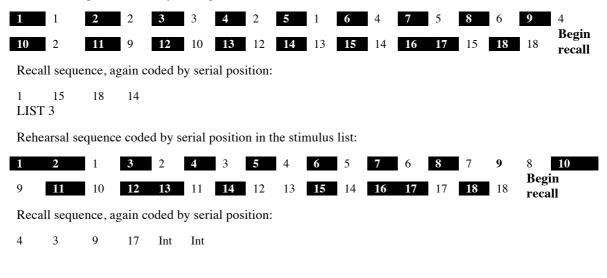


FIGURE 1. Sample data for two successive trials from an experiment by Brodie and Murdock [1977, Expt 1].Words are denoted by serial position within the list. The numbers in inverse contrast denote the presentation of each successive word; other numbers denote overt

rehearsals or recalls. 'Int' denotes a word that was not in the list just presented.

Ordinarily, in free recall experiments, each recall addresses the immediately preceding list (and rehearsals during its presentation). In Shiffrin's procedure, on the other hand, the recall following List 3 would have addressed List 2. The anonymous reviewer envisaged that at the recall following List 3 the instruction to recall the list-before-last would have directed retrieval specifically to List 2, and that unwanted retrievals from List 3 would have been much fewer in number than if List 3 had been addressed directly in the ordinary way. Spontaneous retrieval, on the other hand, supposes that at the recall following List 3, words would have been retrieved from *both* lists and, moreover, that the pattern of that retrieval would have been unchanged by the instruction to recall the list-before-last. That instruction merely means that words from List 3, if they are identified as such, are suppressed rather than being recalled.

### **3. INTRUSIONS IN FREE RECALL**

Since Shiffrin's paper amounted to two pages only in *Science*, only recalls from list N-1 were reported. For intrusions from list N I rely on comparison with an experiment by Murdock & Okada [1970]. Murdock and Okada also presented lists of 20 words, again visually at one word per s, followed by one minute for recall. There is therefore an exact comparison between lists of length 20 followed by another list of length 20 in Shiffrin's Experiment 1 (in which the recall period was always one minute) and data from Murdock & Okada [1970], specifically for lists presented at one word per s, but with recall requested from the last-presented list<sup>2</sup>. In the Murdock and Okada data (the raw

 $<sup>^2</sup>$  The data from the other conditions in Shiffrin's Experiment 1 and from his other two experiments are very similar, so that the selection of this particular condition, for the sake of exact comparison, does not gloss over any important differences.

data are available for analysis) it is possible to distinguish between original recalls, repetitions and intrusions, and the average numbers per list for each category of response in these two experiments are set out in Table 1.

Category of response	Shiffrin (1970)	Murdock & Okada (1970)			
	Responses	Responses	Estimated retrievals		
Original recalls (List N)		6.5			
Repetitions (List N)		0.2 <sup>3</sup>	1.2 <sup>3</sup>		
Original recalls (List <i>N</i> –1) or Intrusions (List <i>N</i> )	2.8	0.6 <sup>3</sup>	<i>4.2</i> <sup>3</sup>		
Total		7.3			

TABLE 1. Mean numbers of responses and estimated retrievals per list in the experiments by Shiffrin [1970] and Murdock and Okada [1970]

Note Data from Shiffrin [1970] and Murdock and Okada [1970].

Murdock & Okada's 36 participants (1 word/s presentations only) made an average of 7.3 responses per list. Of that 7.3, 6.5 were original recalls, 0.2 were repetitions and 0.6 intrusions. However, it is known that once a word has been recalled, its probability of a second recall is much reduced, implying that, while it may be retrieved a second time, if that retrieval is identified as a repetition, it is suppressed. Table 2 sets out the data from Murdock & Okada (1970, 1 word/s). Each row categorises recalls in the  $k^{th}$  output position. Each recall is either a repetition of a recall originally output somewhere in  $1^{st}$  to  $(k-1)^{th}$  position or is an original recall (i.e., not a repetition). A quick visual examination reveals that the number of repetitions is much less than the number of original recalls, but this is complicated by the fact that one cannot have a repetition until the word that is repeated has already been recalled. The Appendix sets out a randomisation test which compares the frequency of repetitions with the frequency of original recalls, circumventing the difficulty that a repetition cannot occur until the original recall that is being repeated has itself been uttered and that repetition frequencies are necessarily lower on that account<sup>4</sup>. The same procedure can be utilised to estimate the median number of retrievals prior to suppression of recall by asking what number of repetitions would equate to a significance level of 0.5. That median (895 compared with 131 repeated recalls) suggests an average (Table 1) of 1.2 repeated retrievals per list. That is, about six repeated retrievals out of seven<sup>3</sup> are estimated to have been suppressed. Applying the same rate of suppression to intrusions (words from previous lists which, again, were not required to be recalled) suggests that the 0.6 per list that were recorded resulted from about 4.2 retrievals. The immediately preceding list would have been the principal source of those 4.2 intrusive retrievals.

 $<sup>^3</sup>$  The statistics in Table 1 have been rounded to 1 d.p. To 3 d.p. the statistics are 0.182, 1.243, 0.611, 4.175.

<sup>&</sup>lt;sup>4</sup> The significance level attaching to the hypothesis that the probability of repetition in Table 2 is not reduced is calculated by this procedure to be exp(-967.9).

Output position	Original		Output position of original recall								Total			
of repetition	recall	1	2	3	4	5	6	7	8	9	10	11	12	recalls
2nd	719	0												719
3rd	715	0	0											715
4th	705	1	0	0										706
5th	649	2	5	0	0									656
6th	548	10	8	5	1	0								572
7th	424	10	7	8	2	0	0							451
8th	283	6	11	5	5	0	0	0						310
9th	187	5	3	5	1	3	1	0	0					205
10th	99	6	2	1	2	2	2	1	0	0				115
11th	52	0	1	1	1	2	0	0	0	0	0			57
12th	23	1	0	1	1	0	0	0	0	0	0	0		26
13th	3	2	1	0	0	0	0	0	0	0	0	0	0	6

 TABLE 2. Numbers of repetitions as a function of output position in the recall sequence and of output position of the potential source.

*Note* Data from Murdock and Okada [1970].

Murdock & Okada's participants averaged 7.3 recalls per list, and are estimated to have averaged 11.9 retrievals. The idea that Shiffrin's participants averaged only 2.8 *retrievals* in the same period of time (1 min) is simply implausible; there must have been intrusions, not reported because anything other than recalls from list N-1 was irrelevant to Shiffrin's purpose in writing. Transposing the Murdock and Okada averages to Shiffrin's experiment suggests about 7.7 retrievals (6.5 + 1.2, including repetitions) from the last list presented. Since in terms of Shiffrin's task these were intrusions, they would have been suppressed, if recognised as such by the participant, and simply not reported (by the experimenter) otherwise. A comparison of the 2.8 original recalls from list N-1 with the 4.2 intrusive retrievals estimated from the Murdock & Okada data provides no ground for proposing an enhanced retrieval of words from list N-1. Shiffrin's tally is about what one would expect if retrieval from memory was spontaneous, indifferently from list N and listN-1.

# REPLICATIONS OF SHIFFRIN'S EXPERIMENT

Ward & Tan [2004] have replicated Shiffrin's experiment and his result. They have also repeated the experiment with 4 s between successive words and instructions to rehearse aloud. Participants do rehearse the words from the list (N-1) that is to be recalled and, analysis of rehearsals in free recall [Laming, 2006, 2008] shows that the subsequent recall of rehearsed words is enhanced. Analysis of the recall sequences from Murdock & Okada [1970; see Laming, 2010] indicates some small level of rehearsal even at 1 word/s, so this is a contributory factor in both the experiments by Shiffrin [1970] and Murdock & Okada [1970]. Jang & Huber [2008, Expt 2] have also repeated Shiffrin's experiment (with slightly different design parameters) and have shown that there are indeed recalls from the intervening list, notwithstanding an instruction to suppress such recalls. Very recently Unsworth, Spillers & Brewer [2012] have replicated Shiffrin's

experiment in some detail and confirmed his results. They reported intrusions from both the intervening List N and from the preceding List N-2, with relatively more intrusions from List N when that list was of length 20 rather than 5. The proportion recalled from List N-1 was substantially reduced below the proportion obtained from a control list (i.e., List N with instruction to recall the list just presented) and the mean latency was increased in proportion. These results are all consistent with the idea that the pattern of retrieval is unchanged by the instruction to recall the list-before-last; only the selection of retrievals for output is different.

The key to this reassessment of Shiffrin's [1970] data is that recall of the listbefore-last produces notably fewer original recalls than recall of the ultimate list. Harbison, Dougherty, Davelaar & Fayyad [2009] have reported a related experiment with four 8-item lists presented in succession, each list accompanied by a cue word. After a second presentation of the four lists, participants were asked to retrieve one list at a time, the designated list being indicated by its cue word. Participants were, indeed, able to recall some words from the list four places back in the order, but only averaged two words out of 8.

Ward & Tan [2004, p. 1197] commented:

... it becomes apparent that the participants (in Shiffrin [1970]) found the task exceedingly difficult...

I question whether this was actually so. To be sure, the participants recalled *between* one and three items from the TBR lists, but whether that represents an experienced difficulty, or simply a natural rate of retrieval indifferently from all preceding lists, is unclear.

### 4. DISCUSSION

The ultimate question here concerns the mechanics of retrieval. Is retrieval from memory effected by a cue or is it spontaneous, with a cue (if one may still call it a "cue") acting only in retrospect to identify the material required after it has been retrieved? The process of retrieval is not observable, so the answer to this question must depend on how well the alternative answers accommodate experimental findings.

The reviewer for the *Psychological Review* mistakenly equated retrieval with actual recall. But we know that participants retrieve some words that they do not utter in recall; first, because they tell us so and, second, because the rehearsal of a word increases its likelihood of subsequent recall, while actual recall has the opposite effect (Table 2). The inference has to be that participants retrieve some words a second or third time (1.2 per list in Murdock and Okada [1970]), but do not utter them because they are identified as 'already recalled'. This means that retrieval is also unobservable. It may be inferred from recall, but recall does not identify all retrievals.

It is widely supposed that retrieval from memory is cued. No cue, no retrieval. This sits well with the analysis of data from recognition memory, paired associates, and serial recall, but free recall presents a problem. Fifty years of intensive research has failed to discover an operative cue. It is common to propose a search of relevant items in memory [Anderson & Bower, 1972; Howard & Kahana, 2002; Mensink & Raaijmakers, 1988; Raaijmakers & Shiffrin, 1980]. As Unsworth, Spillers & Brewer [2012] have recently put it:

In such models it is assumed that the search set is determined, in part, by the match between context stored in the items and the context present during

retrieval such that the greater the overlap between the two, the more likely an item has in being included in the search set and subsequently recalled [e.g., Mensink & Raaijmakers, 1988]. Thus context at retrieval acts as the primary cue and dictates the extent to which relevant items will be recalled [Unsworth, Spillers & Brewer, 2012, p. 1].

This sketch pays lip service to the idea of cued retrieval, but sounds suspiciously like selection of relevant material ("context at retrieval acts as the primary cue") after it has been retrieved. In truth, so long as only frequencies of recalls are available as data, there is no way of distinguishing between a cue eliciting retrieval directly and a cue acting in retrospect to select relevant material after it has been retrieved; except...

... in free recall experiments in which the participants have rehearsed aloud, extrapolation of the pattern of rehearsal beyond the point at which recall is requested predicts the subsequent sequence of recalls with an interesting level of accuracy [Laming, 2006, 2008]<sup>5</sup>. Those predictions assume that stimulus presentations and previous rehearsals are retrieved at random (i.e., spontaneously), subject to a decreasing probability (recency effect).

It is logically possible that some material is cued (in e.g., recognition memory, paired associates, and serial recall) and other material (in free recall) is retrieved spontaneously in the absence of an effective cue. But this is to model the experiment, rather than memory. Memory does not start up at the beginning of an experiment; it is already up and running before ever the participant gets to the experimental chamber. An experiment on memory is therefore an imposition on an already ongoing process. Recall is a function of the experiment – different experiments deliver different kinds of data – but retrieval belongs to the underlying process and, as a metatheoretic principle, should be effected in the same manner for all experimental paradigms.

There seems, nevertheless, to be a wide-spread belief that retrieval from memory is cued. But a reviewer cannot object purely on the basis of belief. There has to be some evidence. Hence,

... in the Shiffrin [1970] recall list N-1 paradigm, something has to be focusing participants on the list before last. If it was purely about accessibility relative [sic] based on temporal order then intrusions of the Nth list would have been catastrophic.

## 5. APPENDIX

A RANDOMISATION TEST OF THE SUPPRESSION OF REPEATED RECALLS<sup>6</sup>

Suppose there are N (= 720) lists of recalls and, amongst those lists, r occurrences of some word W. Assuming those occurrences to be labelled (so that different recalls of W are distinguishable), there are  $N^r$  different possible arrangements of the r recalls amongst the N lists. Let m be the number of recall lists containing at least one recall of W; then (r-m) of the r recalls are repetitions.

If there is no suppression of a response that has already been uttered, then all  $N^r$  arrangements are equally probable; but if, once word W has been recalled, the probability of its being uttered a second time in the same list is reduced, then the

<sup>&</sup>lt;sup>5</sup> The predictions are not exact because individual words are commonly rehearsed several times during presentation and one cannot know which particular rehearsal is the source of the present recall.

<sup>&</sup>lt;sup>6</sup> From Laming, D. Failure to recall. *Psychological Review*, 116(1), p. 157–186, 2009, APA. Adapted with permission.

number of repetitions (r-m) will be less than would otherwise occur by chance and the number m of lists containing at least one recall of W will be increased. The test which follows is based on a calculation of the chance probability that m or more of the N lists contain at least one of the r recalls of W under the null hypothesis that there is no suppression of repeated recalls.

Let  $A_m(r,N)$  be the number of arrangements in which exactly *m* lists are occupied. Any such arrangement can be obtained by adding an *r*th recall of W to an existing arrangement of (r-1) recalls. A total of *m* lists will be occupied by *r* recalls if *either* the *r*th recall is added to one of *m* lists already occupied by *r*-1 recalls *or* if it is added to one of the (N-m+1) lists left unoccupied in an arrangement of *r*-1 recalls in *m*-1 lists. Hence

$$A_m(r,N) = mA_m(r-1,N) + (N-m+1)A_{m-1}(r-1,N)$$
(A1)

This recurrence relation is satisfied by

$$A_m(r,N) = \begin{pmatrix} N \\ m \end{pmatrix} \sum_{k=0}^m (-1)^k \begin{pmatrix} m \\ k \end{pmatrix} (m-k)^r$$
(A2)

a result which can be checked by substitution in Equation A1.

Under the hypothesis that there is no suppression of repeated recalls, all  $N^r$  arrangements are equally probable, and the probability  $p_m(r,N)$  that exactly *m* lists are occupied (*i.e.*, that there are exactly *r*-*m* repetitions of *W*) is equal to  $N^{-r}A_m(r,N)$ . The tail probability of *r*-*m* or fewer repetitions in *r* recalls is therefore

$$\sum_{k=m}^{r} p_k(r,N) = \sum_{k=m}^{r} N^{-r} A_k(r,N)$$
(A3)

For computational purposes it is convenient to note that the number of arrangements of r recalls in exactly m out of N lists is equal to the number of ways of selecting the m lists (N!/m!(N-m)!) times the number of ways of arranging the r recalls within the m lists such that every list is occupied. This second factor is independent of N. Therefore put

$$p_m(r,N) = N^{-r}A_m(r,N) = N^{-r}(N)_m f(r,m)$$
(A4)

where f(r, m) is independent of N. After substitution in Equation A1,

$$f(r,m) = mf(r-1,m) + f(r-1,m-1)$$
(A5)

The quantity f(r, m) is the number of ways of assigning r distinct recalls to m lists, such that every list is occupied, divided by m!. It is equal to 1 if either m = r or m = 1. Values for other integer arguments are most easily calculated by iteration from Equation A5, the  $p_m(r, N)$  from Equation A4, and the tail probabilities (A3) by summation of the  $p_m(r, N)$ .

#### BIBLIOGRAPHIE

ANDERSON J. R., BOWER G. H. (1972). "Recognition and retrieval processes in free recall", *Psychological Review* 79, p. 97-123.

BRODIE D.A., MURDOCK B.B. Jr. (1977), "Effect of presentation time on nominal and functional serial-position curves of free recall", *Journal of Verbal Learning and Verbal Behavior* 16(2), p. 185–200.

HARBISON J.I., DOUGHERTY M.R., DAVELAAR E.J., FAYYAD B. (2009), "On the lawfulness of the decision to terminate memory search", *Cognition* 111(3), p. 397-402.

HOWARD M.W., KAHANA M.J. (2002), "A distributed representation of temporal context", *Journal of Mathematical Psychology* 46, p. 269-299.

JANG Y., HUBER D.E. (2008), "Context retrieval and context change in free recall: Recalling from long-term memory drives list isolation", *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34(1), p. 112-127.

LAMING D. (2006), "Predicting free recalls", *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32(5), p. 1146–1163.

LAMING D. (2008), "An improved algorithm for predicting free recalls", *Cognitive Psychology* 57(3), p. 179-219.

LAMING D. (2009), "Failure to recall", Psychological Review 116(1), p. 157-186.

LAMING D. (2010), "Serial position curves in free recall", Psychological Review 117(1), p. 93-133.

MENSINK G., RAAIJMAKERS J.G.W. (1988), "A model for interference and forgetting", *Psychological Review* 95, p. 434-455.

MURDOCK B.B. Jr., OKADA R. (1970), "Interresponse times in single-trial free recall", Journal of Experimental Psychology 86(2), p. 263-267.

RAAIJMAKERS J.G.W., SHIFFRIN R.M. (1980), "SAM: A theory of probabilistic search of associative memory", in G. Bower (ed.), *The psychology of learning and motivation*, vol. 14, New York (NY), Academic Press.

SHIFFRIN R.M. (1970), "Forgetting: Trace erosion or retrieval failure", Science 168(3939), p. 1601-1603.

UNSWORTH N., SPILLERS G.J., BREWER G.A. (2012), "Evidence for noisy contextual search: Examining the dynamics of list-before-last recall", *Memory* 20(1), p. 1-13.

WARD G., TAN L. (2004), "The effect of the length of to-be-remembered lists and intervening lists on free recall: A reexamination using overt rehearsal", *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30(6), p. 1196-1210.