Clustering effects on the recall
of unrelated words

MARILYN A. BORGES, JOSEPH R. LEVINE, ELLEN M. LeVITA,
and APRIL M. McTAGGERT
San Diego State University, San Diego, California 92182

The purpose of this experiment was to examine directly how various numbers of clusters affect recall performance. Thirty subjects were presented 36 unrelated nouns for six study-recall trials. Each of six subjects was told to organize the words into 3, 7, or 11 clusters. Two control groups were also used: free recall (FR; no clustering instructions) and free cluster (FC; clustering instructions, but no specific number of clusters were assigned). The similarity between total recall and all other measures for the three-cluster, FC, and FR groups suggests that organizing the words into approximately three clusters optimizes recall performance. This study also supports a previous finding that recall increases over trials mainly due to increases in cluster size rather than increases in number of clusters.

Based on numerous experiments in which subjects repeatedly sorted unrelated words into various sizes and number of clusters (piles) prior to a free recall task, Mandler (1967) has suggested that the optimum number of clusters for maximum learning in a free recall task is seven (5 ± 2), with seven (5 ± 2) items composing each cluster. The linearly increasing relationship between number of clusters (up to size seven) and recall found in the sorting experiments held both for subjects determining their own number of clusters and for subjects constrained to using a presassigned number of clusters.

Although number of clusters or categories is regularly measured in free recall tasks involving categorized word lists, unrelated word lists that are subjectively and idiosyncratically organized by subjects have generally been considered unamenable to such measurements. However, Buschke (1977) recently showed that even with unrelated words subjective organization (or number and size of clusters) can be measured in an unrelated-word free recall task by having subjects write their recall in clusters on a two-dimensional (2D) grid.

After 10 study-recall trials of an unrelated 32-item list, the mean number of clusters used by Buschke's (1977) 10 subjects on the last recall trial was nine, with 3 additional unclustered items. He found these subjects' "2D-clustered" recall to be very similar to the recall of a group of subjects given standard free recall instructions. But Buschke's experiment did not provide a test of Mandler's (1967) suggestion concerning the number of clusters/recall relationship, since the number of clusters was not systematically manipulated.

In the present experiment, a modified version of Buschke's (1977) 2D technique was used, and the number

of clusters that subjects used to organize their unrelated word list was systematically varied. These data, then, should provide a direct test of the number of clusters/recall relationship proposed by Mandler (1967).

METHOD

Design
Five groups of subjects were presented 36 unrelated nouns for six study and recall trials. Three groups of subjects were constrained to organize the words into 3, 7, or 11 clusters. Two control groups were also used: free recall (FR; no clustering instructions) and free cluster (FC; clustering instructions, but no specific number of clusters were assigned).

Subjects
Thirty undergraduates from San Diego State University participated as subjects. Six subjects were randomly assigned to each of the five recall group conditions.

Materials
Thirty-six nouns with frequencies of occurrence ranging from 103 to 319/million were selected from the Kučera and Francis (1967) norms.

Procedure
All subjects were told the purpose of the experiment was to learn the list of 36 words. They were also informed that the usual method people use to learn a list of words is to group or organize the words into smaller subgroupings of items that "seem to go together." The subjects in the three constrained conditions were further instructed to write the words in 3, 7, or 11 groups or clusters at recall. FC subjects were also told to write their recall in groupings, but no particular number was assigned or suggested. All cluster subjects were told they could arrange the word groupings on their recall sheets however they wished. FR subjects were told just that they would be writing down all the words they could remember, in any order they wished, after each list presentation.

The 36 test items were presented on slides at a 4-sec rate. They were presented in a different random order on each of the six study trials, with all subjects seeing the same ordering of words per study trial. After list presentation, subjects were given a blank sheet of paper for recall, and the constrained

clustering subjects were again told the particular number of groupings they were to use. All subjects said each word aloud as they wrote it during recall, and their responses were tape-recorded. Each recall trial was terminated after 4 min, and the next presentation or study trial immediately followed. The same procedure was followed for all six trials.

RESULTS

Total Recall

An analysis of variance on the total number of words recalled by each subject per trial indicated that both trials and groups were significant main effects [$F_{(5,125)} = 303.72$ and $F_{(4,25)} = 3.17$, respectively; both $ps < .05$]. The interaction of trials and groups was not significant. As shown in Figure 1, mean total recall increased across groups in the following order: The most words were recalled by the constrained three-cluster group, the FC group, and the FR group. The means for these groups were greater than the 7-cluster group, who recalled more than the constrained 11-cluster group.

Numbers of Clusters and Items per Category

A cluster was defined as the number of separate word groupings used by each subject per group per trial. The FR group did not physically separate their recalled words into clusters; thus, their number of clusters was not considered here. An analysis of variance showed that trials, groups, and Trials by Groups were all significant effects [$F_{(5,100)} = 31.54$, $F_{(3,20)} = 120.30$, and $F_{(15,100)} = 7.70$, respectively; all $ps < .05$]. As shown in Figure 2, by Trial 3, Groups 3, 7, and 11 were using approximately the number of clusters they were instructed to use. Of particular note was that the FC group chose to use a relatively constant number of

Figure 2. Mean number of clusters used over trials for four of the subject groups. (Groups 3, 7, and 11 were constrained to use the designated number of clusters; FC is the free clustering group.)

clusters (mean = 4.75) on the last four of the six trials. Examination of Figure 2 shows that the interaction of Groups by Trials was solely due to Groups 7 and 11's using fewer than the assigned number of clusters on the first two trials.

The number of items per cluster that was recalled by each subject per trial was computed. An analysis of variance on the items per cluster data showed trials, groups, and Trial by Groups to be significant effects [$F_{(5,100)} = 27.16$, $F_{(3,20)} = 76.60$, and $F_{(15,100)} = 15.79$, respectively; all $ps < .05$]. Again, the number of items per cluster for the FR group was not considered, since their recall was not physically separated into clusters. The number of items per cluster for Groups 3, 7, and 11 was constrained by their assignment of a particular number of clusters. Since total recall is equal to the number of clusters multiplied by the number of items per cluster, the increases in total number of words recalled by the 3-, 7-, and 11-cluster groups after attaining their assigned number of clusters on Trial 3 were due strictly to increases in number of items per cluster. Of particular note was that since the FC group's number of clusters were approximately constant over the last four trials, their increases in total number of recalled words were also due to increases in the number of words per cluster.

List Organization

If the list was perfectly organized, it would be expected that subjects would recall all items from one cluster, then all items from another cluster, and so on. That is, no words from a given cluster would be retrieved separately from all the other words of the cluster. The actual output order of each subject's recalled words (which had been tape-recorded) was compared with the subject's written clusters. The number of times each written cluster was initiated more than once was computed for each subject per trial. An analysis of variance showed
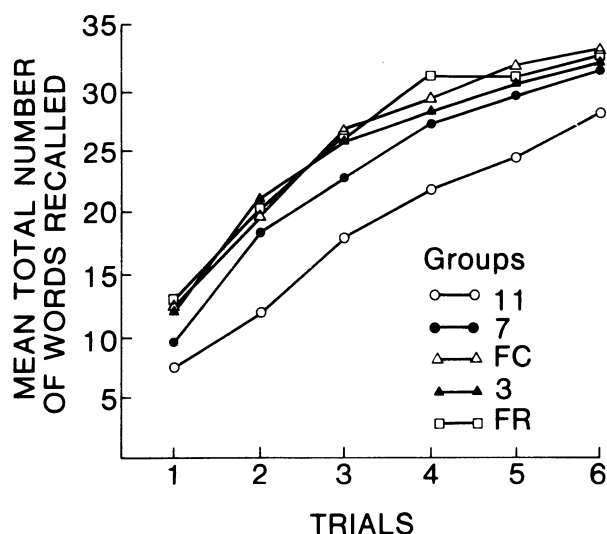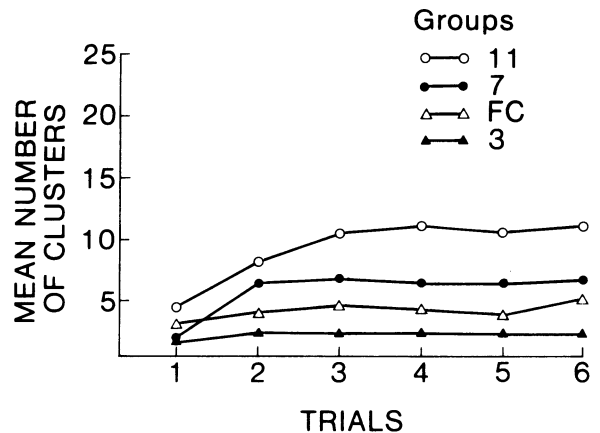

Figure 1. Mean total recall over trials for each of the five subject groups. (Groups 3, 7, and 11 were constrained to use the designated number of clusters; FC is the free clustering group, and FR is the standard free recall group.)

that these "reorderings" did not differ among the four clustering groups compared, but that reorderings did increase from a mean of 1 per subject on Trial 1 to 4.7 per subject on Trial 6 [$F(5,100) = 7.57$, $p < .05$].

However, since the reordering measure does not take into account that more reorderings are possible when more clusters are used, RRR measures (Bousfield & Bousfield, 1966) were also computed to equate the different numbers of clusters per group. "It is the ratio of the actual proportion of repetitions [Total repetitions (R)/Words recalled (N) − 1], over maximum possible RR [Number of words recalled (N) − Number of categories recalled (N_c)/Words recalled (N) − 1], which reduces to RRR = R/(N − N_c)" (Mandler, 1969, p. 533). The RRR measures showed that constrained Group 11's recall was less organized than that of Groups 3, 7, and FC. The RRR means for Groups 3, 7, and FR were .83, .79, and .79; Group 11's mean was .69.

DISCUSSION

The main purpose of the present study was to test Mandler's (1967) hypothesis that total recall would increase as the number of clusters in which subjects encoded items increased from three to seven and that recall would decrease if either more than seven clusters were used or the size of the clusters exceeded seven. The results of the present experiment did not support his prediction. The subjects constrained to using 3 clusters recalled more than subjects constrained to using 7 clusters, although subjects using 11 clusters did recall fewer items than either of the two other groups. The superiority of the three-cluster group was unexpected, since their clusters contained approximately 11 words by Trial 6—many more than Mandler's upper limit of 7 words per cluster. In fact, the likelihood that specific words recalled on one trial would be forgotten on the following trial did not increase with the number of items per cluster ($p > .90$).

The lack of any significant differences between the constrained three-cluster group's total recall and that of the FC and FR groups' total recall suggests that the clustering or constraining conditions per se did not significantly affect the above recall results. In fact, Mandler (1967) found his number of clusters/recall relationship to hold both when subjects were constrained to use a particular number of items per cluster and when subjects were free to choose their own number of clusters or categories.

Mandler (1967) does suggest that when clusters are very large, subjects may break up the clusters into smaller sequential subclusters. (These subclusters then act like additional separate clusters in affecting recall.) Although such subclustering could account for some of the constrained Group 3's superior recall

performance, it does not explain why Group 3's recall exceeded that of the "hypothesized-optimal" Group 7's recall. Additionally, Mandler states that clusters up to size 10 to 15 are better considered as single clusters and only past that size is subclustering likely to occur. Thus, although Mandler's number of cluster/recall relationship certainly describes the recall of words well organized by several prior sorting trials, the present results suggest that his number of cluster relationship does not generalize to the standard free recall paradigm.

Certainly much evidence from free recall studies using both unrelated and categorized word lists supports the supposition that subjects organize their words into clusters. But the present results suggest that in a free recall task of unrelated words, the optimum number of clusters may be approximately three to five. (The FC group's recall was similar to Group 3's, and the FC group used approximately five clusters.) However, as the number of clusters exceeds five, recall decreases monotonically as the number of clusters increases. (Recall was highest for Groups 3 and FC, lower for Group 7, and lowest for Group 11.) Group 3's recall results also suggest that up to 11 words can be encoded in a single cluster without inhibiting recall performance.

The present results support Buschke's (1977) finding that increases in recall can be mainly attributed to increases in the number of items per category rather than to increases in the number of clusters used. On Trials 3-6, the FR group's total recall did not differ from that of the FC group (who chose to use a relatively constant number of clusters) or that of the constrained Group 3 (who were constrained to use a constant number of clusters). Thus, these groups' increases in number of recalled words over the last four trials can be attributed to increases in the number of words encoded per cluster, not to increases in the number of clusters used.

REFERENCES

Bousfield, A. K., & Bousfield, W. A. Measurement of clustering and of sequential constancies in repeated free recall. *Psychological Reports*, 1966, **19**, 935-942.

Buschke, H. Two-dimensional recall: Immediate identification of clusters in episodic and semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 201-215.

Kučera, H., & Francis, W. N. *Computational analysis of present-day American English*. Providence, R.I: Brown University Press, 1967.

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory*. New York: Academic Press, 1967.

Mandler, G. Input variables and output strategies in free recall of categorized lists. *American Journal of Psychology*, 1969, **82**, 531-539.