THE OCCURRENCE OF CLUSTERING IN THE RECALL OF RANDOMLY ARRANGED WORDS OF DIFFERENT FREQUENCIES-OF-USAGE

W. A. Bousfield

B. H. Cohen

Technical Report No. 2

Office of Naval Research Contract Nonr--631 (00)

UNIVERSITY OF CONNECTICUT

STORRS, CONNECTICUT

September, 1952

The Occurrence of Clustering in the Recall of Randomly Arranged Words of Different Frequencies-of-Usage

In an earlier study the writers (1) undertook to determine the influence of reinforcement on the incidence of clustering occurring in the recall of a randomly arranged list of stimulus-words comprising four categories of items. Clustering was defined as the occurrence of sequences of two or more words in the same category. Reinforcement was controlled on the basis of the number of presentations of the stimulus-words during the learning period. It was found that under these conditions reinforcement was effective in producing predictable changes in clustering. Logically we may state that the purpose of the present study is also to investigate the influence of reinforcement on the clustering of verbal responses in recall. As generally used, the term reinforcement refers to an action or condition having as one of its results the strengthening of habits. In dealing with the learning of words in a laboratory situation, it is necessary to recognize that varying amounts of reinforcement have taken place prior to such additional reinforcement as may be added during the controlled learning. In our earlier experiment, we dealt with the results of imposing varying amounts of additional reinforcement during a controlled learning situation. The present study differs in that it undertook to make use of the varying degrees of prior reinforcement. This was done on the basis of selecting words for the stimulus-lists according to their frequency-of-usage counts. These frequencies were obtained from the Thorndike-Lorge tables (3). We recognize that the use of word-counts from written material as an index of their prior reinforcement is a somewhat crude undertaking. It appears empirically justified, however, and has the advantage of operational precision.

While we may specify different operations for achieving reinforcement, we can assume that the measurable effects on resultant habit strength remain essentially the same regardless of how the strength was acquired. On this assumption we undertook to test the same hypotheses relating to the effects of reinforcement on clustering as were proposed for the earlier study. These hypotheses were stated as follows:

1. The total amount of clustering in the recall of randomized verbal associates should be a positive function of the degree of their reinforcement.

2. The progressive changes in clustering from the beginning to the end of recall of a group of randomized associates should be modified by reinforcement in the following ways:

    a. The initial level of clustering should be a positive function of reinforcement.

b. There should be a positive relationship between reinforcement and the speed of attainment of maximum clustering.

c. For all levels of reinforcement, clustering should tend to decrease as the supply of available associates approaches exhaustion.

METHOD

Subjects. The subjects were 150 undergraduate students in introductory psychology enrolled in six laboratory sections ranging in size from 22 to 29 members. They were divided into two experimental groups, each comprising the students in three laboratory sections, and each having a total of 75 students.

Stimulus-Material and Apparatus. Two lists of stimulus-words were prepared from the Thorndike-Lorge tables (3) to meet the following specifications: (a) each list was to comprise 60 nouns with 15-each in the following four categories, namely, animals, names, professions, and vegetables; (b) one list was to contain words of high frequencies, and the other, words of low frequencies; (c) the mean word-frequency for each category should be the same within each list; (d) for each category the words in each list should match as closely as possible with respect to number of syllables. We had originally hoped to employ a three-fold frequency range, but this proved to be impossible because of our rather rigid specifications for the construction of the lists. The two lists of stimulus-words appear in Table 1 along with their syllabic counts and Thorndike-Lorge frequencies. The items from each of the two lists were randomized and transferred in their random orders to slides. To expose the words on a screen, we used a projector with a movable mask. The words were projected singly and at the rate of three seconds per word. Mimeographed data sheets were prepared for use by the subjects in writing the words they were able to recall. These sheets were 8½" X 11" in size and divided into five columns numbered in consecutive order.

Procedure. As already indicated, the subjects were divided into two experimental groups, each with three sub-groups. One experimental group received the low frequency stimulus-words and the other the words with high frequencies. The experiment was conducted in the small lecture room where the students in each sub-group assembled during the scheduled period. The procedure conformed with that described in detail in the earlier study (1), and was as follows. The experimenter distributed the data sheets and requested the cooperation of the subjects. He informed them that a list of words would be projected on the screen one at a time and that following the projection they were to write as many words as they could recall in the order in which they were recalled. He explained further that at a given signal they were to start writing in the first column of the data sheets and that at intervals they would be told to shift to the next column and continue with more words. The experimenter then went to the projector and exposed the complete list

TABLE 1

Lists of Stimulus-words

A. Low Frequency List

Animals			Names			Professions			Vegetables		
Word	Syls.	Freq.	Word	Syls.	Freq.	Word	Syls.	Freq.	Word	Syls.	Freq.
stag	1	7	Rex	1	2	chef	1	2	kale	1	1
mink	1	5	Clyde	1	3	plumber	2	2	cress	1	2
ape	1	6	Saul	1	4	typist	2	1	squash	1	6
doe	1	6	Luke	1	2	milkman	2	1	chard	1	1
walrus	2	2	Wesley	2	1	draftsman	2	1	parsnip	2	1
giraffe	2	1	Amos	2	1	brakeman	2	1	eggplant	2	1
bobcat	2	1	Rufus	2	1	rancher	2	1	rhubarb	2	3
poodle	2	1	Gerald	2	1	postman	2	4	garlic	2	3
gazelle	2	1	Harvey	2	1	broker	2	67	radish	2	4
llama	2	1	Wallace	2	7	jockey	2	3	melon	2	5
bullfrog	2	1	Byron	2	3	printer	2	7	mustard	2	7
zebra	2	2	Angelo	3	1	machinist	3	1	horseradish	3	1
oppossum	3	1	Everett	3	1	caretaker	3	1	broccoli	3	1
hyena	3	2	Theodore	3	5	bookkeeper	3	3	artichoke	3	1
kangaroo	3	2	Achilles	3	6	contractor	3	5	cauliflower	4	2
Means	1.93	2.60		2.00	2.60		2.20	2.60		2.07	2.60

TABLE 1*

(Continued)

B. High Frequency List

Animals			Names			Professions			Vegetables		
Word	Syls.	Freq.	Word	Syls.	Freq.	Word	Syls.	Freq.	Word	Syls.	Freq.
wolf	1	A	James	1	A	priest	1	42	beet	1	11
skunk	1	13	Ralph	1	21	chauffeur	2	9	pea	1	30
coon	1	13	Roy	1	14	fireman	2	9	bean	1	43
colt	1	21	Tim	1	11	dentist	2	9	corn	1	A
lion	2	A	David	2	46	baker	2	10	carrot	2	9
donkey	2	16	Duncan	2	10	archer	2	11	turnip	2	10
otter	2	8	Milton	2	9	grocer	2	11	mushroom	2	10
weasel	2	9	Gordon	2	9	waiter	2	13	lettuce	2	12
badger	2	11	Leonard	2	10	surgeon	2	16	pumpkin	2	13
turtle	2	13	Moses	2	12	shepherd	2	32	cabbage	2	16
camel	2	18	Peter	2	A	artist	2	A	pepper	2	27
monkey	2	23	Benjamin	3	11	editor	3	A	celery	3	9
coyote	3	9	Columbus	3	29	explorer	3	10	spinach	2	8
buffalo	3	19	Samuel	3	15	inventor	3	14	tomato	3	11
elephant	3	35	Abraham	3	11	policeman	3	22	potato	3	A
Means	1.93	23.87		2.00	23.87		2.20	23.87		1.93	23.93

* Where A is listed as the frequency of a stimulus word, it indicates a Thorndike-Lorge frequency-of-usage count of between 50-99 occurrences per million words. In computing the means, all A's were assigned a frequency of 75.

of words once. Three seconds after the final word had appeared on the screen, the subjects were told to start writing the words they were able to recall. The signals for shifting to succeeding columns were given at two-minute intervals, thus allowing a total of ten minutes for recall. In this way each of three sub-groups received the list of high frequency words and three sub-groups received the low frequency words.

RESULTS

Preparation of Data. Several steps were taken in order to facilitate the statistical analysis of the data. First, the words appearing in each list of recalled items were numbered in consecutive order. Each word was then labelled as follows to indicate its classification: A, for animals; N, for names; P, for professions; V, for vegetables; I, for words which were illegible, or if legible, could not be classed in any one of the four experimental categories. Items which were classed as A, N, P or V, but did not appear in the list of stimulus-words, were given the additional label NL. This classification identifies two types of errors, or what we would prefer to term intrusions. We shall here refer to the NL-words as categorical intrusions and the I-words as irrelevant intrusions. Having labelled all items according to this scheme, all sequences of two or more words in the same category were bracketed to indicate clustering. In doing this we made no discrimination against the NL-words. This is consistent with our definition of clustering and the existence of such intrusions is evidence of the potency of the organizing tendency under investigation.

The analyses made of our data followed in general the plan of the earlier study (1).

Relationship between Word-Frequencies and the Mean Number of Words Recalled. If we have correctly assumed that word frequency-of-usage counts may serve as an index of reinforcement, it would follow that with the opportunity to memorize held constant, the amount of recall for the high frequency words should be greater than that for the words of low frequencies. This expectation was verified in the data. For the high frequency words the mean number recalled was 25.55 with a standard deviation of 5.32. The corresponding results for the low frequency words were 22.18 and 5.95. The critical ratio of these means is 3.66. This indicates that the mean number of high frequency words recalled was about 15 per cent higher than the mean for the words of low frequency and that this difference is statistically reliable.

Relationship between Word-Frequencies and Intrusions Appearing in Recall. We have defined two types of intrusions or errors appearing in the lists of recalled items. Categorical intrusions were words not given in a stimulus-list, but classifiable as belonging to one of the experimental categories. The mean number of such intrusions per subject for the high frequency words was 1.27; for the low frequency words the mean was 1.37.

If we designate these intrusions in terms of percentage of total items recalled, we find that 5.4% of the high frequency items and 5.7% of the low frequency items were categorical intrusions.

Irrelevant intrusions comprised all items not classifiable as belonging to an experimental category. The mean number per subject of these less frequent intrusions was 0.45 for the high frequency words and 0.73 for those of low frequency. Expressed as percentages of the total items recalled we have 2.0% for the high frequency words and 3.0% for those of low frequency.

We would hesitate from these data to assume a relationship between the incidence of intrusions and the word-frequencies of our two lists of stimulus-words. It may be noted, however, that the incidence of categorical intrusions is greater than that of the irrelevant type. It also appears that intrusions of both types constitute a relatively small proportion of the data.

Relationship between Word-Frequencies and Total Clustering as Indicated by the Index of Repetition. The first of our experimental hypotheses predicts greater total clustering for the high frequency words than for those of low frequencies. To test this hypothesis our first step was to employ the formula for the index of repetition, IR, which we had also used in the earlier study (1). This index shows the extent to which the incidence of clusters exceeds chance expectation. Its statement is as follows for our four-category lists of items:

$$IR = \frac{4r}{\sqrt{3n}} - \sqrt{\frac{n}{3}} \qquad (1)$$

In this formula the term r represents the number of repetitions of words in any category; n is the total number of sequences in a list of recalled items, or the total number of items minus one. The advantage of this formula is that the index it provides is based on the assumption of a distribution with unit standard deviation about a mean of zero. We computed the IR for each of our 150 lists of recalled items. For the 75 lists based on the recall of low frequency words the mean IR was 1.14 with a standard deviation of 1.52. For the 75 lists based on the recall of high frequency words, the mean IR was 1.92 with a standard deviation of 1.94. The critical ratio of these means is 2.76. We may conclude that our experimental hypothesis receives reasonable support from these results.

Relationship between Word-Frequencies and the Number of Items Occurring in Clusters of Various Size. This analysis was undertaken as a supplement to that of the IR. It is based on the assumption that the amount of clustering should relate positively to the number of items appearing in the higher order clusters. The results of this treatment appear in Table 2 where we also show the size of clusters based on chance expectation. The figures for chance were derived from an artificial experiment involving the drawing of capsules in four colors without replacement. The largest clusters we obtained from nearly 2500 drawings of capsules for 100

TABLE 2

Relationship between Word-Frequencies and the Mean Number of Items Occurring in Clusters of Varying Size*

Word Frequencies		Size of Clusters									
		Single Items	2	3	4	5	6	7	8	9	10
Low	Subjects	9.88	5.68	3.52	1.76	0.53	0.40	0.19	0.11	0.12	—
	Chance	12.90	6.09	2.32	0.64	0.18	0.05	—	—	—	—
High	Subjects	9.72	5.73	4.44	2.03	1.53	0.72	0.65	0.21	0.24	0.27
	Chance	14.86	7.02	2.67	0.74	0.20	0.06	—	—	—	—

* The minimal size of a cluster is two items. The means in the first column are therefore for single (unclustered) items.

hypothetical subjects did not exceed six items. The figures for chance in Table 2 have been adjusted to the number of items recalled because the number of words recalled was higher for the high frequency list than for the low frequency list. It is evident that these results confirm those obtained from the IR.

Relationship between Word-Frequencies and Progressive Changes in Clustering During Recall. Our experimental hypotheses predicted several modifications in the progression of clustering from the beginning to the end of the period of recall. In comparison with the progression of clustering for the recall of low frequency words, the progression for the high frequency words should show a higher initial level of clustering and quicker attainment of maximum clustering. Furthermore, the clustering in both cases should drop in the direction of randomness as the supply of recallable items approaches exhaustion and the habit strengths of these words become reduced. In order to test these hypotheses, it was necessary to employ an index capable of representing the progressive changes in the amount of clustering during recall. It was also necessary to employ a suitable method of dividing the lists of recalled words into comparable units. To indicate the degree of progressive clustering we used the density ratio, which we define as the decimal fraction obtained by dividing the number of clustered items in a specified group of items by the total number of items in the group. We used the Vincent method in dividing all lists of recalled words into successive decile groups. The density ratios for the successive deciles of items for both the low frequency and the high frequency words are shown in Table 3. The curves for these data appear in Figure 1 along with a reference line for the density ratios expected on the basis of chance. This line represents the mean density ratio of .417 which was obtained for the successive deciles of items in the artificial experiment to which reference has already been made. Inspection of the two progressions of density ratios derived from our data on recall reveals general confirmation of the experimental hypotheses. In comparison with the density ratios for the low frequency words, those for the high frequency words show a higher initial level (.523 as compared with .351), and a quicker attainment of maximum clustering (in the 3rd decile as compared with the 5th). We also observe that the curve for each stimulus-list shows an approach toward randomness in the higher deciles where the subjects were also approaching exhaustion of their supplies of words available for recall.

DISCUSSION

We regard this study as being closely related to the one previously reported (1) in that both dealt with the influence of reinforcement on the incidence of clustering in the recall of randomly arranged words. Both employed the same basic methods for inducing and measuring clustering. Both tested the same hypotheses. They differed, however, in the techniques used for controlling reinforcement. The paradigm of what we are dealing with is the experimental situation in which a subject learns a list of meaningful but randomly presented words and subsequently recalls as many as he can. The order in which our subject recalls the

TABLE 3

Relationship between Word-Frequencies and the Density

Ratios per Decile of Words Recalled

Word Frequencies	Decile									
	1	2	3	4	5	6	7	8	9	10
Low	.351	.561	.684	.665	.713	.612	.540	.577	.528	.317
High	.523	.646	.761	.731	.741	.656	.637	.581	.471	.461

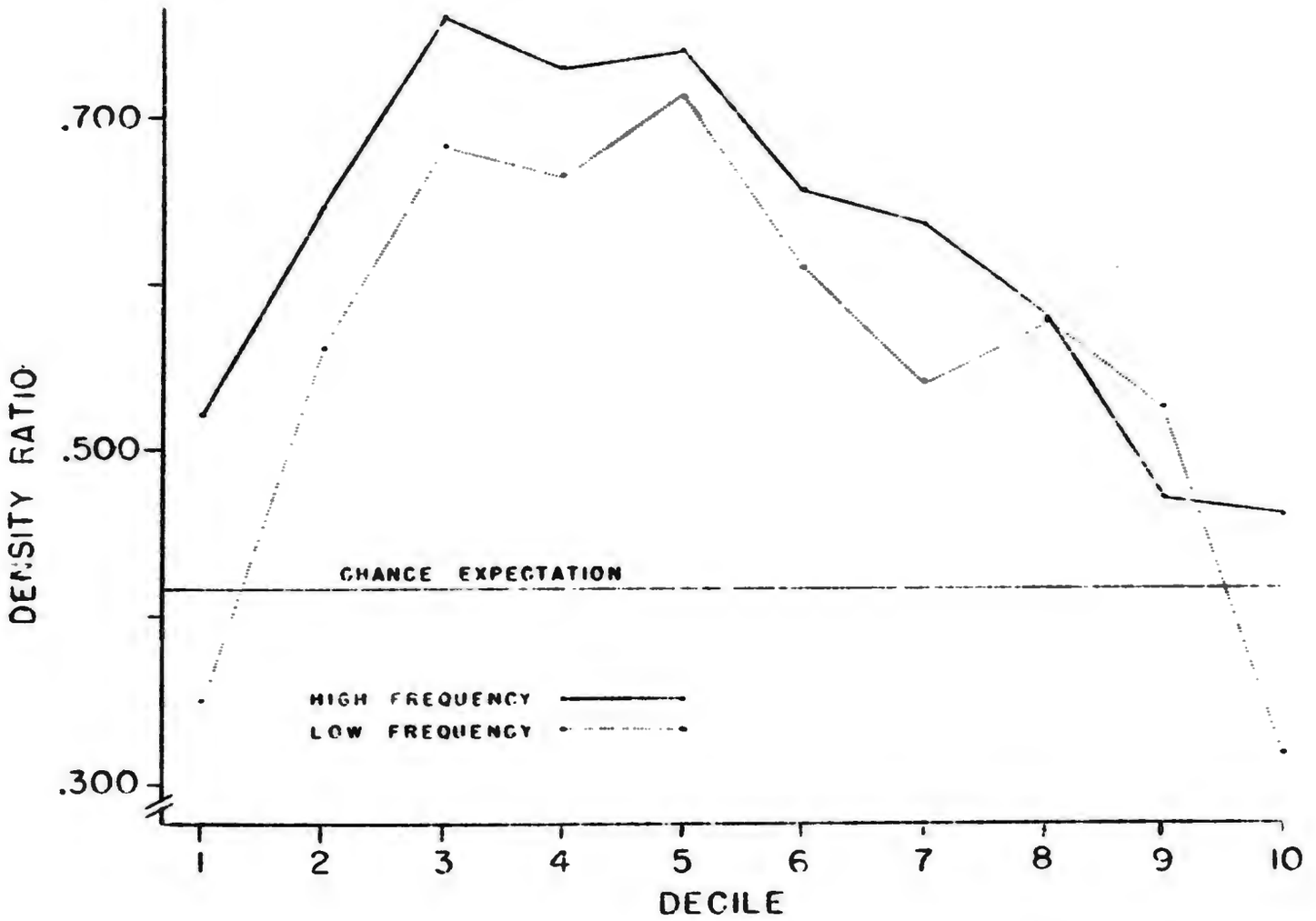

FIGURE 1

PROGRESSIVE CHANGES IN CLUSTERING DURING THE RECALL OF WORDS AS A FUNCTION OF THEIR FREQUENCIES-OF-USAGE

words reflects not only their habit strengths, but also a process of organization which we infer from the tendency to produce sequences of related words. Our technique permits a type of quantification of the amount of organization that takes place. Reinforcement, as we conceive it, should be recognized for experimental purposes as occurring at three separate phases of our paradigm situation, namely: the pre-experimental, experimental, and recall phases. During the pre-experimental phase the words as verbal habits have received varying amounts of reinforcement through differential use. Additional reinforcement of the words occurs in the process of memorization during the experimental phase. It is also necessary to assume that still further reinforcement is added to the words produced during the recall phase. In terms of this analysis the present study dealt with differences in reinforcement during the pre-experimental phase. The earlier study undertook to manipulate reinforcement during the experimental phase.

To account for the phenomena of clustering, it appears necessary to go beyond the building up of habit strengths of isolated word-habits. Such habit strengths should be distributed randomly with respect to the categories. As shown in our earlier work it was necessary to postulate higher order mediating processes capable of adding to the strength possessed by the individual word-habits. To develop this conception we have relied heavily on Hebb's (2) interpretation of the development of superordinate perceptions. In essence, Hebb proposes that the repeated arousal of a group of related subordinate perceptions has as one of its major consequences the development of a superordinate structure whose activity constitutes the perception of the whole. As a result of such learning the occurrence of a single subordinate perception may activate the superordinate structure, which in turn will facilitate the responses of its subordinates. Transferring this notion to our experimental situation, we would regard the perceptions of the words as subordinate perceptions. The categories of our words correspond to superordinate structures. Thus, for example, the occurrence of the word leopard in recall results in the activation of its superordinate, namely, that corresponding to animal. This in turn facilitates its subordinates, and increases the likelihood of the subjects naming other animals. The strengthening of both the subordinate as well as the superordinate structures should occur at each of the three phases of our experimental situation. It happens, however, that during recall the subject tends to give first the words having relatively high habit strength. As recall proceeds, the subject draws on words with progressively lower habit strength. We thus assume that the superordinate structures gain strength during recall while the habit strengths of the specific words decline. The range of possible growth of the superordinate structures during recall depends on their levels of development at the beginning of recall. If the superordinate structures had already developed their maximum strengths this range would be zero. On the other hand, if the superordinate structures were initially weak, the range of possible growth would be large. We next assume that the potency of the superordinate structures to facilitate effectively their subordinates depends on the strengths of the subordinates. As long as the subordinates are of

optimal strength, any growth that takes place in the superordinates will result in an increase in the incidence of clustering. Below this optimal level the superordinates lose progressively their effective facilitating influence. Our experimental hypotheses are all derived from these assumptions. At the beginning of recall, the strengths of both the subordinate and the superordinate structures for the high frequency words were stronger than was the case for the low frequency words. This difference would favor greater general clustering for the high frequency words and also greater initial clustering. Maximum clustering should be attained more rapidly during the recall of high frequency words because their superordinate structures were nearer their maximum strengths. In both instances, however, the subjects were permitted to approach exhaustion of their supplies of words for recall and also at this time had available only the words of increasingly low habit strength. The result was an approach toward randomness in the incidence of clustering for both types of words.

We may take stock of this theorizing and indicate its essential points. In order to account for the organization indicated in clustering it appears necessary to go beyond the postulated habit strengths of individual words. We must further assume the functioning of higher order mediating processes which we have regarded as dependent on superordinate structures. Such structures develop through reinforcement and act to facilitate their subordinate word-habits. Thus memory, which in these terms amounts to the reactivation of both subordinate and superordinate structures, is conceived of as being selective in its operation. If our analysis is valid, we have been dealing not only with organization as a function of higher mental processes, but we have also been dealing with the problem of meaning. The occurrence of meaning would depend on the activation of superordinate structures. Our interpretation has relied heavily on Hebb. The best we can hope for is validation and extension of this type of interpretation. The next best is an alternative interpretation which will better subsume our experimental results.

SUMMARY

This study was undertaken with the purpose of investigating the incidence of clustering in the recall of two stimulus-lists of randomly arranged words, one list having words with low, and the other with high, Thorndike-Lorge frequencies-of-usage counts. Clustering is defined as the occurrence of sequences of two or more words in the same category. Both lists comprised 60 nouns, including 15-each, in the following four categories: animals, names, professions and vegetables. The low frequency list was presented once to a group of 75 undergraduate student subjects and similarly the high frequency list to a different group also comprising 75 undergraduate student subjects. Immediately after the presentation of the lists of stimulus-words, the subjects were asked to list the words they were able to recall. The results of the study were

interpreted as supporting the following experimental hypotheses:

1. The total amount of clustering occurring in the recall of randomized verbal associates should be a positive function of the degree of their reinforcement.

2. The progressive changes in clustering from the beginning to the end of recall of a group of randomized verbal associates should be modified by reinforcement in the following ways:

    a. The initial level of clustering should be a positive function of reinforcement.

    b. There should be a positive relationship between reinforcement and the speed of attainment of maximum clustering.

    c. For all levels of reinforcement, clustering should tend to decrease as the supply of available associates approaches exhaustion.

This study is presented as a correlate of an earlier research undertaken by the writers in which reinforcement was controlled by varying the number of presentations of the stimulus-word list. It tests the same hypotheses and similarly depends for its theoretical interpretation on Hebb's account of the development of superordinate perceptions.

REFERENCES

1. Bousfield, W. A., & Cohen, B. H. The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates. Technical Report No. 1 under contract Nonr-631 (00) between the Office of Naval Research and the University of Connecticut.

2. Hebb, D. O. The Organization of Behavior. New York: Wiley, 1949. Pp. 335.

3. Thorndike, E. L., & Lorge, I. The Teacher's Word Book of 30,000 words. New York: Teachers College, Columbia University, 1944. Pp. 274.