DOCUMENT RESUME

ED 070 024 CG 007 710

AUTHOR Schmuller, Joseph A.; Farley, Frank H.
TITLE Toward a Multi-Dimensional Measure of Clustering in Free Recall.
INSTITUTION Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
REPORT NO WRDCCL-WP-95
BUREAU NO BR-5-0216
PUB DATE Dec 71
CONTRACT OEC-5-10-154
NOTE 16p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Cluster Analysis; Instrumentation; *Learning; *Measurement Techniques; *Recall (Psychological); *Retention

ABSTRACT

Various reported measures of clustering in free recall are reviewed under categories of algebraic versus probabilistic approaches. Shortcomings in these measures are outlined and a new multi-dimensional measure is advanced which overcomes many of the deficiencies noted. (Author)

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This working paper is from the Motivation and Individual Differences in Learning and Retention Project in Program 1, Conditions and Processes of Learning. General objectives of the Program are to generate knowledge about concept learning and cognitive skills, to synthesize existing knowledge and develop general taxonomies, models, or theories of cognitive learning, and to utilize the knowledge in the development of curriculum materials and procedures. Contributing to these Program objectives, this project has these objectives: to determine the developmental role of individual differences and motivation-attention in the learning and memory process and to ascertain at what age certain individual differences become important in learning and memory and at what age certain motivation-retention relationships emerge;

to develop a theory of individual differences and motivation in learning and memory; and to develop practical means, based on the knowledge generated by the research, as well as synthesized from other sources, to maximize the retention of verbal material.

ACKNOWLEDGMENTS

Appreciation is expressed to Katherine Feldman and Thomas J. Fischbach for their comments on various aspects of this analysis.

CONTENTS

ABSTRACT

Various reported measures of clustering in free recall are reviewed under categories of algebraic versus probabilistic approaches. Shortcomings in these measures are outlined and a new multi-dimensional measure is advanced which overcomes many of the deficiencies noted.

CLUSTERING MEASURES

In a free-recall experiment, S is presented with a list of items which he then is instructed to recall. If the original items can be classed into mutually exclusive categories, it has been found that S usually arranges these items by category when he recalls them. This is known as clustering.

Historically, indices of clustering in free-recall experiments have taken what is herein referred to as the "algebraic" approach. That is, some index or ratio is derived from certain characteristics of the data such as repetitions of items from a given category, runs of item types, lengths of runs, etc. The ratio is obtained by comparing these numbers to some ideal such as maximum possible runs, maximum pairs, etc. A review of this approach and the various types of indices it has produced up to 1969 is presented by Shuell (1969).

In a paper which appeared after Shuell's review, Dalrymple-Alford (1970) utilized the algebraic approach and derived the index C, which is:

$$C = \frac{\text{(Repetitions - Minimum Repetition)}}{\text{(Maximum Repetition - Minimum Repetition)}}$$

Dalrymple-Alford also gives equations for finding maximum and minimum repetitions.

More recently, Frankel and Cole (1971) have arrived at what will be referred to in the present paper as the "probabilistic" approach. Using statistics of sequential lists, they calculate the mean and variance of runs (strings of items from the same category) expected in a given

list of recalled items. Using these statistics and the observed number of runs, a z score may be obtained whose probability of occurrence may then be read from a standard normal table.

Alternatively, a chi-square approximation may be used which removes a good deal of the computation involved in obtaining z.

$$X^2_{(1)} = \frac{(O_R - M_R)^2}{M_R}$$

The associated probability may then be found in a X^2 table.

Whether algebraic or probabilistic, an index of clustering ought to be sensitive to differences in cluster length. An example will be given which shows that both the Dalrymple-Alford measure and the z score are fallible in this respect.

Given a presented list of 16 items, four of which fall into category a, four into category b, four into c, and four into d, S might possibly recall five items. Three of these possible strings are:

(1) a a b c d,

(2) a a a b c,

(3) a a a a b.

For each of these the Dalrymple-Alford C index is 1.00, implying perfect clustering. Interesting results are also obtained by the probabilistic method: for (1), $z = -1.22$; for (2), $z = -1.33$; and for (3), $z = -1.22$. Thus, the latter method does not discriminate between the least clustered case and the most clustered (most clustered being defined as items from a category occurring together most frequently), while the former makes no distinction at all.

The strong point of the probabilistic approach is that is yields a conditional probability; i.e., the probability that the observed amount of runs will occur given the items that have actually been recalled. The z score, however, does not take into account the structure of the original population of items. Thus, the obtained value of z for (3) does not reflect the fact that S has clustered all the items from a.

In developing a conditional probabilistic measure, then, the structure of the sampled-from list should be taken into account as well as the recalled list. One ideal measure would be

$$\Pr(\text{observed recall list}) = P(A|B)$$

where (A) = recalling a given string of items

(B) = recalling a given set of items

from a list whose composition has been defined by E. An example will make this clearer. Suppose S recalls a a a a b. We then wish to find the probability of this particular string given that four a's and one b have been recalled from a list containing four a's, four b's, four c's, and four d's. From the basic laws of conditional probability,

$$P(A|B) = \frac{P(A)\ \ P(B|A)}{P(B)}$$

In this case $P(B|A) = 1$, since if we are given that a certain string has occurred, we know that a list composed of the items in that string must have occurred. Therefore,

$$P(A|B) = \frac{P(A)}{P(B)}$$

By the hypergeometric distribution,

$$P(B) = \frac{\prod_{i=1}^{j} \binom{Ti}{ni}}{\binom{T}{N}}$$

$$\text{where } \sum_{i=1}^{j} Ti = T \text{ and } \sum_{i=1}^{j} ni = N$$

P(A) is represented by the joint probability of recalling n items, and attaining n - b runs; e.g., for (3) this would be p (number of items = 5 and number of runs = 2). If we assume that items recalled are normally distributed as well as runs, we can calculate these probabilities after calculating mean and variance of items recalled and mean and variance of runs. These probabilities can then be multiplied to obtain P(A), which when divided by P(B) yields $P(A|B)$. A string which has a low conditional probability may then be said to be "significantly" clustered.

Another way of viewing clustering is to see it as a deviation of observed runs from the maximum possible number of runs, given that n items have been recalled. Maximum runs are then simply equal to n, and we have

$$X^2_{(1)} = \frac{(\text{observed runs} - n)^2}{n}$$

A low probability would then indicate significant clustering.

Another way is to use a X^2 approximation in conjunction with the number of observed and expected repetitions in the recalled list. From Dalrymple-Alford (1970),

$$E(\text{Reps}) = \frac{n_1^2}{N} - 1$$

in which n_1 is the number of items of type $\underline{i}$ in the recalled list. Then

$$X^2_{(1)} = \frac{O(\text{Reps}) - E(\text{reps})^2}{E(\text{reps})}$$

The process of clustering would seem to be dependent on at least three subprocesses--concentration within categories, grouping, and recall. Measures of clustering thus far proposed yield a result which might be arrived at in a variety of ways and which does not reflect performance on these three processes. Since this is a problem intrinsic to any single-valued measure, an ideal clustering measure ought to result in more than one value which could then be collapsed into a single result if the investigator wished.

The measures to be proposed have zero as their ideal cases. This runs counter to Frankel and Cole's (1971) contention that a cluster measure should increase as clustering increases.

To measure <u>concentration within categories</u>, the variance of recalled proportions may be used. This is simply

$$\text{var} = \frac{\sum_{j}^{J} (n_j)^2 - J\bar{X}^2}{J}$$

in which n_j = number of items recalled from category j

J = number of categories

$\bar{X}$ = mean proportion recalled

The measure of concentration is thus

$$V = 1 - \frac{\text{var}}{\text{max var}}$$

The variance is divided by the maximum possible variance in order to equalize the scales on all three measures.

The measure of groupingness will simply be the mean number of terms between an item and the next item in the string from the same category. Thus, for a b c b b c a, there are five terms between the two a's, one term between the first two b's, no terms between the second two b's, and two terms between the two c's. Thus, our preliminary M (groupingness measure) = 2, since there are eight "between" terms and four pairs of similar terms. Note that the second b was counted twice (as the second member of the first pair and the first member of the second pair.) In order to keep our scales equivalent, we divide this by the maximum possible result, which can be shown to be J - 1 (where J = number of categories). M for this example, then, is .667.

Our measure of recall is

$$R = 1 - \frac{\text{number recalled}}{\text{number in list}}$$

Thus, this procedure yields three results which may be thought of as describing a point in three-space. Perfect clustering occurs at the origin of this space. If a single value is desired, it seems logical to use the distance from the obtained point to the origin -- which is, of course,

$$D = \sqrt{V^2 + M^2 + R^2}$$

A single result is thus obtained whose components are easily retrieved.

REFERENCES

Dalrymple-Alford, E. C. Measurement of clustering in free recall. Psychological Bulletin, 1970, 74, 32-34.

Frankel, F., & Cole, M. Measures of category clustering in free recall. Psychological Bulletin, 1971, 76, 39-48.

Shuell, T. J. Clustering and organization in free recall. Psychological Bulletin, 1969, 72, 353-374.

GPO 826-076