Psychological Reports, 1963, 13, 359-363.

AN EXPERIMENTAL HYPOTHESIS OF INTRALIST GENERALIZATION[1]

IRIS C. ROTBERG

Human Resources Research Office, The George Washington University

Summary.—Gibson (1940) has hypothesized that stimulus generalization during discrimination learning must increase before it can decrease. This hypothesis can be either supported or rejected, depending on the procedures and measures used in testing it. The present paper suggests a different approach to the measurement of the trend of generalization during discrimination learning. The proposed methodology compares similar and dissimilar confusion errors on the first learning trial and the rates of decrease of the exponential functions of the two error types on subsequent trials. The implications of the methodology for transfer and predifferentiation studies are discussed.

Gibson (1940) has suggested that stimulus generalization during discrimination learning must increase before it can decrease. Murdock (1958) takes a different view. In order to demonstrate that generalization begins at a maximum and steadily decreases, he has restructured the experimental procedure. His procedure eliminates intrusions and omissions by informing *Ss* in the instructions what the responses are and then requiring them to respond on the first trial as well as on succeeding trials.

A series of papers (Battig, 1959; Gibson, 1959; Runquist, 1959) answering Murdock have argued that Murdock thereby predetermined his result by equating errors with generalization. The papers have stressed that certain procedures are artifacts which determine whether an initial rise in generalization will occur. For example, if response learning is a factor in the early trials, *Ss* might not make generalized responses in these trials simply because they do not yet know the responses. Therefore, *Ss* under a free-response procedure would be less likely to make generalized responses during the early trials than would *Ss* under a multiple-choice procedure. The artifacts also include the measure of generalization used. The measure might be absolute number of generalization errors or it might be generalization errors relative to total number of correct responses for each trial.

It is possible, therefore, to either support or reject Gibson's hypothesis, depending on the procedures and measures used in testing it. We need a reformulation of the statement of the problem and a more meaningful measure of the trend of generalization during discrimination learning. The purpose of

[1]The research reported here was conducted by the author while she was employed by The George Washington University, Human Resources Research Office, operating under contract with the Department of the Army. Opinions and conclusions are those of the writer and do not necessarily represent views of the University or the Department of the Army.

this paper is to describe a method for comparing the trends of different levels of generalization. The method employs a categorization of learning errors into similar and dissimilar confusion errors. Similar errors are considered a measure of high generalization; and dissimilar errors, a measure of low generalization.

Method

The trend of generalization during discrimination learning can be effectively measured by a comparison of generalization curves for different degrees of similarity. For the comparison, a paired-associate list is used in which each stimulus is similar to certain stimuli on the list and dissimilar to others. Each stimulus thereby provides the opportunity for both errors indicating confusion between similar stimuli and errors indicating confusion between dissimilar stimuli. In this connection, Gibson (1942) classified errors as confusions between similar or dissimilar items. Her findings, which were confirmed elsewhere, showed that error responses tended to be those responses which would have been appropriate to a stimulus similar to the one presented.

Over a series of learning trials, therefore, we expect the total number of similar errors to be greater than the total number of dissimilar errors. The proposed method delineates the manner in which this total difference occurs by comparing the generalization curves for similar and dissimilar errors. Similar and dissimilar errors are compared on the first learning trial, and the rates of decrease of the exponential functions of the two error types are compared on subsequent trials. The total number of similar errors can be greater than the total number of dissimilar errors at the completion of the learning trials if one or both of these differences occur: (a) similar errors are greater than dissimilar errors on the first trial; (b) the rate of decrease of the exponential function is slower for similar than for dissimilar errors on subsequent trials.

Because the learning curve is typically exponential, exponential rather than linear functions are employed for the comparison of similar and dissimilar errors. A comparison of rates of decrease of linear functions would probably indicate only that the curve with the greater initial number of errors showed the greater absolute amounts of decrease between trials.

The proposed method avoids certain problems that prevent a meaningful experimental test of Gibson's hypothesis. First, Gibson's hypothesis does not specify the degree of similarity necessary before an initial rise in generalization occurs. Since some generalization can be expected during the learning of any set of stimuli, the hypothesis seems to predict an initial rise in every learning curve. In fact, the rising curve of generalization for the first three-tenths of learning that was indicated by Gibson's study (1942) was based on both similar and dissimilar errors. The present formulation compares generalization curves for different degrees of similarity, rather than predicting the shape of a single curve. If the findings indicate that similar errors decrease more slowly than

dissimilar errors, we can then determine where and in what manner the curves differ.

Second, the methodological artifacts previously described largely determine the outcome of experimental tests of Gibson's prediction. The occurrence of an initial rise in generalization depends largely on the number of intralist responses that the nature of the task permits *Ss* to make in the early trials. An initial rise also depends on whether the error measure for each trial is the absolute number of confusion errors or confusion errors relative to total number of correct responses for that trial. These artifacts are irrelevant to the proposed comparison between generalization curves for different degrees of similarity.

Specific Applications

An inspection of data from Gibson's study (1942) provides indirect evidence that similar errors decrease more slowly than do dissimilar errors. Her results indicate that more similar than dissimilar errors were made in both the first and second halves of learning and that this difference was relatively greater in the second half.

Another indirect test is provided by a further experiment performed as a follow-up to a study on similarity grouping (Rotberg & Woolman, 1963). In this follow-up experiment, a nine-word list was used, consisting of nonsense syllables paired with English words. The nonsense syllables formed three similarity categories of three terms each. These terms were used in either the stimulus or the response position. Either similar or dissimilar items were grouped for learning. There were four experimental groups: similar stimuli grouped, dissimilar stimuli grouped, similar responses grouped, dissimilar responses grouped. The associated English words were neutral with respect to the similar-dissimilar factor.

There were 10 trials. Each trial included (a) a learning period during which either similar or dissimilar items were grouped, with correct answers following *S*'s responses; (b) a test period in which stimuli were presented randomly and correct answers were not given; (c) a test period in which responses were presented randomly and *Ss* were required to give the matching stimuli. *Ss* did not respond during the first learning trial.

The following comparisons were made. (1) Trial 1 comparisons between number of similar and number of dissimilar stimulus errors. Similar errors were weighted by three for the comparisons because their chance probability of occurrence was one-third the probability of the dissimilar errors. (2) Trend comparisons between the rates of decrease of the exponential functions of the two error types.

The comparisons were made separately for treatments within each of the experimental periods. Differences were evaluated by means of t tests. Trial 1 comparisons for the stimulus groups indicated significantly more similar than dissimilar errors for five of the six comparisons ($P < .05$). Trend comparisons indicated that similar errors decreased significantly more slowly for two of the comparisons ($P < .05$). Similar errors decreased faster in one comparison ($P < .10$). The other three comparisons were nonsignificant.

The present method can be applied to response similarity as well as to stimulus similarity. Trial 1 comparisons for the response groups indicated significantly more similar than dissimilar errors for two of the comparisons ($P < .05$) and significantly more dissimilar errors ($P < .05$) in one comparison. Trend comparisons indicated that similar errors decreased significantly more slowly for two comparisons ($P < .05$). The difference approached significance in a third comparison ($P < .10$).

The Trial 1 comparisons clearly indicate more similar than dissimilar stimulus errors but are unclear for response errors. The trend comparisons do not support a conclusion that similar errors decrease more slowly than do dissimilar errors, although this might be true under certain experimental conditions. The comparisons suggest that, at least for some conditions, the rates of decrease do not differ. It appears that for certain groups the difference between similar and dissimilar errors occurs on Trial 1 and that this difference is passed on from trial to trial. However, a more direct test is necessary before firm conclusions can be reached. For the most direct test, the similar and dissimilar items within the learning list would be randomized within trials. The Trial 1 and trend comparisons would be based on the number of similar and dissimilar learning errors made on each trial.

Predictions of Transfer and Predifferentiation

Applications of Gibson's hypothesis to transfer and predifferentiation have resulted in predictions that are subject to the same problems of interpretation as the original hypothesis. Gibson (1940) predicted that when stimuli are similar and responses different, the amount of interference between two lists would follow the curve of the first-list generalization errors. According to Gibson's hypothesis, maximum negative transfer would occur up to the peak of generalization for a second list introduced after the learning of the first. After the peak of generalization, there would be less negative transfer as learning progressed.

Gagne and Foster (1949) have made a similar prediction with respect to predifferentiation. They hypothesize that, "If the amount of practice on an initial task involving predifferentiation is increased, the learning of the final task will at first be slowed because of an increase of internal generalization. Following this, the learning of the second task will be made more rapid as the amount of generalization produced by the initial practice passes through a maximum and decreases" (Gagne & Foster, 1949, p. 48). Murdock (1958), on the other hand, indicates that generalization is greatest at the start of learning and that positive transfer resulting from stimulus predifferentiation may be expected after a small number of trials.

The present method of analysis avoids this type of dichotomy, which depends largely on procedural factors. Instead, the method can be applied to compare transfer for different degrees of similarity. For the comparison, the initial and final tasks used include stimuli forming two or more similarity categories. Trial 1 and trend comparisons can be made between: (1) similar and dissimilar errors within the initial list, (2) similar and dissimilar errors within the final list, (3) similar and dissimilar errors between the two lists. This type of analysis can be made for different levels of learning of the initial and final lists. The relative importance of each type of error at a series of proficiency levels can thereby be determined.

References

BATTIG, W. F. Comment on "Intralist generalization in paired-associate learning." *Psychol. Rev.*, 1959, 66, 338-339.

GAGNE, R. M., & FOSTER, H. J. Transfer of training from practice on components in a motor skill. *J. exp. Psychol.*, 1949, 39, 47-68.

GIBSON, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, 47, 196-229.

GIBSON, E. J. Intra-list generalization as a factor in verbal learning. *J. exp. Psychol.*, 1942, 30, 185-200.

GIBSON, E. J. A re-examination of generalization. *Psychol. Rev.*, 1959, 66, 340-342.

MURDOCK, B. B., JR. Intralist generalization in paired-associate learning. *Psychol. Rev.*, 1958, 65, 306-314.

ROTBERG, I. C., & WOOLMAN, M. Verbal paired-associate learning as a function of grouping similar stimuli or responses. *J. exp. Psychol.*, 1963, 65, 47-51.

RUNQUIST, W. N. Remarks on "Intralist generalization in paired-associate learning." *Psychol. Rev.*, 1959, 66, 343-344.

Accepted July 12, 1963.