A COMPARISON OF GROUP AND INDIVIDUAL RESPONSE PROBABILITIES IN CONTINUOUS VERBAL ASSOCIATION

by John J. Fleming

Thesis presented to the School of Psychology and Education of the University of Ottawa as partial fulfillment of the requirements for the degree of Doctor of Philosophy

Erie, Pennsylvania, 1964

ACKNOWLEDGMENT

This thesis was prepared under the direction of Professor Lawrence T. Dayhaw, Ph.D., of the School of Psychology and Education, University of Ottawa.

CURRICULUM STUDIORUM

John J. Fleming was born September 13, 1918 in Erie, Pennsylvania. He received the Bachelor of Arts degree in Psychology from Gannon College, Erie, Pennsylvania, in 1946. He received the Master of Arts degree in Psychology from Fordham University, New York, New York, in 1950. The title of his thesis was An Evaluation of a New Hapto-Kinesthetic Gestalt Test.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

A basic interpretative problem revolves around the extent to which measurements made on a group of subjects are directly or univocally referable to each individual in the group. Very frequently statistical quantities such as the mean or variance are treated solely as group measurements and the individual's performance is related to these quantities simply according to some relative or deviational scale.

It is conceivable, for example, that each individual possesses within himself a "moment" or point of mean intellectual functioning, but we would not be prepared to identify such a hypothetical quantity with the mean intellectual performance of his age group. Yet it is essentially this kind of transposition in meaning that is often made in the case of verbal associative-strength. That is to say, that the frequency with which a response occurs in the group is commonly taken as a definition (as well as a measurement) of associative-strength as applied to the individual subject.

In the Kent-Rosanoff list of words, for example, the response, "chair", to the stimulus, "table", occurs about 267 times out of 1000 in the original standardization data. This can be expressed as a probability (0.267) and as such is interpreted as a measure of associative-strength in the individual subject. The following question immediately arises: If "chair" is the most probable, i.e. the strongest response associatively speaking, then why is it not given every time? If a subject responds to "table" with "room" the probability is 0.001. Yet one must conclude, it seems, that at the moment of response this word was the strongest or at least the most insistent reaction to the stimulus.

There are different ways of handling this apparent dilemma. One might simply prefer to suspend judgment on the significance of such probabilities in the individual and just deal operationally with group frequencies. Such an approach is experimentally admissable but one is left without any real semantic link between group and individual performances. That is, the concept of associative-strength might have different meanings when asserted concerning the group as compared to the individual. To state that frequency of response has a certain quantitative value is a far different kind of assertion than to say that it is by definition a measure of "associative-strength" or "habit-strength" in the individual. These latter concepts have definite connotations in the psychology of association and learning.

Another way of approaching the apparent dilemma presented above would be to adopt an analysis similar to that of some learning theorists. This would assume that habit-strength or associative-strength remains fairly constant but that its manifestation in behavior fluctuates in a mathematical manner as a function of drive level, incentive, and so on. Some psychologists have come to feel that such a list of intervening variables introduces a needless multiplication of assumptions and postulates. At any rate in the case of pre-formed verbal associations these various intervening variables cannot be effectively measured.

The present thesis is an attempt to cast some light on this group-individual problem in verbal association. In order to do so a certain presupposition has entered into the experimental design. It is essentially this: if associative-strength, defined as response probability in the group, is to be applied univocally to the individual's verbal associations, then it must be demonstrated that such statistical properties do in fact apply to an existing population of responses within each individual. This means in turn that in each individual subject a population of verbal responses should be generated which is equivalent in range to the group population of responses. As a result, continuous-free and continuous-controlled association will be employed in this study because they are the only types which will readily permit the generation of such individual response populations under equivalent conditions.

This study will be concerned basically, then, with an attempt to compare the mathematical probabilities of a group population of verbal responses with those from equivalent individual response populations. Interpretation of such data will be made statistically and within the additional framework of unrestricted versus restrictive categorization of response, as well as on the basis of formal as compared to semantic probability.

CHAPTER I

EVOLUTION OF THE THESIS PROBLEM

A. Some Historical Notes on Association.

A history of the concept of association can be traced back to Aristotle who first formulated the classical laws of similarity, contrast, and contiguity. Not much emphasis however was given to association as a dynamic force in mental life until the 17th and 18th centuries. During this period renewed interest was created by Hobbes, Locke, and other British Empiricists[1]. In emulation of physics and chemistry they sought for principles of mind which would be both analytic and explanatory.

Most of the British Empiricists beginning with Hobbes attempted to reduce all associative principles to that of contiguity. If such a reduction had been successful, then mental philosophy would have been in possession of a unifying principle analogous to Newton's law of attraction in physical bodies. At a later time Herbart also conceived of a force between associated ideas which could be either attractive or repulsive. This was analogous to electrical and magnetic forces in physics (which unlike gravitation are negative as well as positive). Herbart is to be noted also for his

---

1 E. A. Burtt, Editor, The English Philosophers from Bacon to Mill, New York, The Modern Library, 1939, xxii-1041 p.

pioneering attempt to treat mental association in mathematical terms. His equations were rigorous enough from a quantitative point of view but unfortunately could not be put into a form that was experimentally testable. As the logician would say, the equations possessed formal truth (validity) but the material truth or falsity remained indeterminate.

The attempt on the part of the British Empiricists to derive analytic and explanatory principles of mind from the concept of association alone was found to be insufficient to deal with the sheer complexity of man's mental life. However it is of interest to note that the concept of association in one form or another has continued to shape psychological theory up to the present time.

Around the turn of the century Freud and the psychoanalysts had already employed free association not only interpretatively but also as an empirical tool with which to evoke non-conscious mental content. During this same period Hollingworth[2] wrote a general psychology centered around Sir William Hamilton's principle of redintegration. He defined the mental as redintegrative whereas for him the physical was non-redintegrative. Such a concept is far more comprehensive than the contiguity principle of Hobbes or the conditioning principles of Pavlov which also appeared around this time. As a matter of fact the laws of association and

2 H. L. Hollingworth, Psychology Its Facts and Principles, New York, D. Appleton and Company, 1928, 534 p.

conditioning can be construed as operating within the overall context of redintegration. To continue with previous analogies, Hollingworth's notion of redintegration might be compared figuratively to a field concept in physics.

Modern learning theory beginning with the early work of Ebbinghaus on memory represents an experimental study of association conceived of within the framework of initial acquisition. As McGeoch expresses it, "the psychology of learning studies the characteristics of such associations among psychological events and the conditions of which their acquisition and retention are a function"[3]. One might add that recall is one index of learning as well as of retention, and, association can in turn be considered as a form of recall.

The more recently developed field of psycholinguistics is to a considerable extent a study of the associational properties of language. However the data which are analyzed in psycholinguistics as compared to that of associative learning experiments are quite different. In psycholinguistic study language as a cultural phenomenon is already given. It possesses a definable structure of word combinations, pre-formed in experience, and it becomes the task of psycholinguistics to study the properties of this structure so as to relate them to determinable psychological laws. This

---

3 J. A. McGeoch, The Psychology of Human Learning, New York, Longmans, Green and Co., 1948, p. 26.

goal has not as yet been realized to any significant extent. Some psychologists such as Osgood[4] conceive of this integration as occurring between linguistic structure and current learning theory. It is the opinion of this writer however that in addition to learning theory psychologists must also seek for other, perhaps as yet unformulated, dimensions of language as a reference point for such integration. Because of its theoretical importance this statement will be given further elaboration in the following section.

B. General Background of the Present Problem.

Science might be described as the systematic formulation of human experience. Initially human experience implies conscious experience. It is one of the tasks of any science to begin with what is phenomenally given or observable in conscious experience, to classify, to establish relations, and then to make inferences about that which is not immediately given. We are thus confronted with a manifold of conscious events consisting of a variegated procession of sense perceptions, images, feelings, and thoughts. The physical, biological, and social sciences arbitrarily parcel out different sectors of this conscious manifold for study.

---

4 C. E. Osgood, Method and Theory in Experimental Psychology, New York, Oxford University Press, 1953, 800 p.

Both physics and biology deal with the sensory manifold as a beginning point for the abstract inferences which will constitute their scientific concepts and principles. As Einstein[5] has stressed, even though the abstract inferences of physics are mental constructs, they can be verified by returning to the reference point of the sensory manifold through the avenue of predictive inference. In other words both physics and biology find as accessible an objective reference point outside the mind which is sufficiently differentiated to provide independent verification for hypothetical inferences.

By comparison, however, psychology is faced with a methodological problem which to this day has not been satisfactorily resolved. Most definitions refer directly or indirectly to psychology as a study of inner experience and outer behavior. Emphasis has quite understandably been given to behavior because this can be treated as an objectively observable series of events which are relatively independent of the mind and against which the psychologist can check his inferences.

Some however point out that behavior is not a coherent system for purposes of determining cause and effect relationships because much of the internal process intervening

---

5 A. Einstein, Out of My Later Years, New York, The Wisdom Library, A Division of Philosophical Library, 1950, 247 p.

between stimulus and reaction is unconscious. Thus the suggestion is made that we turn to brain physiology for our objective reference point analogous to the sensory manifold of the physicist and biologist.

Now when we consider psychology as the study of inner experience, or experiencing self, or "mind", or mental processes, we run into a more difficult problem. That is to say, that we are not in possession of an objective reference point independent of the mind which is comparable to the relationship that neural events bear to behavior. Some inner conscious events can be correlated with neural physiology but such relationships are extremely meagre and manifestly incomplete.

Freud expressed the problem quite clearly as follows:

> Every science is based upon observations and experiences arrived at through the medium of our psychical apparatus. But since our science has as its subject that apparatus itself, the analogy ends here. We make our observations through the medium of the same perceptual apparatus, precisely by the help of the breaks in the series of (conscious) mental events, since we fill in the omissions by plausible inferences and translate them into conscious material.[6]

Without an adequate reference system independent of mental process itself we may turn to a purely descriptive phenomenology of conscious events, or, as in the case of many systems and theories of personality we can begin with a series

---

6 S. Freud, An Outline of Psychoanalysis, New York, W. W. Norton and Company Inc., 1949, p. 36.

of inductively derived postulates against which we attempt to validate subsequent inferences. But even though we attain some measure of formal consistency by such a procedure, we have as yet achieved no real formulation of fundamental psychic laws.

Hence psychology interpreted as an experimental study of the inner nature of experience itself is for the most part undeveloped. The criticism that psychology has not been successful in deriving an experimental science of mind is to this extent probably justified. There is of course no shortage of empirical theories and systems which attempt to explain "mind", "experience", or the "experiencing self" but these are generally limited by adopting a reference point embedded in inductively inferred postulates which Madsen terms axiomatized deductive systems.[7]

Is there, then, an objective, coherent reference system which is independent of the mind and in relation to which the psychologist can test the material truth of his hypothetical inferences, not about behavior, but rather about the process of experiencing itself? It is possible that language may at least in part fulfill such a requirement. In support of this it can be pointed out that language is uniquely human. It evidently reflects all facets of human

---

7 K. B. Madsen, Theories of Motivation, Cleveland, Howard Allen, 1961, 353 p.

experience, otherwise we could not even employ it to discuss the problem of experience. But more significantly language at least from a syntactical point of view is an objective system of signs and symbols with its own coherence and laws which are independent of the mind. Now if inner experience is reflected sufficiently in language, then perhaps we might be able to deduce and verify some of the underlying laws of experience and "mind" by a deeper analysis of language.

However such an analysis would depend upon concepts and dimensions which as yet have not been fully formulated with respect to language. Linguistic analysis up to the present has dealt with the structure of language, semantics, and general semantics. Recently the anthropologist and psychologist have also shown interest insofar as language determines one's very conception of reality as well as his thought processes. And what has come to be known as psycholinguistics deals with language as a communicated system of symbols with statistical properties. All of these trends are converging more and more. Perhaps out of such a convergence will come new concepts and dimensions rich enough to cast light on the laws of inner experience itself.

When the physicist "returns" to the sensory manifold to test out the hypothetical inferences which he has derived, he employs the dimensions of space and time and univocal concepts such as mass, momentum, etc. as a means of analysis and description. Can we derive dimensions and univocal concepts which can be applied to language conceived of as a reflection of inner experience? Hopefully this may come to be.

Let us return now to the point in our discussion which initiated the rather lengthy consideration of science, psychology, and language. One can agree with Osgood that psycholinguistics should be integrated with learning theory, and, in addition, with any other area of psychology that will cast further light on the nature of experience and behavior.

One of the distinct appeals that the study of language as a communicated product has for the psychologist is that it is relatively easy to quantify and deal with statistically. Consequently, even though the various known properties of language have not as yet been effectively integrated with psychological theory, the ease of quantification holds promise that fruitful generalizations will eventually emerge.

The present study deals with word association which has been found quite useful in the analysis of linguistic products and the quantification of meaning. Unlike Herbart's unsuccessful rational equations, modern researchers have found many useful equations of a stochastic or statistical type that can be applied to linguistic and associational sequence. If word association reflects mathematical properties of a uniform sort, then it is not unreasonable to suppose that a fundamental lawfulness in linguistic behavior is being probed. It is at this juncture of course that the psychologist becomes interested in what light such lawfulness might shed on mental and behavioral processes.

Now the statistical properties of verbal association are derived basically from group data. Association or associative-strength is usually expressed as the probability that a response will occur in a given stimulus situation. For example in the original Kent-Rosanoff list of words 267 persons out of 1000 responded to "table" with "chair". Thus we can express the probability of such a response in a similar group as Pr = 0.267. Repetition of the list with similar populations has confirmed the predictive accuracy of such probabilities.

However, if mathematical properties are to be meaningful in interpreting associative sequences within the individual subject, then it should be demonstrated that group mathematical properties do in fact apply to corresponding populations of responses within the individual subject. Although the experimental evidence tends in this direction, it is not unequivocally established. For unless we can make univocal assertions from the group to the individual, our conclusions will be arbitrary. Now it is exactly this kind of assertion that frequently is made concerning verbal associative-strength. Consider this formulation from Osgood on "associational structure":

> (...) the strength of associations has been inferred from group data (...). Is there any basis for expecting the same laws to apply to the availability of associations within the single individual? Thumb and Marbe (1901) obtained a logarithmic relation between individual reaction times for free associations and the frequency of occurrence of these responses in a population. This has come to be known as Marbe's Law and has been substantiated by a number of other investigators (...). But if this be the case, shouldn't the most frequent response be the only response to a given stimulus? Shouldn't the response having the strongest association with the stimulus always be made by every individual? There are several reasons for variability: the habit strengths of the potential responses show continuous oscillation (cf. Hull, 1943); the actual stimulus is a compound affair varying with the momentary context; and there remain real individual differences in the nature of mediation processes that have been established for the same signs.[8]

Much of what Osgood states here may be materially true. Future research will probably be the final arbiter. However his line of reasoning is quite hypothetical and as it stands is subject to considerable reservation. Because it may help to sharpen the focus of the present study, let us briefly consider the principal notions from this statement by Osgood, particularly because it represents a good deal of current thinking about verbal associative-strength.

If we accept group frequency of response as a measure of associative-strength, this becomes in effect a quantitative definition. As it stands this definition is neither true nor false. The question which arises is whether the definition applies to verbal associations within the individual subject.

8 C. E. Osgood, Method and Theory in Experimental Psychology, New York, Oxford University Press, 1953, p. 722.

Now speed or latency of response as expressed in Marbe's Law is inversely related to group frequency of response. That is, the more frequent responses are in general associated with shorter reaction times. However these two determinations are not independent of one another. Response latency is generally correlated with frequency of response with the result that we could in principle select either one as an independent variable. So we are really thrown back on group frequency both as a definition and as a measure of associative-strength.

The next question which Osgood considered is in effect: why does not "chair" occur in response to "table" every time in each subject if it is by definition the strongest as well as the most probable response? He then lists some intervening variables which may insert themselves somewhere between stimulus and response including Hull's concept of reaction potential ($_sE_r$). This is not the place to attempt an evaluation of Hull's elaborate system of postulates and intervening variables. For our purpose here it should be sufficient to point out that Hull's concepts are perhaps meaningful in experimental situations where associative learning is being formed for the first time. In dealing with pre-formed word associations such a system of constructs is not directly applicable and thus remains hypothetical. A similar kind of reservation applies to the other intervening variables which Osgood suggests.

So the question remains: does frequency of verbal response (or associative-strength) in the group apply univocally to the available population of responses within each individual? The present thesis represents an attempt to study this question under the conditions to be specified in the following section.

C. Concepts, Definitions, and Frame of Reference.

Four basic categories of verbal association are clearly definable: (a) discrete-free, wherein a single stimulus word is presented and the subject responds with the first word that comes into his mind (the Kent-Rosanoff list is an example of this category); (b) discrete-controlled, wherein a single stimulus word is administered and the subject is instructed to respond with a word in a particular class relation (for example, the subject might be asked to respond with "the opposite of" or "a part of", and so forth); (c) continuous-free, wherein the subject is instructed to continue responding with any words at all that may come into his consciousness (this is somewhat the same technique as that which the psychoanalyst employs in what he terms "free association"); and (d) continuous-controlled, wherein the subject is asked to continue giving as many words as he can within a given class (for example, the subject may be instructed to name as many cities, or carpenter's tools, or birds as he can within a given time period).

As indicated earlier, the general purpose of the present study is to determine whether group-frequency characteristics of association are applicable to the range of potential responses within the individual subject. In theory, to examine such a question, it would be sufficient to generate a population of responses in each subject and then determine whether the statistical properties of this individual population are similar to those of the group.

Now in order to generate in the individual subject populations of responses which potentially would be comparable to the range of group responses, this study will deal with continuous-free and continuous-controlled categories of word association. The reason for this is that discrete-free and discrete-controlled association do not lend themselves to the generation of sufficient populations of responses in the individual subject. Consider the Kent-Rosanoff list of discrete-free associations. In order to produce a population of responses in a single subject which would be equivalent to the original group data, one would have to administer the list to a single subject one thousand times. And in the process extraneous learning factors, reinforcements, serial effects, and so on, might enter to modify the task in a critical manner.

In addition, the present study will deal interpretatively with formal probability of response as well as semantic probability of response. These terms are definable as follows: formal probability (or frequency) is the number of verbal responses which occur in a given period of time irrespective of the content of the responses; and semantic probability on the other hand refers to the probability (or frequency) of obtaining a particular word as a response. In the Kent-Rosanoff list of responses, for example, the stimulus "table" elicited "round" ten times out of a thousand. The semantic probability of obtaining "round" is therefore 0.01. The formal probability is also 0.01 but it applies to any and all responses which occur exactly ten times out of a thousand.

Formal probability evidently possesses in one sense a greater generalizing value and at the present stage of research it is rather intriguing to contemplate the possibility that such generalized mathematical quantities may begin to describe formal properties of mind insofar as they are reflected in verbal-linguistic behavior.

Finally the term "response" is to be defined as any emitted word which follows upon the initial instructions to associate freely, or, to associate continuously within a given class of terms. This definition thus differs from the notion of response in discrete-free association, for example, where the subject replies to a specific verbal stimulus with a single word.

D. Previous Research and Thesis Hypothesis.

Some of the earliest and most significant studies concerning the mathematical properties of language were carried out by Zipf[9]. He analyzed fairly large samples of Peiping Chinese, Plautine Latin, and American newspaper English. Among other things he found a common logarithmic relation between frequency of occurrence and rank order of frequency in these various samples of written language. Thus the lower the rank order the greater the number of different words found at that rank. When Zipf plotted this data on double log coordinates he obtained a clearly linear relationship. He expressed these results in the following form: $a \cdot b^{2.0} = K$ where b is the frequency with which a word occurs in the sample, a is the number of words of the same frequency, and K is a constant.

Subsequently an analysis was made of Old High German which yielded a value of $a \cdot b^{1.98} = K$. An additional study of the Gothic bible (as translated by Ulfilas in 310 A.D.) resulted in a value of $a \cdot b^{2.025} = K$. All of these numerical values are remarkably close to 2.0.

It is of interest that these quantitative properties of written or oral speech emerge from fairly large samples of word populations and assume the characteristic of what might be termed "meta-language" i.e. a property of language distinct from lexical and syntactical form.

---

9 G. K. Zipf, The Psycho-Biology of Language, Boston, Houghton Mifflin, 1935, 718 p.

The formal property in language was even more strikingly demonstrated by Skinner[10] in his experiment with the verbal summator, an apparatus which emitted random, meaningless, phonetic sounds. At a given distance the sounds resembled overheard conversation. The subjects then related what they thought was being spoken. When Skinner plotted his data against Zipf's the fit was just about as good. At least it was clearly of the same mathematical form.

From these results Skinner concluded that Zipf's mathematical law did not refer to a property of behavior, such as the conscious selection of an appropriate word, but rather to the condition of associative-strength in which latent speech exists. Presumably he meant that, given a gradient of strength in association, words will be emitted in the long run according to the obtained logarithmic form and independently of the external situation to which the person is responding. From a quantitative point of view this formal property is indeed a remarkable finding.

Skinner's theoretical orientation would almost certainly prevent him from allowing for mentalistic variables such as selectivity or set in the above findings. However it should be noted that the logarithmic function which was obtained might conceivably reflect just such a formal property of

10 B. F. Skinner, "The Verbal Summator and a Method for the Study of Latent Speech", Journal of Psychology, Vol. 2, 1936, p. 71-107.

selectivity in mental operation. t least this hypothesis has not been fully explored as yet.

There are many other measures of linguistic properties which lead to formal mathematical expression. However we are unable to discuss all of them here.

In a study relating to discrete-free word association Skinner[11] ranked responses on the Kent-Rosanoff list according to frequency of occurrence. He then combined ranks for all items, i.e. all first ranking responses were combined, all second ranking combined, and so on. The average frequency of occurrence was found to be a logarithmic function of rank somewhat similar to Zipf's law. One of the values Skinner reported was the following: $A \cdot B^{1.29} = K$ where A is the frequency with which a given association occurs, B is the rank, and K is a constant.

As noted earlier, these quantitative findings, in language and association, point to systematic mathematical properties. The frequently recurring logarithmic relationship seems to make a definite assertion about associative sequences. That is to say, that the rate of association at any point is a function of the number of remaining associations; or, the number of different responses increases logarithmically as the associative-strength decreases.

---

11 B. F. Skinner, "The Distribution of Associated Words", Psychological Record, Vol. 1, 1937, p. 71-76.

Bousefield and Sedgewick[12] generated a population of responses in individual subjects with the method of continuous-controlled association and then analyzed the properties of such responses. They employed various classes of terms such as birds, cities, animals, names of fellow students, and so on. The group means of a cumulative number of responses were plotted as a function of elapsed writing time. Their data was fitted by the following log type curve: $n = c(1 - e^{-mt})$ where n is the cumulative number of responses, t is the elapsed writing time, e the logarithmic base and c and m are constants. In another study, Bousefield and Barclay[13] found that responses given more commonly by the group were also those which occurred earlier in the individual lists. They concluded that ordinality of position is then another criterion of associative-strength. This latter result can be interpreted as being consistent with Zipf's data and Skinner's data.

The two studies just referred to did generate a fairly extensive population of responses in the individual subjects prior to statistical analysis. However it is felt that a crucial experiment in this respect has not been performed as

---

12 W. A. Bousefield and C. H. W. Sedgewick, "An Analysis of Sequences of Restricted Associative Responses", Journal of General Psychology, Vol. 30, 1944, p. 149-165.

13 W. A. Bousefield and W. D. Barclay, "The Relationship between Order and Frequency of Occurrence of Restricted Associative Responses", Journal of Experimental Psychology, Vol. 40, 1950, p. 643-647.

yet. By a crucial experiment is meant one in which a population of responses is generated in the individual subject with the requirement that this population of responses is equal at least numerically to the group population of responses. If the definition of associative-strength derived from group responses is to apply univocally to the individual, then it seems that at least numerically equal populations of responses must be assumed before a judgment can be made.

Now we cannot simply assume that, given the opportunity, an individual subject will give the same word content as found in the group even with unlimited time. But of course this is one of the questions that requires clarification.

Referring back to the experiments of Zipf and Skinner it is of interest to observe that they obtained similar mathematical curves with language data that varied considerably in content. So it is conceivable that formal mathematical probabilities will obtain in continuous-free and continuous-controlled verbal association when the semantic probabilities do not. Again this remains to be determined. But if such a difference did obtain, it would at least lend some precision to the interpretation of experimental data even though it would greatly limit the current definition of verbal associative-strength.

In addition the present study will compare group-individual verbal frequencies with unrestricted and restricted task sets. The operation of these variables may also add some clarification to the general problem of interpretation.

The following general hypothesis is to be tested in the present thesis: there are no significant differences in statistical frequencies between a population of responses obtained in a group of subjects as compared to individual populations of responses obtained under similar conditions. The method to be employed in an attempt to test this general hypothesis will be presented in Chapter II, "Experimental Design".

CHAPTER II

EXPERIMENTAL DESIGN

A. General Procedure and Method.

As indicated in Chapter I the design of experiment to be employed in this study is based on the following preliminary considerations: we can make the assumptions that the frequency with which a verbal response occurs in a group of subjects can be considered as (a) a numerical measure of habit-strength or associative-strength, and, (b) this measure of associative-strength applies to each individual subject in the group. Now if both these assumptions are correct, then by generating a sizeable population of verbal responses in each of the subjects comprising the group, similar frequencies or proportions of response should be found in the individual subjects as are obtained in the group as a whole. If such similarity is obtained, then associative-strength can be not only defined in terms of probability but it can also be univocally applied to group as well as any individual subject.

In order to generate a sizeable population of responses in each individual subject continuous-free and continuous-controlled association will be employed. Discrete-free and discrete-controlled association, as pointed out earlier, do not lend themselves readily to the generation of such individual response populations.

The comparison of group and individual response populations will be based upon two criteria, one of which is a logical-empirical criterion, the other, a statistical criterion. The logical-empirical requirement is that responses common to all subjects must be given. Otherwise the generalizing value of such a measurement or definition is so restricted that its usefulness is to be questioned. So, given common responses, the second or statistical criterion is to be invoked in order to determine whether or not the obtained group-common responses are significantly different in occurrence from individual-common responses

In the continuous-free (c-f) associative task each of ten subjects is asked on the first trial to associate freely until he records in writing one hundred responses. After each successive two-minute time period the subject places a slant mark (/) to indicate such intervals of time. This task is repeated for nine successive trials until a total of one thousand responses is recorded for each of the ten subjects, or, ten thousand responses in all.

In the continuous-controlled (c-c) associative task each of the same ten subjects is asked on the first trial to write down the names of as many four legged animals as he can. Following Bousefield and Sedgewick's method each subject continues to record these verbal responses for a period of eighteen minutes. Preliminary trials confirm the fact that most subjects apparently run out of names within such a time period. This task is also repeated for nine successive trials. The eighteen minute time period in each trial is subdivided into nine consecutive two-minute periods and recorded as such by the subject with a slant mark(/).

The formal design of the experiment can be envisaged then as a ten by ten array of data (i.e. the responses of ten subjects on ten trials), and, the presentation of such an array on two different associative tasks (i.e. continuous-free and continuous-controlled association). And finally such arrays or tables of response frequencies will contain as entries only common responses. That is to say, verbal responses which are given by all ten subjects somewhere within the ten trials.

The array or tables so described can be evaluated statistically according to an analysis of variance, two-way classification, conceived of as a mixed model wherein the subjects are random and the successive trials are fixed.

In Chapter I a distinction was made between semantic and formal probability (or frequency). Semantic probability refers to a particular word and the frequency with which it occurs. We can now further specify its meaning by defining it also as a common response.

Formal probability is the frequency with which words, any words, common or not, are given in time. Interpretatively, in the treatment of results, the interest in formal probability centers on continuous-controlled association. The reasons for this are as follows: (1) in continuous-controlled association all subjects will respond uniformly for nine time intervals of two-minute duration each whereas in continuous-free association one cannot expect a uniform number of time intervals; and, (2) the formal rate of continuous-controlled association will provide direct comparison with other research of the same type—particularly that of Bousefield and Sedgewick at the University of Maryland. Comparability of results on a formal basis can be determined, for example, by an attempt to fit the data of this thesis to the function, $n = c\ (1 - e^{-mt})$, obtained by them in several comparable continuous-controlled tasks.

B. Specific Hypotheses to be Tested.

With respect to continuous-free association:

I (a) There are continuous-free responses given which are common both to individuals and to trials.

I (b) There is no significant difference in the relative frequency of continuous-free common responses in the overall group as compared to those of each subject throughout ten trials.

I (c) There is no significant difference in the relative frequency with which continuous-free common responses are given by all subjects on each of ten trials.

With respect to continuous-controlled association:

II (a) There are continuous-controlled responses given which are common both to individuals and to trials.

II (b) There is no significant difference in the relative frequency of continuous-controlled common responses in the overall group as compared to those of each subject throughout ten trials.

II (c) There is no significant difference in the relative frequency with which continuous-controlled common responses are given by all subjects on each of ten trials.

C. Subjects and Instructions.

The subjects for this experiment are ten college seniors with a mean chronological age of 23.0 years and a range of 22.0 to 24.0 years. All ten subjects have achieved scholastic grade averages of B (85.0) or above. In terms of academic concentration and school affiliation they are to be subdivided as follows: five concentrators in Psychology from Gannon College, Erie, Pennsylvania; one concentrator in English from Gannon College; one concentrator in Electrical Engineering from Gannon College; one concentrator in

Political Science from Georgetown University, Washington, D. C.; and two major seminarians from St. Bonaventure University, Allegheny, New York.

The instructions for continuous-free association are as follows:

> When I say "start" I want you to commence writing down any words whatever that may come into your mind. Let your mind wander completely and record each word that comes no matter how foolish or embarrassing or pointless it may seem to you at the time. Do not think of school subjects or any such categories. Just let your mind wander. Every so often I will say "mark" and then I want you to draw a slant mark like this (demonstrates) immediately after the word you are writing at that time. Continue writing down words at random until all of the squares on the sheet are filled. (Each subject is supplied with an 8½ by 11" sheet of paper on which are marked out one hundred separate squares.) Are there any questions?

The instructions for continuous-controlled association are as follows:

> This time when I say "start" I want you to write down the names of as many four-legged animals as you can, that is, animals that walk on four legs. Every so often I will call out "mark". When I do so, I want you to draw a slant mark after the word you have just completed, or, which you are writing at the time. Continue writing the names of four-legged animals until I tell you to stop. Are there any questions?

CHAPTER III

EXPERIMENTAL RESULTS AND THEIR INTERPRETATION

A. Preliminary Observations.

All in all fifteen subjects were contacted and participated in the experiment to some extent. Ten subjects began initially but three of these dropped out after the second trial. One subject in the early stages simply could not give himself to free association. And another of the original subjects became upset emotionally over the process of free association and requested discontinuance. As these five subjects dropped out others were recruited in order to maintain the total of ten.

The task in continuous-free association was administered first because it was felt that this type of association was more difficult to communicate. In two instances it was quite clear that the subjects did not freely associate the first time they tried it. So the initial trial was disregarded in each of these instances.

It might be pointed out also that the five psychology concentrators who acted as subjects had had previous acquaintance with free association and consequently experienced little or no difficulty.

The experimental sessions varied somewhat as to time of day in order to accomodate the schedules of the students. However for each subject the same hour of the day was employed throughout the experimental sessions. One group of four subjects met regularly. But this was the largest single group. The remaining subjects were scheduled individually or in groups of two. The same room was used for all subjects and precautions were taken to reduce outside noise to a minimum. In addition the subjects were so seated in the room that they were not in each other's view nor in the view of the experimenter.

The experimental sessions continued overall for a period of about six weeks. The continuous-controlled sessions were commenced immediately following completion of the continuous-free association trials. In both sessions at least one full day elapsed between each trial.

The two-minute time intervals in both sessions were called out by the experimenter, the writer of this thesis. For this purpose a standard stop-watch was used. At first the question existed as to whether such a method was sufficiently accurate or not. However once experimenting was begun it became evident that greater precision was not required because the recording of associations by the various subjects was at no time so rapid that a second or two would have made any difference. If a subject had just commenced a word or was in the middle of a word when "mark" was called, he had been previously instructed to draw a slant after that word.

In the following presentation and interpretation of experimental results the continuous-free association data will be treated first and following this a discussion of the continuous-controlled data.

B. Interpretation of Continuous-Free Associative Data From a Semantic Point of View.

The continuous-free associative task resulted in ten thousand responses. The words from each subject were transcribed from the original data sheets, alphabetized in a second transcription, and finally analyzed for semantic similarities. Analyzing of this data for the first three subjects suggested that common responses were highly infrequent and seemed to occur in a random manner. What is indeed quite surprising is that in continuous-free association no responses common to all ten subjects were found. This was determined first by listing the nine common responses found for subjects a, b, and c. Then the responses of subjects d ... j were scanned. None of these nine responses were common to all subjects.

Consequently the logical-empirical criterion of response commonality is not met. As a result hypotheses I (a), and by implication, I (b) and I (c) are not confirmed. It is concluded then that under the conditions present in this experiment group frequencies of common response in continuous-free association cannot be applied univocally to each individual subject comprising the group.

An approximate, empirical probability, P, for common continuous-free responses occurring can be computed from

$$P = \frac{\sum_{i=1}^{i=10} c_i}{\sum_{i=1}^{i=10} n_i} \qquad \text{(I)}$$

where c_i represents the number of common continuous-free responses on any trial and n_i represents the overall continuous-free responses on any trial. These values for subjects a, b, and c are given in Table I. Substituting in equation I yields

$$P = \frac{(9)}{(3000)} = 0.003$$

So three times out of a thousand do subjects a, b, and c give common responses in the same task over ten trials.

A question arises, however, as to how representative subjects a, b, and c are when we make reference to all ten subjects. Some light can be cast on this question by

Table I.-

Number of Responses per Trial (n_i) and Common Responses per Trial (c_i) in Continuous-Free Association for Subjects a, b, and c for each of Ten Trials.

Trial	n_i	c_i
1	300	4
2	300	0
3	300	2
4	300	0
5	300	1
6	300	0
7	300	0
8	300	0
9	300	2
10	300	0

comparing the common responses of subjects a, b, and c in the continuous-controlled association (see Section C, first paragraph). In the latter data we can employ a similar expression for the empirical probability of obtaining a common response,

$$P_c = \frac{\sum_{i=1}^{i=10} c_i'}{\sum_{i=1}^{i=10} n_i'} \qquad \text{(II)}$$

where c' and n' represent corresponding values for the continuous-controlled responses. For subjects a, b, and c there results

$$P_{c(a,b,c)} = \frac{(872.0)}{(1917.6)} = 0.454$$

and for all ten subjects,

$$P_{c(a \ldots j)} = \frac{(2968)}{(6392)} = 0.464$$

In order to compare the sample of responses from subjects a, b, and c with the common responses for all ten subjects the following procedure was employed: the "hypothetical true" proportion of the group as a whole is known i.e. $P_{c(a \ldots j)} = 0.464$. Therefore because 0.464 is the criterion in this case one can estimate the "true" standard error (SE_p) of the sample by employing this "true" proportion as follows,

$$SE_p = \left[\frac{P_{c(a...j)} \, (1 - P_{c(a...j)})}{n_{i(a,b,c)}} \right]^{\frac{1}{2}}$$

$$SE_p = \left[\frac{(0.464)\ (0.536)}{1917.6} \right]^{\frac{1}{2}}$$

$$SE_p = 0.011$$

The original difference between sample and "true" proportions is (0.464 - 0.454 = 0.010). This difference is less than SE_p and we may attribute such a difference to chance. Hence it can be concluded that a proportion based on the responses of subjects a, b, and c is roughly representative of all ten subjects in continuous-controlled association. There is no compelling reason to assume that these same three subjects are not approximately representative of all ten subjects in continuous-free association. It should be stressed however that the above SE_p is highly approximate.

It is of interest to make a further tentative analysis of the common, continuous-free associative responses. As indicated above for subjects a, b, and c an approximate probability for a common response is given by $P = \frac{9}{3000} = 0.003$. The mean number of common responses, per subject, per trial is 0.30. The sum of the probabilities, P, for K common words per subject per trial can be estimated from the Poisson function.

$$\sum_{K=0}^{N} F_K = \sum_{K=0}^{N} \frac{e^{-m} m^K}{K!} \qquad \text{(IV)}$$

where

$$\sum_{K=0}^{N} \frac{e^{-m} m^K}{K!} = e^{-m} \left(1 + \frac{m}{1!} + \frac{m^2}{2!} + \frac{m^3}{3!} + \ldots + \frac{m^N}{N!}\right)$$

e is the Napierian logarithmic base; m is the mean frequency of common responses per subject per trial. When m = 0.30, e^{-m} = 0.74082. Substituting in equation IV there results

$$\sum_{K=0}^{N} \frac{e^{-m} m^K}{K!} = (0.740 + 0.222 + 0.033 + 0.003 + \ldots)$$

The terms on the right hand side give the respective probabilities of obtaining (0, 1, 2, 3 ...) common responses per subject, per trial. Evidently with these three subjects it would be a highly infrequent occurrence in the long run to obtain any common responses in continuous-free association especially when we consider the probability of obtaining zero common responses (0.74) per subject, per trial.

C. Interpretation of Continuous-Controlled Associative Data From a Semantic Point of View.

There were approximately 6,392 responses resulting from the continuous-controlled associative task. These responses were first transcribed from the original data sheets, alphabetized in a second transcription, and then analyzed for semantic similarities. As the responses of each subject were analyzed sequentially it became evident that words common to all subjects were given. All in all, 2,968 such common responses occurred. This latter figure represents the number of times that thirty-four different (common) words occurred in toto. These words are listed in Table II along with the average number of trials in which they appeared. It can be seen that these are all very familiar words although somewhat surprisingly other familiar words were not given by all subjects even though the opportunity presented itself over ten successive trials.

Table II shows that, on the average, common responses are given in the majority of trials. Further analysis discloses that over fifty percent of common responses are given on six or more trials. The common responses constitute about forty-six percent of all continuous-controlled responses.

Table III lists the number of common responses for each subject over ten trials. These are coded, common responses which were arrived at by subtracting 20 from each entry in the table. If hypotheses II (b) and II (c) are valid, then we should find no significant differences between the response frequencies of rows and columns in

Table II.-

Tabulation of Common C-C Responses with Mean Trial Frequency[1] for ten Subjects on ten Trials (Grand Mean = 8.70).

Response	Mean Trial Frequency	Response	Mean Trial Frequency
ant-eater	8.9	horse	9.8
antelope	8.4	lion	9.6
bear	9.2	mink	8.1
buffalo	9.4	moose	7.9
cat	10.0	mouse	8.1
chipmunk	7.5	mule	6.8
cow	9.9	panther	8.3
deer	9.7	pig	9.4
dog	9.6	rabbit	8.4
donkey	7.4	rat	7.6
elephant	9.8	rhinoceros	9.0
elk	8.0	sheep	7.3
fox	8.4	skunk	8.0
giraffe	9.2	squirrel	9.2
goat	8.5	tiger	9.9
groundhog	7.5	wolf	9.2
hippopotamus	8.3	zebra	10.0

1 Mean Trial Frequency $= \dfrac{\sum_{i}^{N} f_t}{N_i}$ where f_t is number of trials in which a common word appears and N_i is number of individuals giving such responses.

Table III.-

Coded Frequency of Common Responses[2] for Ten Subjects on Ten Trials

Subjects	Trials									
	1	2	3	4	5	6	7	8	9	10
a	13	11	14	14	14	14	13	13	14	8
b	1	1	7	10	8	10	10	12	12	12
c	0	3	7	7	6	6	9	6	9	10
d	4	8	5	9	12	11	12	11	11	11
e	8	9	14	13	13	14	13	10	8	13
f	13	11	14	14	14	14	13	13	14	8
g	7	5	8	9	12	12	12	10	11	9
h	8	6	6	7	12	10	11	11	7	11
i	7	11	11	11	12	11	12	10	11	11
j	0	2	3	5	6	4	7	7	6	7

2 For instance, the entry, 13, means that subject a on his first trial gives (33 - 20) or 13 of the 34 common words being studied. A coded frequency is an obtained frequency minus 20.

Table III. Because the same subjects were used for the successive trials, the condition of independence for Analysis of Variance is approached by studying the results at the level of each subject individually. If group frequencies of common responses apply to each and every individual, then the estimated of overall variance from rows and columns should be the same within chance error.

Table III is considered as a two-way classification, mixed model where the subjects are considered random and the successive trials fixed. The maximum coded score in a cell is 14 or (34 - 20).

The sums of squares was analyzed according to the general model:

$$\sum_{r=1}^{R}\sum_{c=1}^{C}\sum_{i=1}^{n}(X_{rci} - \bar{X}_{...})^2 = nC\sum_{r=1}^{R}(\bar{X}_{r..} - \bar{X}_{...})^2$$

$$+ nR\sum_{c=1}^{C}(\bar{X}_{.c.} - \bar{X}_{...})^2$$

$$+ n\sum_{r=1}^{R}\sum_{c=1}^{C}(\bar{X}_{rc.} - \bar{X}_{r..} - \bar{X}_{.c.} + \bar{X}_{...})^2$$

$$+ \sum_{r=1}^{R}\sum_{c=1}^{C}\sum_{i=1}^{n}(X_{rci} - \bar{X}_{rc.})^2 \qquad (IV)$$

where r and R refer to rows; c and C to columns, $\bar{X}$ is a mean value; $\bar{X}_{...}$ is the overall mean; $\bar{X}_{r..}$ is a row mean and $\bar{X}_{.c.}$ is a column mean.

In the instance where there is one entry per cell, i.e. where n = 1, equation IV reduces to the following special case:

$$\sum_{r=1}^{R}\sum_{c=1}^{C}(X_{rc} - \bar{X}_{..})^2 = C\sum_{r=1}^{R}(\bar{X}_{r.} - \bar{X}_{..})^2 + R\sum_{c=1}^{C}(\bar{X}_{.c} - \bar{X}_{..})^2$$

$$+\sum_{r=1}^{R}\sum_{c=1}^{C}(X_{rc} - \bar{X}_{r.} - \bar{X}_{.c} + \bar{X}_{..})^2 \qquad \text{(V)}$$

where the first term on the right is the between-rows sums of squares; the second term is the between-columns sums of squares; and the third term is the residual or interaction sums of squares. Equation V is the model which is to be followed in the present analysis of data.

From Table III the following values are derived,

Row totals, squared and summed: $\sum_{r=1}^{R} T_{r.}^2 = 96{,}291$

Column totals, squared and summed: $\sum_{c=1}^{C} T_{.c}^2 = 92{,}811$

Total cell entry, squared and summed: $\sum_{r=1}^{R} \sum_{c=1}^{C} T_{rc}^2 = 10{,}233$

The sum of squares (SS) in computational form are as follows:

SS rows: $\frac{1}{C} \sum_{r=1}^{R} T_{r.}^2 - \frac{T^2}{N} = \frac{96{,}291}{10} - \frac{900{,}601}{100}$

$= (9629.10) - (9006.01)$

$= 623.09$

Degrees of freedom $= (R - 1) = 9$

s_r^2 = Variance estimate for rows $= \frac{623.09}{9} = 69.23$

SS columns: $\frac{1}{R} \sum_{c=1}^{C} T_{.c}^2 - \frac{T^2}{N} = \frac{92{,}811}{10} - \frac{900{,}601}{100}$

$= (9281.10) - (9006.01)$

$= 275.09$

Degrees of freedom $= (C - 1) = 9$

s_c^2 = Variance estimate for columns $= \frac{275.09}{9} = 30.55$

<u>SS interaction</u>: $\sum_{r=1}^{R}\sum_{c=1}^{C} X_{rc}^2 - \frac{1}{C}\sum_{r=1}^{R} T_{r.}^2 - \frac{1}{R}\sum_{c=1}^{C} T_{.c}^2 + \frac{T^2}{N}$

$= 10{,}233 - 9629.50 - 9281.10 + 9006.01$

$= 328.81$

Degrees of freedom $= (R - 1)\ (C - 1) = 81$

S_1^2 = Variance estimate, interaction $= \frac{328.81}{81} = 4.06$

<u>SS total</u>: $\sum_{r=1}^{R}\sum_{c=1}^{C} X_{rc}^2 - \frac{T^2}{N} = 10{,}233 - \frac{900{,}601}{100}$

$= (10{,}233) - (9006.01)$

$= 1226.99$

Degrees of freedom $= (N - 1) = 99$

Table IV summarizes the foregoing calculations and lists the Variance estimates to be entered into the F ratios. For a mixed model (n = 1) the error term for row effects is s_{rc}^2. Thus $F_r = s_r^2/s_{rc}^2$. The error term for column effects is also s_{rc}^2. Thus $F_c = s_c^2/s_{rc}^2$. In the above data these F ratios are

$$F_r = \frac{s_r^2}{s_{rc}^2} = \frac{69.23}{4.06} = 17.0$$

$$F_c = \frac{s_c^2}{s_{rc}^2} = \frac{30.55}{4.06} = 7.5$$

Table IV.-

Analysis of Variance for Data of Table III.

Source of Variation	Sums of Squares	Degrees of Freedom	Variance Estimate
Rows (subjects)	623.09	9	$s_r^2 = 69.23$
Columns (trials)	275.09	9	$s_c^2 = 30.55$
Interaction	328.81	81	$s_{rc}^2 = 4.06$
Total	1226.99	99	

These obtained values are significant at the 1% level. In other words from F_c we know that as one proceeds from one trial to the next the number of common responses given by the ten subjects varies more than one would expect on the basis of chance fluctuation. A similar judgment can be made on the basis of F_r, i.e. the variation in common responses given by the various subjects is greater than chance expectation.

These results can be applied to an evaluation of the three hypotheses: II (a), II (b), and II (c).

Hypothesis II (a) is confirmed. There are continuous-controlled responses common to all subjects.

Hypothesis II (b) is not confirmed by this data. The number of common (c-c) responses differs significantly among the ten subjects since F_r is significant. With the data obtained there is reason to question the general assumption that group frequencies (c-c) of response (associative-strength) can be directly applied un a univocal manner to the individual subject.

Hypothesis II (c) is not confirmed. If we consider each trial as representing a group of responses these also vary significantly as shown by F_c being significant. So that a group tabulation of common responses obtained on one occasion as compared to another occasion may represent a significant difference. Hence with continuous-controlled associative material as employed in this study there is even question about employing one group sampling of responses as typical of the group in the long run.

However the fact that common responses are given in continuous-controlled association does indicate a semantic link between group and individual. So that the aufgabe for continuous-controlled association does mobilize similar associations in each subject. This fact alone would not be very impressive but, as we have already seen that overall these common responses are quite frequent (46% of the total). In addition, the time rate of emission of common responses for all subjects on all trials conforms to a mathematical function similar to that for the overall or total number of continuous-controlled responses. This latter interpretation will be discussed in Section D.

D. Interpretation of Continuous-Controlled Associative Data According to Their Mathematical Distribution as a Corollary to the Rejection of Hypotheses II (b) and II (c).

The common continuous-controlled responses were made by all subjects on each trial over nine two-minute time periods. From these data it is possible to plot a curve of responses which reflects the rate at which such responses are given by all subjects on all trials. As each two-minute time period elapses the number of responses decreases in a curvilinear manner. Bousefield and Sedgewick constructed such a curve of responses for all the continuous-controlled responses (not common responses). As indicated earlier they found that such a curve was fit by the function, $n = c\ (1 - e^{-mt})$.

It was decided to compare the overall continuous-controlled responses with the common continuous-controlled responses in order to determine whether the same type of mathematical distribution obtains. However before going on to this problem some clarification of Bousefield and Sedgewick's method should be made.

These authors asked college students over successive two-minute time intervals to name fellow students, animals of two or more syllables, cities, carpenter's tools, and so on. Cumulative frequencies of responses given were generally fit by the function, $n = c\ (1 - e^{-mt})$ where n is the cumulative number of responses, t is elapsed time, and e is the Naperian logarithmic base. The constant, c, gives the asymptote or limiting number of responses with unlimited time, and m is the rate of change of the curve to that asymptotic point. Thus the value, c, is inferential; it is never actually obtained. Bousefield and Sedgewick conclude that the rate of continuous-controlled associative responses is proportional to the number available or remaining at any time.

They present a rather lengthy derivation of the equations whereby c and m may be obtained. Their method is essentially one of selected points which are separated sufficiently so that one value, n_3, is at least twice the value of the other, n_2. The asymptotic value for c is obtained by substituting in the following equation,

$$c = \frac{n_2^2}{2n_2 - n_3}$$

The constant, m, is obtained from

$$e^{mt_2} = \frac{n_2}{n_3 - n_2}$$

or, taking logs,

$$m = \frac{1}{t_2} \left(\log_e n_2 - \log_e (n_3 - n_2) \right)$$

These solutions are approximate and like any method of selected points in curve fitting are most accurate only if all the points fall on or very near the theoretical curve which is computed from n_3 and n_2.

As one might expect the continuous-controlled data obtained in the present thesis is such that most of the plotted points do not fall exactly on a curve. Various substitutions in the above equations from Bousefield and Sedgewick yielded constants that even by inspection did not adequately represent the data. The difficulty is two-fold; (1) the continuous-controlled responses in this thesis begin with relatively high frequencies and (2) they decrease quite rapidly. As a result in order to employ an n_3 which is at least twice n_2, the first value of 24.55 responses must be used. And apparently this necessity introduces distortion. Bousefield and Sedgewick encountered a somewhat similar problem in some of their data and resolved it by adjusting the initial, obtained value. It seems, however, that this becomes a little too approximate.

In order to determine a value for c and m in the continuous-controlled data from this thesis, it was finally decided to employ a different approach. The function, $n = c\ (1 - e^{-mt})$, is obtained by integrating the following differential equation,

$$d\,n = m\ (c - n)\ d\,t \qquad \text{(V)}$$

where the symbols represent values given earlier. It is to be noted that d n represents a change in the cumulative number of responses given. An approximate solution of equation V was made by initially assuming an arbitrary value for c. Then in solving for m the criterion becomes

$$c = n + \left(\frac{1}{m} \right) \left(\frac{d\,n}{d\,t} \right)$$

$$\text{and } m = \left(\frac{1}{c - n} \right) \frac{d\,n}{d\,t}$$

where m is a constant, unknown value; n, $\frac{d\,n}{d\,t}$ are given; and c is successively approximated.

For the overall, cumulative, continuous-controlled responses in the data obtained the squares of the residuals were least for c = 66 and m = 0.1552. The values for d n are plotted graphically in Fig. 1. The data are reasonably fit by theoretical values from the differential equation. So as an approximation we can express an integrated form of equation V as $n = 66\ (1 - e^{-0.1552t})$. This is represented in curve (A), Fig. 2. The quantity, 66.0, accurate to within 1.0 response, indicates the average limit for naming four-legged animals with unlimited time and 0.1552 is a measure of the rapidity with which this limit or asymptote is approached.

In analyzing the distribution of common continuous-controlled responses a comparison can now be made with the overall distribution of continuous-controlled responses. These data are summarized in Table V. The cumulative response columns indicate that the common responses represent a reduction of approximately $(1 - \frac{29.68}{63.92})$ or fifty-three percent.

The cumulative distribution of common continuous-controlled responses was analyzed in a manner similar to the above analysis of the total continuous-controlled frequencies per time interval. Both curves on visual inspection seemed to be similar, so the differential equation, $d\ n = m\ (c - n)\ d\ t$, was again employed. The same method of arbitrary substitution

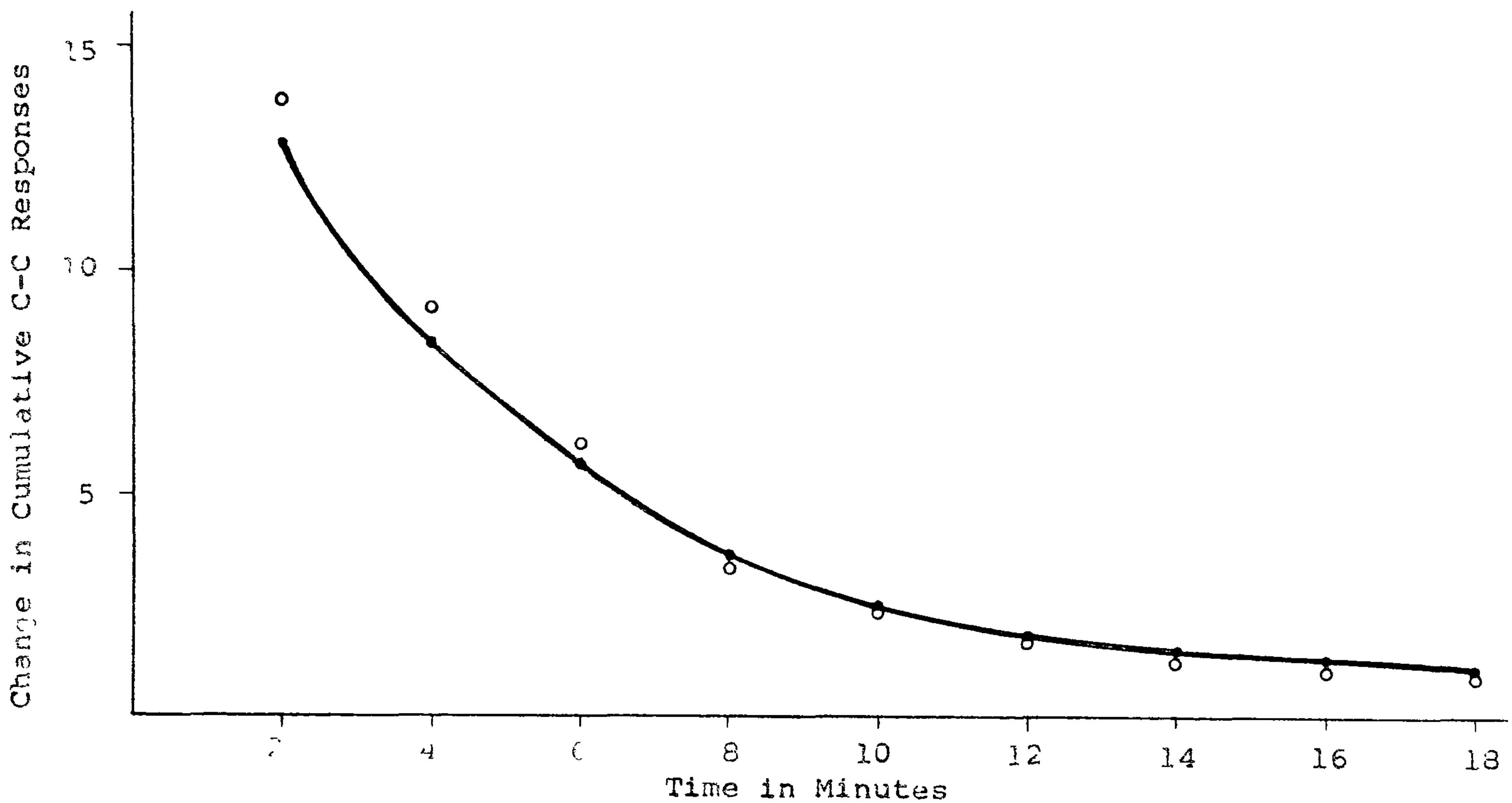

Fig. 1 Plot of $dn = 0.155\,(66 - n)\,dt$. Solid Line Represents Theoretical Values; Circles, Obtained Values.

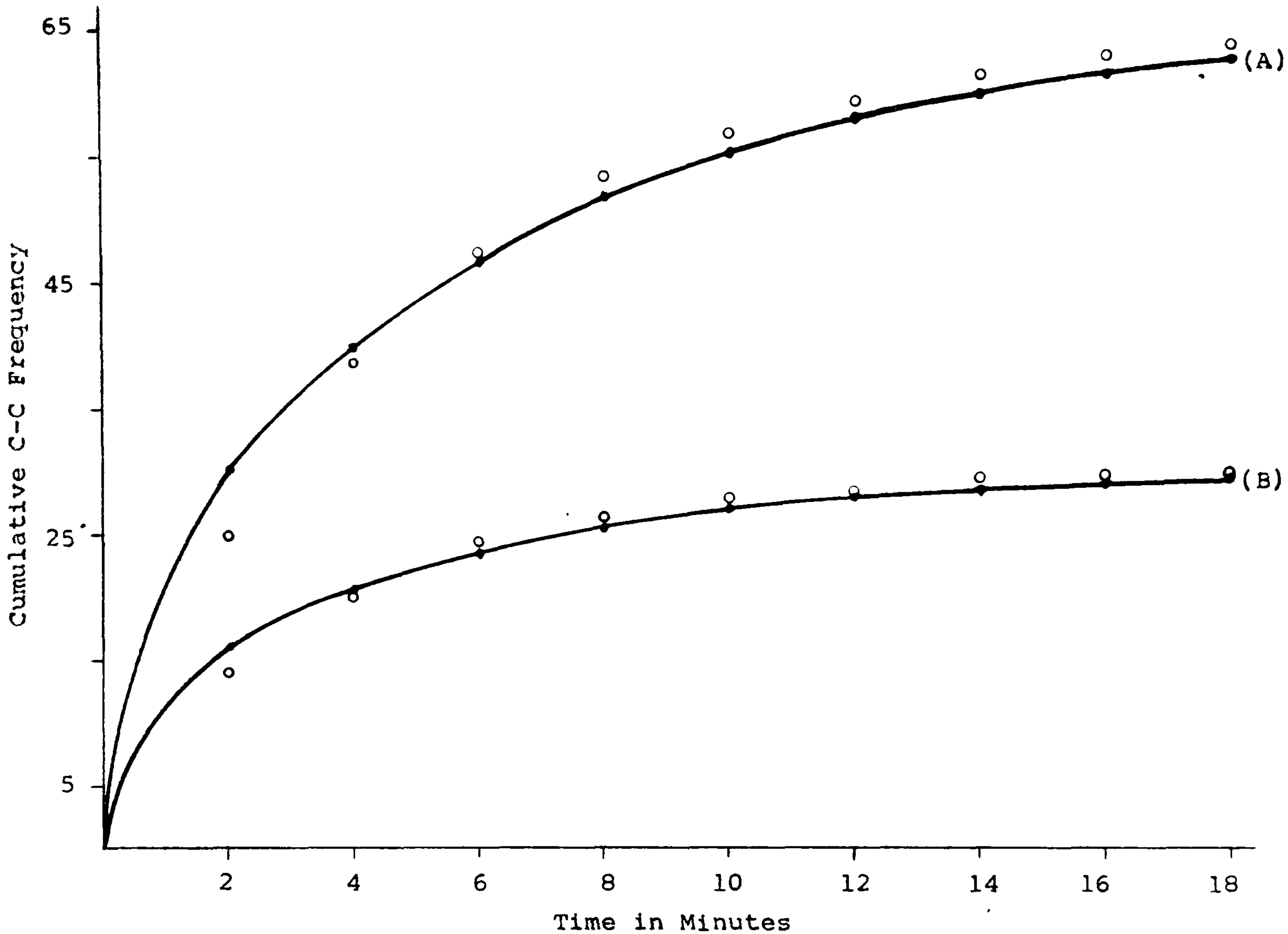

Fig. 2 Cumulative C-C Responses (A) and Common C-C Responses (B) Plotted against Elapsed Time.

Table V.-

Overall Continuous-Controlled Responses along with Common Responses, per Time Interval.

Overall Responses per Time Interval	Common Responses per Time Interval	Overall Cumulative Responses	Common Cumulative Responses
24.55	14.29	24.55	14.29
13.91	6.08	38.46	20.37
9.06	3.77	47.52	24.14
6.10	2.47	53.62	26.61
3.51	1.29	57.13	27.90
2.53	0.72	59.66	28.62
1.81	0.52	61.47	29.14
1.40	0.37	62.87	29.51
1.05	0.17	63.92	29.68

for c and then successive approximations to determine the value of m was employed. The squares of the residuals were least for $m = 0.192$ and $c = 30.0$. Inspection of Fig. 3 indicates a reasonably good fit. So the empirical equation, $n = c\ (1 - e^{-mt})$, becomes $n = 30.0\ (1 - e^{-0.192t})$. Fig. 2 is a comparative plot of this equation in curve (B) along with the earlier, overall, c-c curve (A) where $n = 66.0\ (1 - e^{-0.155t})$.

Curve (B) indicates that the experimental group virtually achieved the theoretical limit of common responses (30.0). The exponential value, $m = 0.192$, indicates that this limit was approached relatively more rapidly than in curve (A) where $m = 0.155$ and where the theoretical limit is 66.0 responses.

So a conclusion which can be formulated here is that the cumulative rate of emission of common continuous-controlled responses can be expressed by the same mathematical function as fits the cumulative emission of the total continuous-controlled responses. This is a rather thought-provoking result in the sense that it tells us that whatever produces common responses in all subjects also does so at a mathematical rate similar to that which produces the total of continuous-controlled responses. The one very general, dynamic variable present in continuous-controlled association is the set or aufgabe which restricts associations

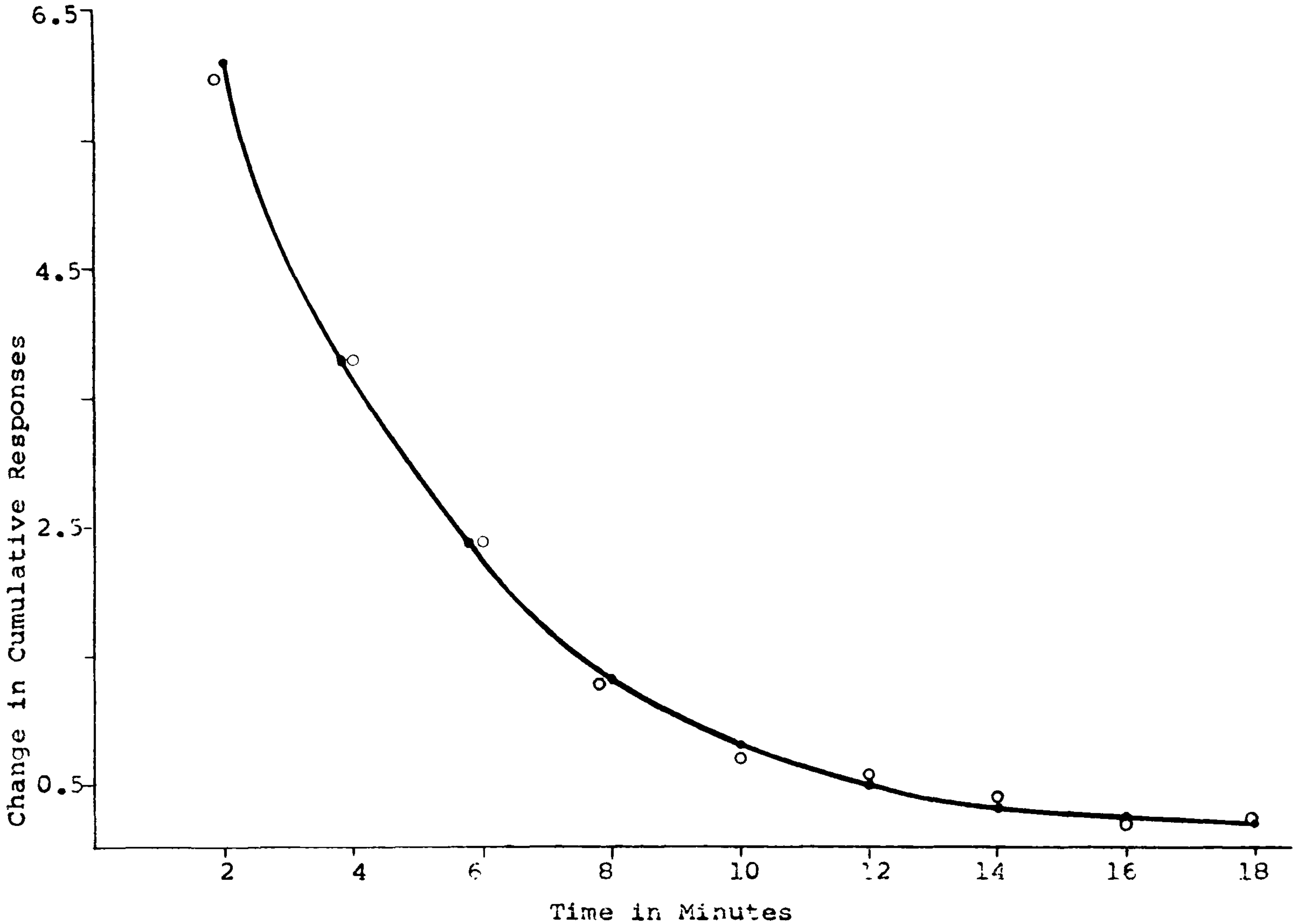

Fig. 3 Plot of $dn = -0.192\ (30 - n)\ dt$. Solid Line Represents Theoretical Values; Circles, Obtained Values.

to one specific category (in the present instance, four-legged animals). And within this Aufgabe the rate of continuous-controlled associations is at each point proportional to the number of associations available or remaining.

SUMMARY AND CONCLUSIONS

The analysis of experimental results from continuous-free and continuous-controlled associative tasks casts some doubt on the feasibility of employing group frequencies of verbal response as a definition or measurement of response strength in the individual subject. It must be borne in mind, however, that the summary of results to follow does not deal with discrete-free and discrete-controlled types of word association.

The general hypothesis of no significant differences in frequencies between a group population of responses as compared to individual response populations was not confirmed in both continuous-free and continuous-controlled associative data.

The criterion of significance was based on two conditions: (1) commonality of response and (2) statistical frequency. As a further condition the criterion of statistical frequency applied only to common responses. Hence, if the commonality criterion was not met, the statistical criteria were per se not upheld.

In the continuous-free associative data commonality of response did not occur. Thus both of the above criteria, (1) and (2), were not supported. Above and beyond this lack of significance a further analysis of responses from subjects a, b, and c was undertaken in order to specify the degree to which commonality occurred even in a small group of subjects. A Poisson function showed that in the long run responses common to even three such subjects was highly unlikely.

In summary, then, with respect to continuous-free association: No common responses occurred. Therefore the hypotheses of no significant differences in frequency of response between subjects and groups (trials) was not confirmed.

In the continuous-controlled associative data commonality of response did occur. Thirty-four common responses occurred in approximately forty-six per cent of all continuous-controlled responses. Analysis of Variance disclosed however that the statistical frequency of common responses did vary significantly from one subject to another as well as from one group (trial) to another.

In general, then, the practice of defining or measuring associative-strength in an individual subject by tabulating group frequencies of response is to be questioned. Such question applies to both continuous-free and continuous-controlled type of data as employed in this study. However, in the continuous-controlled data common responses were given. So, apart from statistical frequency of occurrence, a semantic link between group and individual associative processes does exist and is seemingly related to the Aufgabe or set which the subject imposes on his associative processes.

Despite the lack of statistical significance, and, because common continuous-controlled responses do represent a general semantic link between group and individual associative sequences, further analysis was made. It was found that throughout nine successive two-minute time periods the rate of emission in common responses followed the same mathematical function as did the overall continuous-controlled associations for all subjects on ten trials. Thus the semantic link between group and individual with respect to common responses does reflect a pattern or regularity in this respect.

Many implications for further research suggest themselves as a result of the present findings:

The hypotheses of the present thesis should if possible be tested in relation to discrete-free and discrete-controlled associations. These types of association present distinct methodological problems, however. For example, how do you generate 1000 responses of the Kent-Rosanoff type in a single subject without (a) spending years collecting data and (b) avoiding reinforcement effects, serial effects, and so on, which might constantly modify the task? There is the question whether randomized presentations of stimuli could overcome these objections.

A design somewhat similar to the present study might well be applied to continuous-controlled data wherein varying restrictions of Aufgabe or set are attempted. In this connection it should be remarked that the present thesis as well as the related studies of Bousefield and Sedgewick employed concrete continuous-controlled categories such as cities, birds, carpenter's tools, names of fellow students, and so on. Some insight into higher mental processes might result in establishing whether commonality occurs in more abstract types of Aufgabe or set.

Finally, one wonders what kind of variations in commonality might occur in groups other than college and university students.

BIBLIOGRAPHY

Bousefield, W. A., "An Empirical Study of the Production of Affectively Toned Items", *Journal of General Psychology*, Vol. 30, 1944, p. 205-215.
Cumulative continuous responses for pleasant and unpleasant associations followed a similar type of log curve found in this study.

Bousefield, W. A. and W. D. Barclay, "The Relationship between Order and Frequency of Occurrence of Restricted Associative Responses", *Journal of Experimental Psychology*, Vol. 40, 1950, p. 643-647.
Authors presented a criterion of ordinality for associative-strength in continuous-controlled association. This formed one of the hypotheses of the present study but was not verified.

Bousefield, W. A. and C. H. W. Sedgewick, "An Analysis of Sequences of Restricted Associative Responses", *Journal of General Psychology*, Vol. 30, 1944, p. 149-165.
These authors presented a method of recording and plotting continuous-controlled associations in successive two-minute intervals which was incorporated importantly into the design of the present study.

Cofer, C. N., "Comparison of Word Associations Obtained by the Methods of Discrete Single Word and Continued Association", *Psychological Report*, Vol. 4, 1958, p. 507-510.
Author compared discrete-free and continuous-free associations to the same stimulus word. The number of continuous-free was much less in group but the number of different responses was greater. This points to one of the advantages of continuous type association.

Cohen, B. D., H. I. Kalish, J. R. Thurston, and E. Cohen, "Experimental Manipulation of Verbal Behavior", *Journal of Experimental Psychology*, Vol. 47, 1954, p. 106-110.
This study by Cohen et al showed how completely unaware a subject may be regarding surrounding reinforcers of a verbal and non-verbal sort.

Hall, J. F., "Controlled Association Time as a Function of Word Frequency", *Journal of Psychology*, Vol. 53, 1962, p. 77-80.
With discrete-free associative technique the author found no relationship between frequency and commonality of response. A thought provoking result in light of continuous-free results of present thesis.

Rosanoff, A. J., Manual of Psychiatry and Mental Hygiene, New York, John Wiley and Sons, Inc., 1938.
This work contains original data from one thousand subjects in discrete-free association to one hundred common words. Investigation of this list contributed to formulation of thesis problem.

Wilson, W. C. and W. S. Verplanck, "Some Observations on the Reinforcement of Verbal Operants", American Journal of Psychology, Vol. 69, 1956, p. 448-451.
Authors demonstrated that subtle vocal sounds could act as reinforcers to change significantly the content frequencies in continuous association. Their findings were related to those of Cohen et al. These two studies it is felt helped to avoid bias in subjects' responses due to cues from experimental surroundings.

APPENDIX 1

ABSTRACT OF

A Comparison of Group and Individual Response Probabilities in Continuous Verbal Association[1]

Verbal response frequency in a group of subjects is often employed as both a definition and measurement of verbal associative-strength. A fundamental theoretical question centers around the extent to which such group frequencies or probabilities are directly applicable to the range of responses within individual subjects comprising the group.

In the form of a specific hypothesis the above theoretical question may be stated as follows: Under given conditions there is no significant difference in the frequency with which verbal responses are elicited in a group of subjects as compared to the frequency with which they are elicited in each and every individual comprising the group.

To test this general hypothesis a population of verbal responses was generated in each of ten subjects on ten different trials. This data constituted a ten by ten array such that response frequencies in each subject could

1 John J. Fleming, Doctoral thesis presented to the School of Psychology of the University of Ottawa, Ontario, August, 1964, x-64 p.

be compared with equal frequencies in groups (trials) of subjects.

In order to generate a sufficient population of responses in each subject continuous-free and continuous-controlled association were employed. Comparison of group and individual responses was based upon two related criteria. If group and individual response populations reflected similar frequencies in word association then: (1) there must be responses common to all subjects, and in addition (2) there must be no significant difference in the frequency of common responses between group and individual response populations. Common responses were determined by arithmetic tabulation. Statistical significance was tested by an Analysis of Variance design wherein Variance estimates could be made from rows (subjects) as well as columns (groups or trials).

With respect to continuous-free association no common responses occurred. So it was concluded that in this type of data group frequencies of responses can not be directly applied to individual response populations.

With respect to continuous-controlled association common responses did occur. However the frequency with which they occurred in each subject varied significantly. The same statistically significant result applied to the variation in groups (trials) of responses.

As a general conclusion from the data, then, in continuous-free and continuous-controlled types of verbal association there is real question as to whether group frequencies of response can be directly applied to individual response populations. From the design of the present study however there was demonstrated a semantic link between group and individual responses in the case of continuous-controlled association. A restrictive Aufgabe to name four-legged animals did result in approximately forty-six per cent common responses whose rate of emission conformed to a mathematical function which has been found in other studies to be characteristic of this type of word association.