cy. #1

COMMUNICATION AND LEARNING IN TASK-ORIENTED GROUPS

LEE S. CHRISTIE
R. DUNCAN LUCE
JOSIAH MACY, JR.

LOAN COPY only

TECHNICAL REPORT NO. 231

MAY 13, 1952

RESEARCH LABORATORY OF ELECTRONICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS

The research reported in this document was made possible through support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the Army Signal Corps, the Navy Department (Office of Naval Research) and the Air Force (Air Materiel Command), under Signal Corps Contract No. DA36-039 sc-100, Project No. 8-102B-0; Department of the Army Project No. 3-99-10-022; and by the Rand Corporation.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS

Technical Report No. 231 May 13, 1952

COMMUNICATION AND LEARNING IN TASK-ORIENTED GROUPS

Lee S. Christie
R. Duncan Luce
Josiah Macy, Jr.

Abstract

Chapter 1: Task-oriented groups are discussed generally and definitions are introduced to isolate variables which were experimentally manipulated.

Chapter 2: A general description of the experimental techniques is given.

Chapter 3: The statistical behavior of individual message destination choice is coupled with communication network properties to account for the observed statistics of group performance. Emphasis is placed upon the learning which occurs during trials.

Chapter 4: Individual decision latency is shown to be approximately exponential. A simple theory relates group latency to individual, and these results are used to explicate other experiments for which less complete data are available.

Chapter 5: Attention is turned to noise in the coding-decoding of messages. Group errors are simply explicable in terms of a measure of noise. A mechanism, redundancy, is demonstrated to account for a decrease in noise. In turn, redundancy is related, imperfectly, to several network properties.

Chapter 6: Questionnaire attitudinal data are presented and for some questions a high correlation with network properties is demonstrated. A factor analysis yields four orthogonal factors for the questionnaire used.

Appendix 1: Detailed descriptions of specific experiments are presented.

Appendix 2: An electrical device which controls and records a class of communication experiments is described.

Appendix 3: A human group interpretation of the classical electrical network equations is shown to have very limited applicability.

Appendix 4: Some mathematical results on network topology and their experimental implications are recounted.

An experimental group is shown writing and sending messages. For a description of the apparatus, see section II.

TABLE OF CONTENTS

COMMUNICATION AND LEARNING IN TASK-ORIENTED GROUPS

PREFACE

1. Introduction

This study reports the principal theoretical and experimental developments of the "Group Networks Laboratory" of the Research Laboratory of Electronics, M.I.T., during the two and one-half years from August 1949 to April 1952. To provide a frame of reference for this work, a brief account of the background and development of the theoretical concepts and experimental techniques will be given in this preface. As in all research in a relatively new field, the development of ideas owes much to the influence of a large number of related papers and publications, and to conversations and interchanges of ideas with many workers. We can give explicit credit to only a few of them here; however, specific references are made throughout the text, and a fairly complete bibliography of related publications is included.

2. History of Research

The initial developments in this work can perhaps be traced most directly to Kurt Lewin, whose pioneer contributions to the study of group dynamics were extended over many years and ended only with his death in 1947. Lewin was responsible for the introduction of a particular concept of psychological space and its applications in the study of groups. He strongly emphasized the role of motivational concepts in group studies. At present, this emphasis is more characteristic of the work of the Research Center for Group Dynamics at the University of Michigan, the extension of the group Lewin started, than of the Group Networks Laboratory at M.I.T.

Our work, initiated by Alex Bavelas, one of Lewin's students, has taken a route more influenced by new developments of a mathematical nature and has been less concerned with the dynamics of the psychological situation than the Michigan group. Bavelas, in a dissertation done under Lewin in 1947, examined an aspect of the internal structure of groups (the part we now term the communication network), and he suggested some mathematical measures that might be pertinent to the study of group behavior. In the spring of 1948, Bavelas presented a graduate seminar at M.I.T. on both his and Lewin's ideas on group behavior. The members of this seminar, largely students of engineering and physics, became interested in extending both the mathematical notions and developing experiments that would test their relevance. Before the term was concluded, an apparatus was built and pilot experiments were conducted. Interest in the research remained high, and Bavelas was granted a small budget to continue the experimental work. One member of the seminar, H. J. Leavitt, did his dissertation in this field in the summer of 1948, and his experiments were repeated with female subjects by S. L. Smith, who had been engaged by Bavelas as an experimental assistant early in the summer of 1948. In the fall of 1948, R. D. Luce and A. D. Perry, graduate students in the Department of Mathematics, were engaged as consultants for one term.

Bavelas' seminar, including many members of the original group, met again during the winter of 1948-49. Theoretical discussion continued and several exploratory experiments were carried out. In the fall of 1949, Bavelas obtained a contract from the RAND Corporation for partial support of his work, and early in 1950 he augmented his staff by employing two graduate students of mathematics, R. Abelson and A. Simmel. During this period the work attracted wide interest, and men from diverse fields contributed to the seminars. In particular, in the spring of 1950, O. H. Straus and W. H. Huggins, of the M.I.T. Research Laboratory of Electronics and the Air Force Cambridge Research Laboratory, respectively, attended these seminars, and from their knowledge of systems analysis contributed appreciably to the theoretical notions being developed.

During the summer of 1950, the research group, which then consisted of Bavelas, Straus, Luce, F. Barrett, J. Macy, Jr., and S. L. Smith, became an integral part of the Research Laboratory of Electronics and moved to new quarters in that laboratory. The move permitted the group to make use of the extensive service and shop facilities of R.L.E., and resulted in greatly improved conditions for experimental work. In the fall of 1950, Bavelas and Straus left the group to undertake specialized work connected with the defense program, and Smith left to resume graduate studies. At this point Luce and Macy were named to head the project jointly. In the summer of 1951, Simmel again joined the group, and in the fall of the same year L. S. Christie joined the staff. At the end of 1951, Barrett and Simmel left the laboratory to assume other positions.

3. Development of the Field

Initially, the research in the seminars was exploratory. Bavelas had developed some mathematical parameters for the communication linkage (the network) between members of a group of people. It was conjectured that networks having different values of these parameters would have noticeably different effects on groups. On the basis of the idea that the network is an important determinant of group behavior, the first experiments were designed to demonstrate the existence of such an effect. This general proposition was shown to be true, but not always with the expected relation of performance to the network parameters. This led to additional mathematical investigation to ascertain parameters which did have a relation to the experimental results. As the work progressed, a need was felt for mechanisms which would account for the discovered relationships in more basic terms – in, say, the characteristic behavior of the people composing the group.

At this time considerable attention had been attracted by Norbert Wiener's book Cybernetics which emphasized the parallel of feedback concepts in electromechanical systems and social systems, and which discussed the new ideas of information theory. Gradually, it became quite evident that these concepts might be fruitfully applied to obtain a better understanding of group behavior. Wiener participated in several conferences, and there is no doubt that his ideas have affected the theoretical development of our work. The first actual application of these ideas occurred in the winter of 1949-50 through the participation of several men familiar with the methods and terminology of modern electrical communication theory. Their work included the restatement of the mathematics of electrical network theory in terms of group behavior (see Appendix 3), an emphasis that the group and its environment must be treated as a system, and a discussion of the Shannon-Wiener theory of information. These latter notions allowed, for example, an understanding and analysis of a specific experiment in which noise had occurred in the coding of messages (see Chap. V).

The task of applying the existing techniques of electrical systems analysis was found to be not trivial, for many conditions satisfied in the electrical case are not generally found in human groups. Thus, during this period of development much time was spent considering the extent to which such an application is possible. In particular, the extent to which the psychological make-up of the individuals comprising the group may be neglected in theory and in experiment was studied. These discussions led to the feeling that the results which were of primary importance at this stage of the research were those which are independent of the variation in psychological character of the individuals constituting the group. We felt justified in considering that, in many situations, the effect of the sets of stimuli encountered in groups was roughly the same over a large range of personalities, and that for purposes of our research we could replace the individual with a fictional "normal" man whose responses were statistically distributed. This point is discussed at length in section I. 3. 3.

The most recent phase of the research has been a continued emphasis on the systems approach, with, however, the recognition that direct application of the existing electrical techniques is not likely to be possible. The experimental program has passed from the exploratory stage to more comprehensive examination of specific phenomena. Attempts have been made to obtain data which are statistically significant and these data have been subjected to much more detailed and searching analysis. These new data, coupled with the systems viewpoint, have allowed us to construct some simple probability models for the behavior of the groups which account, at least in part, for some of the relationships first discovered two to three years ago. This recent work has been primarily concerned with the objective phenomenon of message flow and has only incidentally been concerned with such psychological phenomena as the evolution of leadership, the morale of subjects, and the like. Thus, some of the correlations Leavitt obtained between subjects' attitudes and parameters of the network are no better understood today than they were when they were first discovered. The latest work indicates that the notions of systems analysis are a desirable way to describe the operations of the experimental groups, provided one does not try to carry over without change the particular mathematical forms that have been used in other applications, e. g. electrical engineering.

Thus, at the present time, some, but not most, of the problems have been solved.

This report is simply a slice out of a continuing program and not a final rounded summary.

4. Prerequisites

The work we are reporting is most cogently described as psychological, but much of it has been done by people not trained as psychologists. The nonpsychological members of the group have brought ideas and methods to bear on our problems which are often not well known to workers in the behavioral sciences. On the other hand, our psychologists use techniques, e.g. statistics, with which people trained in electrical engineering or pure mathematics are not always familiar. As a result, the methodology of this study is a coat of many colors. We have used at every point that technique which appeared likely to be most fruitful. Whenever the theoretical basis of a technique could not be given without undue lengthening of the study and undue duplication of readily available literature, we have given references.

The first chapter, which defines the area of study, assumes a familiarity with some of the elementary concepts and terminology used in mathematics when speaking of functions. More important, the basic notions of scientific analysis in terms of a "black box" are explained and used. This chapter is thus more readily understood if the reader is familiar with this technique, but such familiarity is not believed to be essential. In Chapters III through VI, many of the standard statistical techniques are freely employed in the analysis of the data, and notions of probability theory which lead to mathematics of varying complexity are occasionally employed. In Chapter V the concepts and some of the measures of communication theory as developed by Shannon and Wiener are used without elaboration.

The appendixes, with the exception of the first, require more specialized knowledge than the body of the report. The second describes an electrical device of some complexity; its technical details will be clear only to one with some electrical training. The third appendix rephrases the mathematics of linear electrical network analysis in terms of the group situation; this requires some knowledge of mathematical analysis and matrix theory. The last appendix is clearer if the reader has some sophistication in abstract mathematics, though this is not essential, since no proofs are given.

5. Acknowledgements

We would like to acknowledge the aid and assistance given us by the staff of the Research Laboratory of Electronics, who have been of great assistance in the design and construction of our apparatus, and have participated in stimulating discussions on our research. In addition to those staff members mentioned in the previous sections, we are indebted to J. C. R. Licklider, Robert Fano, and Samuel N. Mason for helpful theoretical discussion.

Within the project itself, much credit must go to F. D. Barrett and A. S. Simmel. Barrett aided in the design of Experiment 4, ran Experiment 5, and assisted in the preparation of Chapter I and Appendix I. Simmel also assisted on Chapter I and prepared the code for the problem solved on the computer Whirlwind (see Chap. III). In addition, invaluable assistance has been rendered by Deborah Senft, Patricia Thorlakson, and Eleanor Palmer, who performed most of the recent experiments and spent many painful hours reducing the mass of data to the numbers reported in this study; by John B. Flannery, who constructed most of the apparatus; and by Adassa Balaban, who spent long hours preparing type script of this study.

Outside the project, we are indebted to the Joint Computation Laboratory for Nuclear Science and R. L. E. for computation aid, in particular to Hannah Wasserman; and to Project Whirlwind, with special credit to John W. Carr III, and Manuel Rotenberg.

The critical aid of Jerome B. Wiesner, Alex Bavelas, and Joyce Harman, who have read portions of the manuscript, is greatly appreciated.

Finally, we are grateful to the Research Laboratory of Electronics and the RAND Corporation for the financial support which has made this work possible.

CHAPTER I – TASK-ORIENTED GROUPS

1. Introduction

There are found in society human groups whose function is to effect assigned changes in some part of their external environment. Our research is concerned with this class of groups, which have been termed task-oriented (15).* The following sections of this chapter make precise what we mean by the term, task-oriented group, but in this introduction we shall first discuss these groups informally.

The notion that effecting changes is the function of the task-oriented group implies that the members of the group attempt to cooperate with one another; thus any group which is subject to an analysis by the theory of games is not a task-oriented group, and conversely (105). It is true that a competitive group can, according to the game theory, effect changes on its environment, but this is not the primary function of the group; each member of the competitive group has as his goal the maximization of some utility for himself. Although in general a task-oriented group does not have subparts which are competitive, many competitive groups do have subparts which are primarily task-oriented. For example, in a purely competitive economy composed of units at the factory level the workers of a factory form a task-oriented group. The assigned goals of the factory derive from competitive considerations, but the process of attempting to achieve these goals often is not competitive. More technically, coalitions of players of a game are, at least temporarily, task-oriented. It is further often true that a large task-oriented group fractures naturally into several smaller ones which may, to a first approximation, be studied independently. Thus a study of a small group may well provide information about larger systems.

Basic to any notion of a task-oriented group is the concept of communication. A group can only be said to exist if each member is capable, at least indirectly, of a pertinent influence on each other member. Basic to the notion of a task-oriented group is, in addition, communication outside the group, for, as we mentioned, a group will be termed task-oriented only if it has a pertinent influence on the external environment, which in turn has an influence upon the group. We are thus led to make use of the recent developments in the fields of communication engineering, and in particular of the theory of information enunciated by Shannon and Wiener and elaborated by Fano (51, 52, 59, 60, 62, 65). This theory, while a great aid, is not sufficient for our purposes, since it is concerned far more with problems of coding of information and channel capacity than it is with the intuitive notions of information. Often, for the lack of a more sophisticated theory, we shall have to work with primitive concepts of information.

In addition, we shall draw heavily on the treatment of control systems as developed in recent decades by communication engineering. The theory of an automatic control

---

*References to the bibliography will be indicated by numbers in parentheses.

system is concerned with a "master" component which governs the behavior, to some extent, of its "slave" component by means of signals which control the release of energy. The manner in which the master does this is dependent on its fixed structure, on that of the slave, and on signals fed back from the slave. This two-way flow of communication is the essential characteristic of the interactive or feedback system, and as we remarked above, must be an essential property of the task-oriented group; the group being the master and some portion of its environment being the slave. The function of the group is to exercise control over this portion of the environment in order to make it assume some desired state. Effecting this desired state is the "task" of the group.*

The problem of control in a feedback system is comparatively simple if the slave is subject, essentially, to inputs only from the master, and if the effect of these inputs is predictable to a certain degree of accuracy. In electromechanical systems this is generally the case to a good approximation, both by design and application. When the slave is subject to important influences from other sources, the problem then becomes at best one of statistical optimization. Any complete study of task-oriented groups must take into account the existence of unpredictable and uncontrollable influences on the environment of the group.

The organizational structure of many task-oriented groups is specified; certain people are placed in roles of leadership, others carry out more perfunctory, but no less essential, jobs. In some groups restrictions are made as to who may communicate to whom and about what they may talk. Such specifications or limitations are usually imposed to organize the group's behavior so that it attacks its task effectively, and although these imposed restrictions may sometimes achieve this aim, often they do not. Thus one aspect of a study such as ours must be to determine the effects of such structural limitations.

To study such an area it has been necessary for us to select a method of obtaining empirical data and to develop a conceptual framework in which to place such data. There are two methods of obtaining data: controlled laboratory experiments and field studies.** Our choice is the former, and consequently, we have, for practical reasons, restricted ourselves to comparatively small groups. We shall discuss below both the faults and advantages of this choice. The details of the experimental technique employed are the subject, in fairly general terms, of Chapter II and, in more detail, of Appendix 1. The results of the experimentation are given in Chapters III through VI.

The conceptual scheme must have at least the property of including the type of group discussed above and of providing a useful vocabulary with which to discuss our

---

*A more general study of the similarities between electromechanical control systems and organic and social systems is given by N. Wiener, Cybernetics (10).

**Perhaps the most systematic field research on the behavior of organizations has been undertaken by H. A. Simon. See, for example, his A Study of Decision-Making Processes in Administrative Organization (44).

experiments. The most satisfactory form we have devised is comparatively abstract, being based on a few undefined notions in terms of which all other definitions are formulated. The remainder of this chapter presents this scheme.

2. The Relation of the Group to its Environment*,**

2.1. Node and Transfer Function

As the foundation of our treatment, we consider the concept of a "black box." This is considered to be any entity which is unanalyzable by choice or necessity, and which will be studied by investigating the relationships between the effect of the environment of the box on the box itself and the corresponding effect of the box on the environment. Any influence exerted on the black box by its external environment is termed an "input" to it. Any influence exerted by the black box on its external environment is termed an "output" from it. The basic element in this treatment is the assumption that the inputs to and outputs from the black box are related; that is, any set of inputs is operated upon in some manner by the black box to determine a corresponding set of outputs. The function describing the relation between possible inputs and outputs is called the "transfer function" of the black box. No assumption is made that these functions are single-valued, linear, or have any other specific mathematical properties. In many cases of engineering practice, it is possible to express the transfer function of the box in relatively simple mathematical terms, but this is not an inherent assumption. In most of the cases considered in this study, the transfer function will be at best of a statistical nature; in other words, it will relate the probability distribution of a set of outputs to a given input or statistical distribution of inputs.

In this type of analysis, no attempt is made to "open the box;" that is, no effort is made to determine its contents, or to explain why it performs as it does. True, in some cases in engineering analysis the concept of an equivalent circuit is used; this corresponds to the statement, "If the box contained this sort of material, it would act in the way we have observed." This concept is only an aid to calculation or heuristic description, and makes no pretense that the suggested "equivalent circuit" is really inside the box. Its sole purpose is to enable us to obtain and use the transfer function which is, of course, an empirical function. In addition, it is important to remember that this law need not be mathematical in the sense of calculus or analysis; it is indeed pleasant when it is an analytic function relating two variables which are complex numbers, as is often the case in physics; but the transfer function may equally well relate inputs which are sets of conditions on age and education, and lists of people with certain of these

---

*Psychologists may note that the treatment in the following sections is in many ways similar to behavioristic theory, although not identical.

**It is interesting to compare the following formulation with the more general and detailed study by Oskar Morgenstern, Prolegomena to a Theory of Organization (41).

properties, to an output which is a list of people fulfilling the given conditions. This type of transfer function occurs, for example, in many uses of IBM machines.

In accordance with electrical-engineering terminology, we call any such black box a "node," and with each node N a transfer function F_N, which will relate any input I_N to the output O_N which results from it, is associated. In symbols,

$$F_N(I_N) = O_N$$

where it is understood that the O_N so determined may be a unique output or set of outputs, or a probability distribution over a set of possible outputs. The transfer function of a node is thus operationally defined; that is, it is determined purely in terms of observable effects discovered experimentally and makes no assumptions about the "fundamental" internal nature of the node.

2.2. Task-Oriented Groups

Consider now a system composed of two such nodes, denoted C and E, and assume that the system is so constructed that some of the outputs from C are inputs to E, some of the outputs from E are inputs to C, and the remainder are to and from the rest of the environment. This is represented in Fig. I.1, the inputs and outputs being denoted by symbols as indicated. This class of system is quite general; any interactive two-element situation may be represented in these terms.

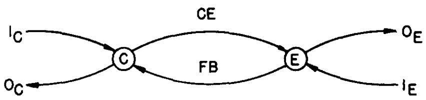


Fig. I.1

We then add some further specifications. If some part of the input I_C is of such a nature that it imposes a requirement to which the output FB from E should conform, we call that part of I_C a "task." We do not define a task as an input to C which states requirements on the output of C, but rather we define it as an "instruction" to C concerning the desired "behavior" of E. Of course, the input I_C may have many other functions as well; we merely define as a task that part of I_C which establishes such a requirement.

If part of the input I_C is a task, we call the node C a "control node." If the transfer function F_E of the node E is such that any change in the input CE to E results in a change in the output FB to C, we call the node E a "functional environment" to C. If the nodes C and E in such a system are a control node and a functional environment, we call the system a "feedback system." Of course, the functions of C and E are determined relative to a given system and situation, so that a given node may be considered a control node in one case, a functional environment in another, or neither in a third case, depending on the point of view and the particular system under consideration.

Using these terms, we may then define the area of our study: A group of human beings, or of human beings and machines, is a task-oriented group if it is the control node of a feedback system. As we remarked above, a given group may or may not be considered a task-oriented group, depending on the point of view under which it is examined. Examples of such task-oriented groups are many: research teams, a city council, a committee with a given objective, or a group of experimental subjects given a group task. A collection of people formed for purely social purposes is not a task-oriented group. Such a collection has no task in the sense we have defined, and so they are not part of a feedback system.

We may now assign descriptive terms to the various inputs and outputs of a feedback system, as follows:

I_C = external input to control
O_C = external output from control
CE = control of functional environment
I_E = external input to functional environment
O_E = external output from functional environment
FB = feedback (from functional environment to control).

2.3. Communication and Information

The process of transferring an output of one node to make that output an input to another node is called "communication." This transfer process must be effected by some physical means. Various methods are possible such as light or sound signals, electrical signals, or the transfer of material objects. The physical form which the communication takes may correspondingly be electromagnetic or sound waves, electrical impulses, marks on paper, and so forth. These physical units of the communicative process which carry semantic content can be described both by their content and by their identifying physical features. When we wish to refer to the former, we shall speak of "symbol contents," and when we wish to refer to the latter, "symbol designs." A set of such symbol designs used in communication from a given node to a given destination within a definite time interval is called a "message." Two such sets of symbol designs, with the same source and destination, are considered as two separate messages, if they are separated in time by an interval during which no communication takes place between that particular source and destination.

Any analysis of the interaction processes in a feedback system must examine the content of a message in addition to its physical form and its time and place of transmission. This examination of the content of a message involves a study of the quality usually referred to as information. It would obviously be desirable at this point to give a clear and inclusive definition of information, but unfortunately this does not seem possible at the current state of scientific knowledge. Let us examine this problem more fully.

Some aspects of the problem of information have been treated precisely in the mathematical theory of information. (See, for example, refs. 51, 52, and 59.) An examination of the elements of this theory will make clear those aspects of the information problem which it does not encompass. A finite set S of symbol designs, from which all messages to be considered are formed, is assumed given. A message is a finite ordered sequence formed from the symbols of S by selection with replacement, or what we shall simply call selection. The probability that any symbol x_i of S will be selected, and the probability of selecting x_i if the previous n selections of the message are $x_1, x_2, \ldots, x_n$, are assumed given. Consider the set M of all possible messages m_i formed from the symbols of S. For each message m_i there is a probability of occurrence deducible from the probabilities coordinated to the elements of S, and for each m_i of n symbol designs there is also deducible a conditional probability that the next selection will yield a given message of n + 1 symbol designs. All other conditional probabilities in the set M are zero. Thus M forms a Markov chain of finite sequences of symbols; this chain we shall call the "message space." To any message space is assigned a single real number H which is called the "average amount of information per symbol" or "entropy" of the underlying set S. Without going into mathematical details, H is found as follows: Encode each message having n symbols by a well-defined binary number which is dependent on the probabilities in the message space. Divide the number of binary digits in this number by n. Form the sum of all such numbers resulting from messages having n symbols, each one weighted by its probability of occurrence. Call the resulting number H_n; then $H = \lim_{n\to\infty} H_n$.

We observe the important point that the amount of information is determined by a mapping from the message space to the real numbers; this measure does not, therefore, say anything about any particular message, but indicates only certain characteristics which are true for all messages on the average. Looked at another way, information theory describes a technique of mapping any given coding system S into a binary code in which each binary digit is equiprobable and such that, on the average, a minimum number of binary digits is used to code the messages of M. Thus, there is a mapping of the code S into one of a class of normal forms, the members of this class being equivalent in the sense that they require a minimum number of binary digits on the average. Two different codes are compared, in information theory, by considering the average number of binary digits required per symbol in the normal form, that is, by comparing the respective H values.

In real communication systems there is, in general, noise in the transmission of any message. Noise is defined as "any phenomenon which during the transmission transforms the message in an unpredictable manner" (R. M. Fano, ref. 70, p. 693). One would expect any measure of information in the noisy case to have the following properties: (a) When the noise is zero, the value of H characterized above is obtained. (b) An increase in noise results in a decrease in amount of information effectively transmitted. If the statistics of the noise are given, such a modification of the definition of amount of information is possible. It is a single number associated with what may now be called the noisy-message space.

This technique is ideal when we are dealing with a situation in which we are concerned with the statistics of the message space, and are not concerned with any specific message that might be formed. We may term this a measure of information at the syntactic level. It is a method which is suitable for problems of channel capacity, efficiency of coding, and the like, for which the theory was designed.

Two principal characteristics of this theory are: (a) It is concerned solely with the originator or effector of a message (only to the extent of assuming that the statistics of the appearance of certain combinations of symbols are known), and some interval during which neither the effector nor the receiver is involved but during which noise may be introduced. (b) It is not directly concerned with specific messages, but only with the ensemble of all possible messages. These two facts demonstrate that information theory does not treat several properties which are characteristic of information and communication in a human group. The particular messages that are sent, and more important, the specific facts communicated, and the relation of these facts to the behavior of the receiver are the aspects which are of greatest concern to psychology.

Before probing this further, it is appropriate to point out that information theory may be applied in principle to the semantic content of messages. We mention this in order to show clearly that a problem still remains for which mathematical information theory is not adequate, even when its base is shifted from symbol designs to symbol contents. We assumed above that the set S consisted of symbol designs, and therefore, for the most part, semantic content will be conveyed in some combinations of symbol designs only and not in others. Nonetheless, we may form a new set S' of symbols such that each one represents exactly one fact that the effector may send. The elements of S' are in one-to-one correspondence with classes of equisignificant messages formed from the elements of S. In principle, it will be possible to determine the statistics of S'; hence, there will be a semantic message space and assigned to it a semantic amount of information. Of course, this space may be noisy. Such a technique, though sometimes useful, does not solve our problem, since the theory is concerned, not with the specific messages or facts, but only with the set of all possible facts expressible by sequences of elements from S.

We can say, therefore, that there is an area which needs theoretical formulation which, when formulated, might well be called psychological information theory. Such a theory is not known to us, and we have not constructed one; however, we may make a few remarks about the problems it must encompass. Let us make two assumptions (the first is not basic, and the other will be discussed later): (a) The messages have been reduced to a semantic code. (b) Messages identical in design are identical in meaning for both the effector and the receiver. Then, at least three aspects of the message are important as determinants of the receiver's behavior: (a) Does the message convey any new information; that is, does the receiver not already have the content of the message?*

---

*The elements of a theory of information which will cope with problems such as this one, and some, though not all, of the others we are discussing has been presented by Y. Bar-Hillel (50).

(b) What truth value does he assign to the message? For example, if he already has the symbol, then the second receipt of it (from another source) may augment, in his mind, the probability of its truth. He must assign a truth value which is based on auxiliary evidence from contextual material he already has, on his evaluation of the effector, and on his estimation of the noise in the transmission. (c) Is the message relevant to his interests and activities at the time? We may briefly summarize these points by saying that the theory must evaluate the message in relation to the state of the receiver.

In actual practice, the situation is more complicated than we have indicated; it is often very incorrect to assume that a given message (sentence) has the same factual content for both the receiver and the sender. This disparity can be considered to be noise in the encoding-decoding process which must be considered to be distinct from the noise in the transmission. It is important to note that we have termed this noise in the joint process of encoding-decoding. Operationally it is impossible to speak of noise in the encoding process alone, or in the decoding process alone, since no operations which do not consist of both an encoding and a decoding can be designed to determine it. The statistics of noise can be defined only in terms of the intended selection from a set as compared with the actual selection; thus both a sending and a receiving instrumentality are required, the former coding and the latter decoding the message.

Without assumptions which seem likely to prove unwarranted, there appears to be a very serious difficulty in trying to treat this noise by information theory techniques. Let us suppose that the effector wishes to send a message expressing a fact which we may assume he has labeled I_e. Suppose he encodes this into the semantic symbol x which is transmitted without distortion, and this is decoded into a fact labeled by the receiver I_r. We should like to say that the encoding-decoding process is noise-free if $I_e = I_r$. There are two problems here, the definition of equality and the determination that equality as defined actually obtains. Assume, for the moment, equality is defined. Then the method used to determine whether equality exists is to have both the effector and the receiver send I_e and I_r, respectively, to an observer. Such a communication may be an ordinary message, but more often it is an observation of behavior which, for our purposes, can be treated as a message to the observer O. Let us, for simplicity, assume that both the effector and the receiver use the same symbol x in this process, and this is decoded by the observer into the idea he labels I_o. If we know $I_e = I_o$ and $I_r = I_o$, then $I_e = I_r$. This necessitates knowing that two other encoding-decoding processes are noise-free. This is not known. The problem is complicated in notation, but not in principle, if the effector and the receiver do not use the same symbol x. A more serious difficulty is the assumption that equality is defined. In general, such equality is defined through an outside person such as the observer and by the same technique as we just described. Thus, the detection of the noise and the definition of equality are not independent. In practice this is dealt with by requiring highly redundant messages (that is, possibly several referring to the same fact) from the receiver and the effector to the observer; this should reduce the amount of noise, and equality is defined by the observer. It is difficult to make this definition of equality more general and less restricted to the specific observer, and also to be certain that the noise has in fact been reduced to a very small amount.

In our work we shall make use of mathematical information theory where it is more readily applicable; that is, at the syntactical level. In particular we report one experiment in which noise was present (see Chap. V), and this data will be analyzed in the terms of information theory. However, since the noise was in the encoding-decoding process, the analysis contains the logical weakness mentioned above. When we must deal with the relation of the semantic content of a message and its receiver, we shall employ techniques developed ad hoc for the specific circumstance.

2.4. Boundary and Initial Conditions

Parts of the input I_C to the control node in a feedback system may consist of restrictions on the possible outputs of the control node. The parts of the input I_C which contain such restrictions are called the "boundary conditions" on the control node. We include under boundary conditions only restrictions on the output of the control node which are directly imposed; for instance, the task may indirectly have the effect of restricting the output of the control node, but since these restrictions are imposed indirectly, the task is not considered part of the boundary conditions. Boundary conditions are usually imposed to make the outputs of the control node conform to the characteristics of a particular functional environment, or to limit the actual functional environment of the control node. For example, if the functional environment contains an amplifier, through which the output CE of the control node is to pass, then part of the boundary conditions would restrict the output of the control node to electrical signals of a prescribed type, sent along a particular wire to this amplifier.

All portions of the input I_C to the control node, except for the boundary conditions, are called the "initial conditions." The task, for instance, is part of the initial conditions, which may also include a priori information about the functional environment and its transfer function, and other information. In general, information about the functional environment is given the group prior to its attempt to effect changes in the environment, but this need not always be true. The group may be continually receiving more information about the functional environment during the course of its interaction with the environment. Furthermore, there are many examples of groups whose task is continually changing. The use of the words "initial conditions" for these categories of information given the group cannot always be taken to imply that this information is received before the group commences its activities.

We shall say that a functional environment of a control node is "determinate" if it is completely predictable relative to the possible outputs of the group on the basis of the initial conditions. In general, the functional environment is not determinate. For example, if a man is cleaning a dirty room and an unexpected breeze blows through a broken window, he will not have complete control, if any, of where the dust will be.

If, on the other hand, the environment is an amplifier, barring power or component failure, the output is for all practical purposes determinate. Indeterminancy of the functional environment may not be a permanent situation in that, by trial and error or by systematic experimentation, the task-oriented group may learn enough about it to make it, thereafter, predictable. The indeterminancy of the functional environment may be a result of incomplete information or the functional environment may be intrinsically indeterminate. Indeterminancy due to incomplete information about the behavior of the functional environment should not be confused with the incomplete information about the current state of the functional environment resulting from noise in the feedback. This, of course, may hinder the completion of the task as much as an indeterminate environment, but it is inherently different.

If the task completely specifies the required feedback, we shall say the task is "determinate," otherwise indeterminate. It seems clear that, in general, the task will not, or cannot, completely specify the desired feedback. For example, the task of a scientific research team is not determinate; nor is a situation in which the feedback may be divided into several classes, only one of which is specified. Specifically, an electrical device which requires that the average power output be a certain amount, but which does not prescribe the particular waveform, is an example of a case where only part of the feedback is specified.

2.5. The Functional Problem of a Task-Oriented Group

In the most general case of a task-oriented group, the functional environment is neither static nor determinate, and the communication is noisy. The control node has a certain set of possible transfer functions available to it, and must select a transfer function from this set which at any instant will best fulfill the imposed boundary and initial conditions. This selection process will usually result in a time sequence of such choices, with the result that the control node should come closer and closer to fulfilling the task as the sequence of choices proceeds in time. In some cases the task itself will also be changing with time, and this further increases the complexity of the problem.

One feature of this process, which task-oriented groups have in common with all feedback systems, is the dependence of the process on the time delays within the system, and the time constants of the nodes. In any interactive two-element system, the time delays may be represented as in Fig. I. 2. The actual sources of such time delays τ may be internal to the nodes C and E, but they may still be represented as part of the communication process. Each node N will also have a time constant τ_N which is the time between the occurrence of the inputs and the occurrence of the corresponding outputs. If the functional environment in a feedback system is responding to other inputs in

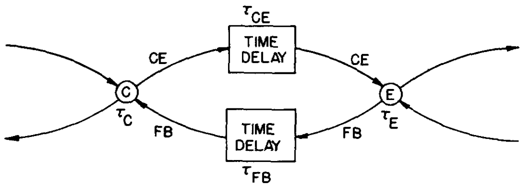


Fig. I. 2

addition to CE, the various time constants cannot be independently chosen and still permit the control node to comply with the task. For example, if τ_E is small, and τ_{CE} and τ_{FB} large enough, no value of τ_C will permit the control node to control FB effectively. This phenomenon is discussed at length in standard works on servomechanism theory (120, 122), and leads to the establishment of various stability criteria for feedback systems. These considerations cannot be overlooked in dealing with sequences in time of such choices of the transfer functions by the control node as are mentioned above.

We return now to the specific problems of the task-oriented group. If an element of indeterminancy remains in the task and in the knowledge of the functional environment, or if there is noise in the feedback, then no single solution can be the optimum behavior of the control group. Rather, a probability distribution of solutions will describe optimum behavior.

The important notions are: (a) The problem of the control group is the repeated selection, from the class of all transfer functions available to the group, of a set of transfer functions which will most nearly fulfill the task, and the organization of action to obtain the corresponding performance. (b) In general, the words "most nearly fulfill" mean an optimization which yields a statistical distribution over all possible activities rather than a unique selection.

2.6. Experimental Limitations

For our experimental studies of task-oriented groups, we have generally restricted ourselves to a very simple set of conditions: a determinate functional environment, a determinate and static task, simple and rigid boundary conditions, and initial conditions which completely specify the transfer function of the functional environment. The tasks used have usually been sufficiently simple to permit successful completion in a few minutes, and the information used in the solution of the task has often been coded in some simple fashion to facilitate subsequent data analysis.

We have in almost all cases tried to produce the simplest possible set of conditions in the laboratory which would permit a task-oriented group to be studied; our emphasis has been on the investigation, under these simple conditions, of the methods used by the task-oriented group to select a suitable transfer function. Variations in experimental conditions have been undertaken to shed light on various facets of this central problem, and the investigation of more complex situations has been deferred to a later date.

This experimental procedure was adopted for two reasons. The first and most obvious was purely practical: we felt that under these simple conditions we could, by using the facilities at hand, successfully set up and complete numbers of experiments and produce data which were statistically coherent and readily analyzable. Second, it is feasible to attack more complex and realistic situations on the experimental level only after initial experiments are carried out under very simple conditions. An understanding of the results obtained from these simple situations will enable us both to attack the theoretical problems of more complex situations and to design experiments which will successfully investigate them.

3. The Internal Structure of the Group

3.1. Introduction

In the previous section the task-oriented group was treated as a black box, and various concepts were defined. It is possible to continue to theorize and to experiment on this level, without attempting to "open the box" and investigate its inner structure. Unfortunately, although this technique seems the most direct, we have been unable to arrive at a set of general rules or laws which will govern the performance of task-oriented groups on this basis. Perhaps such a solution will be possible in the future. It took many years of development before electrical engineering was able to operate on this level of analysis. At present, however, it appears that any study of task-oriented groups made on this level would require an exhaustive series of experiments dealing with each possible type of group. For practical reasons, this is beyond us. Consequently, we have found it desirable to attempt to penetrate the inner structure of the group.

The second level of analysis, using the systems-analysis approach, is to treat the object of study as consisting of several interrelated black boxes or components. The concept of interrelation is essential except when a trivial reduction of the original object of study into several independent and simpler objects exists. For a system having more than one component, the theory has a new aspect: as before, it must include the transfer function of the black boxes with which we are working, but it must also formulate the relations existing among these components. It is the emphasis on the interrelationships between black boxes that distinguishes the systems approach from strict behaviorism.

The choice of components to study is arbitrary, although it is always motivated by some undefined but generally acceptable requirement of "naturalness," and by the pragmatic condition that it must answer more readily the questions asked of the theory. A further breakdown of the original object of study is often helpful because the possible transfer functions for the components are more restricted than those of the whole system and they have more peaked statistical distributions.

3.2 Components of a Task-Oriented Group

When an attempt is made to study the inner structure of a task-oriented group, the choice of the components into which the group is to be decomposed is not trivial. Since by definition a task-oriented group is a set of people and machines, the most obvious decomposition is to treat these people and machines, respectively, as the new black boxes. In some cases, smaller groups of people or machines could also be taken as single components. The choice of this breakdown of a task-oriented group can be supported by the fundamental nature of such a group.

A task-oriented group is formed for one reason only: to create by appropriate interaction a transfer function which is appreciably different from that of any single person. The important feature from a social point of view is that the organization is able to cope with problems which the individual, or a set of independent individuals, is unable to handle. Consider, for example, the air-defense problem. Air defense, quite obviously, cannot be carried out by any single individual, or a set of independent ones, for the possibilities in their transfer functions are too limited. In fact a group of people is not sufficient, for there is noise (clouds) in visual transmissions from the target to the individual. So equipment (radar, and the like) is employed along with communicating people to obtain a less noisy feedback and an appropriate transfer function having an output which will ultimately destroy their targets. The use of a group, rather than an individual, is inherent in this problem. It would seem, from an abstract point of view, that an organization pieces together several transfer functions to obtain a resulting transfer function appreciably different from any of the individual ones. If this is true, a logical way in which to attempt a penetration of the structure of the group is to consider the group as composed of a collection of smaller black boxes, or nodes, whose individual transfer functions add or combine in some way to make the over-all transfer function of the task-oriented group. In general, these nodes will be the individuals comprising the group, but, as mentioned above, some of these nodes may be machines or smaller groups of people if this seems appropriate to the analysis at hand.

If this breakdown is to be useful, it must result in theoretical and practical problems which are easier to handle than the original ones. Is it easier to determine the transfer functions of individuals than the transfer functions of groups? We cannot answer this affirmatively at present, but we observed above that the primary reason for using a group is to make possible a transfer function not available to an individual. This suggests that the transfer function of the individual does not have as great a range as that of the group, and may therefore be easier to determine. Nevertheless, an extremely complex problem remains, which probably can be attacked only by using a number of drastic simplifying assumptions in both experimental and theoretical work. The basic theoretical effort of such research must be the development of such assumptions and the conclusions that follow from them.

3.3. Psychological Factors

In the development of further analysis of the task-oriented group, and in the theoretical and experimental treatment of such groups, we have generally made the assumption that the individuals in the group are statistically alike. That is, we regard them as having the same probability distribution for any given type of behavior, thus effectively ignoring the differences which exist between individuals, except as they contribute to the calculated probability distribution for the "average" man. This assumption has been criticized by those who say that the psychological differences between individuals are too fundamental to be treated in this manner, even for the purposes developed in this study. The critics have emphasized that people have idiosyncrasies; they strive towards the most varied goals propelled by drives whose definition is a subject of constant dispute among psychologists; they differ in altogether too many respects. Can we propose to study groups of people without taking these factors into account?

The ultimate answer to such criticism is whether work in the direction we are outlining leads to a body of theory which explains the observed data and predicts new results. It is clearly impossible to state what the most profitable research emphasis is until one knows the relative importance of differences among group members and of group structure. We know only that both factors affect group behavior.

Our decision has been to examine the effects of group structure. Specifically, we have sought to study experimental problems in which the structural factors seem to be dominant and which do not tend to stress individual differences such as intelligence and emotional reaction. Since it would be quite impossible to expect complete success in the reduction of individual differences, the residue of variance has been considered simply as a statistical distribution.

Several experimental techniques have been used to insure that individual variability is not a major factor in the work done. Consequently, our results cannot be expected to yield much light on problems in which the effects of individual variability are dominant. Although these techniques will be examined more fully in Chapter II, it is appropriate to mention them briefly here. First, the tasks were sufficiently simple so that anyone of average intelligence had no difficulty in understanding them. Second, the tasks required only a modicum of dexterity and ingenuity. Third, the situations created in the task situations have been of such a nature that they would not evoke strong emotional responses from the subjects. Fourth, the motivation given the group was designed to produce similar responses from all members of the group, and to remain at a fairly constant level throughout the series of trials; at the same time the motivation was not sufficient to cause strong emotional involvement on the part of the subjects.

It is to be hoped that psychologists interested in the area of explaining and developing theories to account for the variability of human behavior will probe more deeply the problems of an individual in a group situation. Such theories, were they formulated in terms of the dependence of the transfer function of an individual on various psychological variables, would be of inestimable value to a systems analysis of group behavior. Until this state of sophistication is reached, it may not be inappropriate to divide the study into two stages: one concerns itself with the effects of group structure when the individual variance is small, and the other concerns itself with an explication of the individual variance. Assuming this so, we have restricted our efforts to the former problem at present.

3.4. Communication Network and Structure

One large category of restrictions imposed on task-oriented groups has been mentioned previously, that is, those on the outputs of the control group, the boundary conditions. One important way to effect these restrictions, in practice, is to impose boundary conditions on the transfer functions of the individuals composing the group, since it is often easier to specify such restrictions than to restrict the whole group. In addition, the group itself, in the course of completing one or more tasks, may induce restrictions on its own members. This effect is not part of the boundary conditions of the group, but it is very similar in its results to a boundary condition which restricts the individual transfer functions.

Of this class of restrictions we have isolated those that deal with the source and destination of communication. There are two classes of such restrictions: (a) those imposed as part of the boundary conditions, and (b) those evolving within the group. In the first case, that part of the boundary conditions which determines who may not communicate to whom is termed the "communication network," while the "communication structure" of the group is the actual set of such restrictions that are effective during a particular phase of the group's activity; that is, who does not communciate to whom. To give a more positive definition, the communication structure is that part of the communication network which is actually employed.

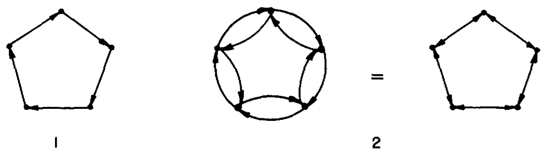


Fig. I. 3

One-way links are drawn as in 1. Two-way links are simplified as in 2.

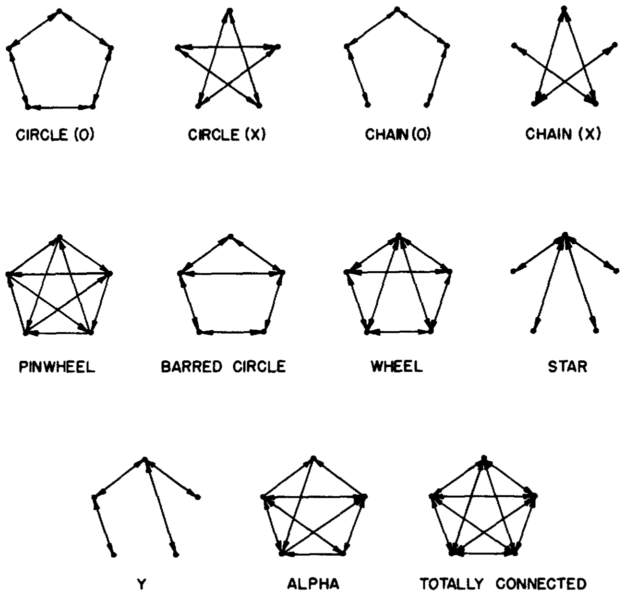


Fig. I. 4

Communication networks and structures will be represented in this paper by diagrams of the form shown in Fig. I. 3. A line with an arrowhead from node a to node b indicates, in the case of a communication network, that communication may occur from a to b. In the case of a communication structure it indicates that communication has occurred. The lack of such an arrow indicates that no message may occur and no message has occurred, respectively. In Fig. I. 4 we present diagrams and names of all networks studied experimentally.

For some purposes it is convenient not to use such a restrictive definition but to treat the case of no channel present as if it were a channel so noisy that a message cannot pass, and to treat the case of a channel in which every message can pass unchanged as noise-free. Obviously, there may be channels which have varying amounts of noise and hence have varying effects on the percentage of a message which will be correctly passed. The usual measures of information theory (59) are adequate to characterize such communication links. Since all communication links have been noise-free in our experiments, it will be necessary to consider communication network and structure only in the sense that we have first defined them.

3.5. Constraints and Differentiation

Besides the communication network, there are other boundary conditions that restrict the class of transfer functions that some or all of the individuals employ. Any such restrictions, except for those imposing the communication network, are called "constraints." For example, a man not in a command position in a military situation may not employ the transfer functions which will result in an order. If all the individuals in the group have the same restrictions externally imposed on them, or if a group has no constraints imposed, the group will be considered "homogeneously constrained." For the most part the groups we have studied have been homogeneously constrained. Imposed constraints are aimed at, and do result in, simplifying the group's problem of selecting the proper transfer functions for the task. In many actual situations the existent constraints do not properly serve their purpose, since the task has changed while the imposed restrictions have not.

The set of limitations on the individual transfer functions which naturally arise in the solution of the problem, but which are not imposed by boundary conditions, are called "differentiation." Differentiation will arise under a wide variety of circumstances; for example, as a function of the intelligence and ability of the people in the group, the location or distribution of information about the functional environment within the group, the task, the functional environment, the initial information, and the like. This important problem has been only briefly explored in our experimental work.

3.6. Group Size

The size of the group makes a difference not only in the class of problems it may handle, but also the way in which it handles them. Our discussion has made no mention of this, although it is apparently true that the nature of an individual's transfer function is dependent upon the size of the group he is in. Obviously a given set of transfer functions may be combined in many more ways in a large group than in a small one. The relationship between these two factors, and the nature of the change in transfer functions as the size of the group increases is not known. The work of this laboratory sheds light on this only in a restricted and incidental way (see sec. III. 10), for we have restricted our studies almost entirely to groups of five people (see sec. II. 1. 3).

4. Problems of Analysis

When the individuals of a group are treated as the nodes, there are three aspects to the problem of analysis: first, the statistical determination of the transfer function of the node; second, the statistical determination of the group transfer function; and third, a method for combining the transfer functions of the nodes to obtain the transfer function of the group.

These aspects of the analysis are certainly not independent in fact, but to some extent they may be considered as if they were. The first two problems are essentially psychological and are primarily empirical. The third problem is essentially mathematical. They are interrelated since the mathematical formulation, which is based on experience and intuition, includes parameters which must be measured. These measurements being made, a prediction of the group transfer function should follow from the mathematics which, if it does not accord with observation, will necessitate a reformulation of the mathematical problem.

In their full generality these problems must be of enormous complexity since the nodal transfer function will vary with individual psychological variables which themselves will be dependent on the on-going group process. However, as we mentioned in section I. 3. 3, we are treating the nodes as statistically average individuals and so some of the complexity is eliminated. We should then look for statements of transfer functions which are dependent on inputs to the nodes and on time, but not explicitly dependent on the individuals. In addition, we should expect to find results about the group transfer function, and we should attempt to relate the individual statistical results to the group statistical results.

The results reported in Chapters III through VI fulfill these expectations to varying degrees. Transfer functions for both the individual and the group are presented. The very methods of distilling these functions from the raw data embody the assumption that individuals are statistically identical. Interwoven in this empirical data are both mathematical formulations which attempt to relate the individual to the group result, and deductions which relate certain aspects of the individual transfer functions to more basic considerations. At the present time many mathematical difficulties have not been overcome, making it impossible to answer definitely the question of whether the approximations to the individual transfer functions are sufficiently good to predict the group results. However, the direction of effort to be made is more than clear.

Other mathematical programs related to the problem of building the group from its nodes, but which do not have direct application to our experiments, are presented as appendixes. The one more intimately concerned with the above problem is Appendix 3 which presents a model for group behavior which is formally identical to linear electrical network theory. It has the serious drawback that it does not have a close relationship to most actual situations, including our experimental ones. Essentially, two assumptions are made which invalidate it for group applications, but which are essential for a mathematical solution in the conventional manner. They are discussed in detail in that appendix, and it will suffice to say here that we assume there is but one particle of information, replicated many times, moving in the system, and the response of the node is linear. The model is valuable in that it indicates how such a model may be built, points up the role of the transfer function, and suggests the type of question that may be asked of such a model.

More distantly related to the dynamic group problem is the study sketched in Appendix 4 which is concerned with the abstract properties of the network system itself. This, depending on how one looks at it, is a study of the topological or algebraic aspects of a dynamic problem. Its importance lies in the notion that the psychological reaction of people in a network may be as much influenced by variables best described in topological terms as by the dynamics of the communication process. This, however, remains only a pious hope for there cannot be said to be any really substantial work supporting this view.

In addition to the type of analysis problems we have discussed, there is a class of synthesis problems which is among the most important of the applied problems. No attempt will be made to formulate any of these problems precisely, for we have done no work in this area; however, two loosely phrased examples may give an idea of the type of problem we have in mind. Given a communication network for a group, a desired group transfer function, and a set of available individual transfer functions, determine the location of the individuals in the network to give the desired transfer function, if possible. This problem is made markedly more difficult in reality, for the individuals are liable to pass through a transient learning period which changes their transfer functions (see Chaps. III and IV). A second example is to fix the individual transfer functions and find the communication network to achieve the desired group transfer function. This is subject to the same discussion of learning as the first example.

5. Summary

The purpose of this chapter has been to introduce to the reader the concept of a task-oriented group. The discussion began with the "undefined" notion of a node having inputs and outputs, the relation between them being called the transfer function. From two such nodes we formed the most general interactive two-element system, which is called a feedback system if one of the nodes has an input which imposes a requirement on the feedback from the other node. In the feedback system, one node is differentiated from the other by the introduction of the task; the node into which the task is introduced is called the control node, and the other the functional environment. When a group of people is the control node of a feedback system, it is called a task-oriented group. The input to the control node is classified into the exclusive categories of boundary conditions and initial conditions. Initial conditions which provide both a determinate environment and a determinate task are the only kind used in the work reported here.

The second major section of the chapter was devoted to a discussion of the task-oriented group and its internal structure. It was pointed out that the boundary conditions are often imposed in the form of boundary conditions on the individuals within the group. One set of restrictions is the communication network and another, which includes the hierarchical structure of the group, consists of the constraints on the individual transfer functions. In addition, certain modes of behavior arising from within the group were classed as the communication structure and as the differentiation of the nodes.

CHAPTER II — EXPERIMENTAL TECHNIQUES

1. Introduction

1.1. A Typical Experiment

Our experimental program is to investigate the concepts set forth in the previous chapter and their interrelations. Because complex group situations are so enormously difficult to analyze, we have conducted our experiments subject to the restrictions of section I.2.6 which were designed to so simplify the situation as to make it amenable to analysis. It is our conviction that the study of restricted situations will provide us with the opportunity to develop techniques which can be appropriately applied to more complex situations. We shall, in this chapter, discuss the technical means by which the desired restrictions were imposed. Since these means were quite similar in all the experiments to be reported, our discussion will apply to the whole program rather than to each experiment individually. The detailed conditions of each experiment are described in Appendix 1.

Let us first describe briefly and informally the conduct of a typical experimental session. The subjects came to the experiment with little or no knowledge of what they were to do except that they knew approximately the time it would take. They were given the necessary instructions by the experimenter. This briefing included explanations of the aspects of the apparatus which had bearing on the behaviors permitted the subjects, the task to be performed, the means for communicating within the group, the initial information they would be given on each trial, the way they were to signal the task solution when reached, and the rules for handling the materials which would constitute the record of the experiment. The description of the task included telling the subjects what would constitute a successful performance so that, insofar as they were cooperative, they would be motivated to reach the intended goal and, consequently, be rewarded whenever they reached that goal.

The experiments consisted of a fixed number of trials (between fifteen and thirty), and each trial consisted of the solution of a simple problem. In a typical trial problem each member had a box containing five colored marbles, with only one color common to all the boxes. The task was to identify the common color by sending and receiving written messages, and to report this color to the experimenter. The subjects sat at a round table which was designed to enforce the various restrictions of the experiment: a communication network, visual separation of the subjects, and the like. Under the surveillance of the experimenter, who enforced the rules stated during the briefing period, the subjects began the experiment.

Three kinds of data were collected: (a) the written messages which, besides their content, provided identification of author and receiver, and the sequence in which they were sent; (b) errors in the solution recorded by the experimenter; (c) recorded times of group response.

At the end of the sequence of trials each subject was given a questionnaire designed to determine his motivation during the experiment, his concept of the communication network, and whom he believed to be the leader in the group.

1.2. Subjects

We have drawn our subjects from the following groups of people: M.I.T. undergraduates, both volunteer and paid; Harvard and Radcliffe undergraduates, volunteer; enlisted Naval personnel, Receiving Station, First Naval District, Boston; and enlisted Army personnel, Fort Devens, Ayer, Massachusetts. Table II.1 indicates which categories of subjects were used in specific experiments.

There is a question whether results from samples drawn from one of these populations are comparable to results from a different population. In fact, it is generally accepted that intelligence level and cultural background of subjects can have a marked influence on the observed behavior. Thus, for example, in a task demanding a certain type of ingenuity, the M.I.T. subjects could be expected to make a better showing than the military subjects. In the experiments reported very little ingenuity was required, and it appears that the military subjects are just as satisfactory as M.I.T. subjects and, since they are more representative of the U.S. population, more appropriate subjects than are college students.

It has been found in two similar experiments, 1 and 4, that the times necessary to finish an experimental trial differed. In part this was due to somewhat different experimental situations, and it was possibly due in part to the different categories of subjects used: M.I.T. in the former and military in the latter. However, the rank ordering of these times as a function of the different communication networks were the same in the two experiments. It seems reasonable that when the populations from which the subjects are drawn are not very different, the magnitudes of structurally determined measures vary, but their order does not.

Equally well, it is not unreasonable to assume that there are categories of subjects so different that differences in kind as well as in degree will result. There are probably subjects from radically different cultures who would reject the entire experimental situation, and possibly subjects who would accept the situation, but who would yield orderings in certain measures different from those given by the groups we have studied. It remains a question whether such reversal effects can be found between two subject populations, both included in the class of "Citizens of the United States," who are above some minimum educational level. Until more is known about the dependence of individual transfer functions on various psychological variables in the individual, it will be difficult to specify with assurance the population for which our results are valid.

1.3. Experimental Conditions

Our experiments have all been run in several rooms set aside for this purpose in the laboratory. These rooms are fitted with the necessary equipment for the experiments,

Table II. 1

Summary of Experiments Reported

All Experiments In This Table Were Performed Using
Written Messages, Determinate Task and Environment, Homogeneous Constraints

Expt. No.	Name	Year Run	Run By	Subject Population	No. of Groups	No. of Networks	Total No. of Subjects	Message Content Allowed	Action Timing	Coding Noise	Task
1	Common Symbol	48	H. J. Leavitt	M.I.T. Volunteer	20	4	100	Free	Free	No	Finding Common Symbol
2	Common Symbol	49	S.L. Smith	Radcliffe Volunteer	8	2	40	Free	Free	No	Finding Common Symbol
3	Noisy Marble	50	S.L. Smith	M.I.T. Volunteer	12	3	60	Free	Free	Yes	Finding Common Marble
4	Quantized Number	51	G. N. L.	Army and Navy	90	7	450	Coded Restricted	Action Quantized	No	Collecting Five Numbers
5	Quantized Marble	51	G. N. L.	M.I.T. Paid	19	3	95	Free	Action Quantized	No	Finding Common Marble
6	Noisy Marble	52	G. N. L.	Army and Navy	20	4	100	Free	Free	Yes	Finding Common Marble

and the subjects are in them only for the duration of the actual experiment. We normally asked subjects to perform a series of trials under set experimental conditions. Such a series usually lasted about one and one-half hours, and was run continuously.

In all cases reported, we used a different set of subjects for each experimental run. This was necessary since one of the primary problems we have been investigating is the learning that occurs from the time a subject is first introduced into a network until the end of a certain number of trials. It is clear, then, that if the subject is run through the experiment again in a different network, he will carry to this experiment information and ideas that the naive subject would not. (It may be interesting to study the effects of such conditioning, but we have not been pursuing this problem.) On the other hand, since it is usually easier to obtain one set of subjects for a long period of time than it is numerous sets, each for a short duration, it would be desirable to have a method which would permit their use in a number of different experiments. As yet no such method has been devised.

All the experiments to date have been designed for groups of two to six subjects; the work reported is based on five-man groups. This number was picked for purely practical reasons. For any larger number the possible actions within a given network are so great as to require an inordinately large number of experiments. At the other extreme, the possible networks on two- and three-man groups and the possible experimental actions within these networks are so few as to result in experimental distributions on only, say, two values of a variable. It is quite likely that in the future we may use four-man groups, and even three-man groups, to check hypotheses we have obtained from the larger groups, for as the Heise-Miller experiment (29) shows, certain types of design allow results from even the simple three-man groups. Ultimately, work will have to be carried out on a wider range of group size to see which results are independent of size and which are not.

2. Initial Conditions

2.1. Task

Throughout these studies the task has been determinate, and has been chosen to obtain a simple well-specified feedback from the determinate environment. In Experiment 4 the required feedback was a bell ring; in Experiments 1, 2, 3, 5, and 6, a response from the experimenter that the required conditions have been met.

The actions required by the task have been extremely simple and straightforward: Each member of the group was given an item of information, and the task of the group was to collect all these items in one or more places and perform some simple operation on them. In several experiments each member of the group was given a small number of different types of objects, only one type being given to all; the operation required was merely to determine which type was held in common. In general, tasks were chosen to be as simple as possible in their demands on intellectual powers, and at the same time to require as much interaction within the group as possible or as was thought to be analyzable. Our aim in every case was to devise the task so that the intelligence or speed of reasoning of any individual in the group would not be a limiting factor in the performance of the group. A general feature of all the experimental tasks has been that an individual, substituted for the group, would have found the task trivial.

2.2. Functional Environment

The functional environment of all the experimental groups was determinate. It consisted of a simple electrical or mechanical apparatus whose properties were explained to the subjects by the experimenter, and sometimes included the experimenter himself. In the latter case he completely described what his pertinent behavior would be with respect to outputs from the group so that the functional environment was still completely determinate. One of the primary roles of the pre-experimental briefings was to make clear to the subjects the determinacy of the functional environment and the nature of this determinacy.

The communication between the group and its functional environment depended on the nature of the task and the details of the apparatus, but in general it has taken the form of signals from the individuals of the group to some memory or data-recording device which gave a feedback to the group either that the task had been completed, or stated the deviation from correct task completion. Communication from the group to the functional environment entailed the use of switch boxes in each of the sections of the table at which the subjects were seated (see sec. II.3.3), or mechanical devices, such as marbles and tubes through which the marbles could be passed.

In addition to the determinate mechanical aspects of the environment, during the course of the experiment the subjects were given some informational inputs for each trial of the experiment; these, too, are part of the environment. These inputs were the source of the problem which the group was to solve; for example, sets of marbles similar to those mentioned in section II.1.1 were used in Experiments 3, 5, and 6; in Experiments 1, 2, and 4, the inputs were a set of symbols, such as numbers.

These simple and determinate functional environments are certainly open to criticism; principally since few real task-oriented groups have a determinate functional environment. Again the defense is based on the hope that insights derived from simple situations may be applicable to more complex situations which at present are too subtle and complicated to unravel. Certainly, a possible next stage of development in this direction is the introduction of simple indeterminate functional environments. The results on the development of group leadership from one pilot experiment in which this was done indicate that this is a fascinating and fertile area.

An important problem in the realm of initial conditions is the difference, if any, of the groups' methods of solution when the task and the functional environment have the same logical structure but different material implementation.

3. Boundary Conditions

3. 1. Constraints

The experimental groups reported on in this study were homogeneously constrained but were far from free of restrictions. That is to say, the transfer functions of the individuals were all externally restricted in exactly the same way. In none of the experiments were the individual transfer functions completely free of all restrictions; in fact, it is unlikely that one can ever find a task-oriented group in which no imposed restrictions occur. For the experimental work, however, these restrictions have been more stringent than they would be in most real situations.

3.2. Methods of Communication

In all of the experiments reported in this study, communication within the group has been restricted to written messages, and communication between the members of the group and the functional environment has been by means of simple electrical or mechanical signals. There are several reasons for this selection of restrictions on communication. The recording of experimental data in such a fashion that the process of communication in the group can be reconstructed is greatly facilitated when the messages are written. The enforcement of the various experimental restrictions, which are discussed below, is very simple in the case of written messages. Further, it was felt that the introduction of uncontrolled psychological variables would be at a minimum in the case of written messages. For example, if we had allowed oral communication over telephone circuits, vocal characteristics and intonations might have had a serious and unknown influence on the results. Granted that in real situations such characteristics are of great importance, it was felt that such an uncontrolled factor should be eliminated for the present.

In some of the experiments, particularly the earlier ones, the written communication of the group members was not further restricted. They were allowed to discuss the problem at hand, the method of solution, the differentiation (organization) of the group, or anything else they wished to. The data thus obtained is fascinating and is discussed at some length in Chapters V and VI, but it is exceedingly difficult to analyze. This difficulty may be inherent in data of this type, or it may be that we do not know what to look for. In the hope that the latter is the case, we then ran an experiment (No. 4) in which the allowable written messages were restricted to input information. The results of this restricted experiment give some insight into the more complex data, but by no means complete clarification, since all messages about group differentiation were eliminated.

The use of written messages has serious drawbacks, and because of these we cannot know whether our results generalize to groups with different types of communication. Written messages often make the time for the completion of the simplest task inordinately long, so that any reasonable sequence of runs on an experimental group tends to approach the interest and/or fatigue limit for the members of the group. Also, it is doubtful if the same type of emotional reactions are elicited by written communication as by oral. In addition, the use of written messages tends to give a very artificial aspect to the situation which probably tends to reduce the motivation of the group (motivation is discussed more fully in sec. II.4.2).

Experiments using other methods of communication, such as banks of switches and telephone lines, are projected and the apparatus nearly completed (see App. 2 and sec. II.3.3). These new methods should shed considerable light on the influence of the method of communication on group performance.

3.3. Communication Network

In all the experiments, strong and inflexible restrictions were placed on who could communicate with whom; that is, there were imposed networks. This has been considered to be one of the primary variables of our studies; for it was noted early in the work that slight changes in the communication network made appreciable differences in the way in which the group coped with a task, and it is well known that in many business and military organizations such networks do in fact exist. In the earlier studies, in which the group was generally treated as a black box, these effects were noted, and in the later studies we have attempted some explanations of the differences. The value of such explanations is the possibility of their conceptual, if not quantitative, generalization to situations quite removed from the specific ones studied.

The apparatus to impose the network was based on a large round table around which five subjects were seated. Five radial wooden partitions effectively separated the subjects, and they were prevented from circumventing the imposed restrictions by the presence of the experimenter. At the center of the table a block of wood appropriately slotted for the transmission of message cards allowed experimental control over the communication network. The photograph of the table (see frontispiece) will give a clearer understanding than many words.

This apparatus is crude but reasonably flexible. An attempt to replace it by a more sophisticated electrical apparatus has been one of the major operations of the laboratory in the past year. This new device, nicknamed "Octopus," consists of five stations, each having two banks of switches all connected to a central control station and recording apparatus. At the stations, messages of the form "from man A to man B that man C has a 1" may be sent by the subject, and their receipt acknowledged. Each message contains information regarding the sender, the destination, the person referred to, and the item this person possesses, which is either a 0 or a 1. In any given experimental situation we may interpret 0 and 1 as we will. The control section of Octopus allows any of the twenty possible links between ordered pairs of people to be open or closed, and may be readily so modified as to permit a link to be open with probability p. It also controls the input of "initial information" to each station, either a 0 or a 1. For a more complete discussion of this apparatus and its recording device, see Appendix 2.

The advantages of Octopus are that it permits much more rapid communication than can occur with written messages but preserves the simplicity of coded written messages, that it provides a simple and complete record of the data, that it permits simple and flexible control of the communication network, and that it permits the introduction of a measurable amount of noise into any communication link. For ease and speed of experimentation, it sacrifices the flexibility of possible problems, inputs, and messages. It remains to be seen if the gain offsets the loss. We envisage the possibility that Octopus may be used to greatest advantage in conjunction with other means of communication.

3.4. Action Quantization

In Experiments 4 and 5 the subjects were required to act simultaneously in the following sense: After the trial began, each subject prepared his first message, decided where he would send it, and then pressed a button to signal that he was ready to send. Each button activated one of five relays wired in series; hence, when all five subjects had signaled that they were ready, the circuit was closed. This in turn activated a two-second delay relay which, when it closed, activated a buzzer. Each subject sent his message at the sound of the buzzer. The second set of messages was prepared and sent following the same routine, and so on, until the trial was completed. From the beginning of the trial through the first sending of messages is called the "first act," from then to the second sending of messages is called the "second act," and so forth. This set of boundary conditions on the group is called "action quantization."

The purpose of such action quantization is to cause the times for decision and the actual decisions as to where to send a message to be independent, although generally they are not. It causes a marked reduction in complexity of the entire communication process, which results in an experiment amenable to both mathematical analysis and to obtaining adequate data in a relatively short time. It is clear that this is an artificial situation rarely, if ever, encountered outside the laboratory; nonetheless, we feel that it has more than served its purpose in allowing us to obtain approximations to certain transfer functions which would be much more difficult to obtain in the unquantized case and, in addition, it has suggested methods of analysis that are to varying degrees suitable to the more general case. Whether we are justified in this belief is best evaluated after reading Chapters III and IV.

Thus, the experiments we have run fall in two categories: (a) the experiments with a continuous time scale, that is, with message sending times completely unconstrained; and (b) the experiments with the group action quantized in the time dimension. These two schemes do not exhaust the possibilities, and since the jump from the group action-quantized case to the continuous case is so great, it seems wise to consider the possibility of a case which will fall between these two cases. We propose in future experimentation to quantize the individual action times but drop the requirement that every member of the group send on signal. If the signals for "permission to send" are spaced appropriately, we will have the condition which was so desirable in the former quantized case: the experimenter can know what information the subject had before him when he was preparing a given message. However, the subjects will not be required to send messages at times when it is obvious there is nothing to be sent that will help the group to reach its goal. Situations of this sort do arise and are particularly annoying to the end men in the star, for example. It is for just this reason that in Experiment 4 no action-quantized experiments were run with the star pattern. This scheme might be thought of as a speeded-up analog of communication by mail. The postman comes at regular intervals and hence sending and receiving occur on a time-quantized scale, but letters are only sent at those of the possible times when there is reason to send a letter. We believe that this scheme is an important step in the direction of face validity which retains, so far as analysis goes, most of the features of action quantization.

4. Relationship of Subjects and Experimenter

4.1. Role of the Experimenter

Several times in the above discussion we have mentioned the roles that the experimenter plays during an experimental run. Sometimes he is part of the functional environment, but at all times he is a source of initial information and instructions about the experimental situation. We have always attempted to standardize these instructions and roles, but this has not been completely successful. For example, each group will ask questions about the apparatus and procedure, which must be answered to be certain that the functional environment is in fact determinate to all the members of the group. These answers, of course, will put the experimenter in a slightly different relation to that group than to one which asked different questions. Furthermore, the experimenter serves as a policeman, imposing any rules of the game which are not built into the equipment; when the policeman is forced to act, his emotional relation to the person acted upon is probably changed. The control over this variable has been only partial, since these actions on the part of the experimenter were necessary to complete the experiment smoothly and successfully.

We have hoped that the nature of the design would, to a large extent, eliminate the effects of this variable. However, we cannot overemphasize the importance of the relationship between experimenter and group, for we have watched various experimenters at work with groups, and the difference in success and cooperation is marked. It may be desirable to develop experimental procedures which further reduce the role of an experimenter, e. g. by developing equipment which is its own policeman.

4.2. Motivation

Another important way in which the experimenter influences the subjects is the introduction of motivation to carry out the task as well as possible. The tasks are simple and artificial; they do not relate very closely to the subject's daily existence or to his more important goals. On the other hand, the subject is ego-involved to some extent;

he is in a group and under the surveillance of the experimenter and so has some incentive to do well. Further, during the instruction period the experimenter attempts to induce motivation as he explains the functional environment and the task. Several appeals have been used for this purpose: competition with the various other experimental groups who have performed the same task, the importance of science, the relation of these problems to the very important problems confronting industry and the military, and personal challenge. For the most part one concludes that the motivation has been successful or not, only by intuition and observation. We observed that the subjects are reasonably attentive to the problems, they evince great interest in the work and its meaning following the completion of the experiment, and almost no complaints have been received. In addition, questionnaires designed to obtain rough measures of motivation during the experiment were filled out at its completion by the subjects and an examination indicated reasonable motivation.

Nonetheless, we do not feel that we have been completely successful in this direction. One continually has the feeling that experiments having a less artificial aspect would introduce new degrees of motivation which might possibly affect the results. Unfortunately, it is difficult to reconcile appreciable restrictions and high face-validity. This suggests that an appropriate experimental program might include both experiments having high face-validity which will be used to suggest new ideas and directions of study and, on the other hand, very restricted experiments which will attempt to bring forth explanations of the more complex phenomena.

5. Data Measurement

The primary source of data was the cards on which the subjects wrote their messages. These were identified as to source by their color and as to receiver by labeling after the experiment. The subjects were required, at the end of each trial, to group all the cards received and label the package with the trial number. In Experiments 4 and 5 serial numbers on the cards indicated on which act of the trial that card had been sent.

In Experiments 1, 2, 3, 5, and 6 the subjects signalled an answer to the experimenter who recorded it on paper. In addition, in the first three experiments the experimenter recorded the time for the trial. In Experiment 4 the answer was signalled electrically and recorded on an Esterline-Angus pen recorder. The times per act were recorded in both Experiments 4 and 5; however, we failed to record the times at which the subjects signalled they were ready to send a message, and this, as we shall see in Chapter IV, was a mistake.

Following the completion of the trials the subjects were given questionnaires to fill out; samples of these are contained in Appendix 1. The purpose of these questionnaires was to discover what the subjects knew about the network they were working in, to determine what, if any, organizational structure had developed, i.e. what differentiation had occurred, and to ascertain the morale of the subjects. The design of this questionnaire is only partially satisfactory. The most glaring flaw is the difficulty of scaling the results. A new questionnaire is now in service which, it is hoped, will give more satisfactory results.

CHAPTER III – LEARNING

1. Introduction

In all of the experiments listed in Table II.1 learning has occurred, but in three experiments (1, 2, and 4) it has been of primary importance and has been examined in some detail. The first two experiments were of the same structure; the subjects were free to send messages at any time and with any content. Two measures of efficiency were used: the number-of-messages-taken-to-solution in a trial and the group time required to complete a trial. Both of these measures are group measures, and do not relate the group results to the individual performances in the group. In fact, in these experimental situations, it is impossible to do so, since the wide freedom in message content and sending times makes it impossible to know what information man A had when he sent message x. Compounded into the first measure were messages about the way to solve the problem (organizational messages), messages directly related to the problem (problem messages), messages about the influence of the network and messages about the effect of feedback of previous results to the members of the group. As may be seen in Leavitt's paper (34), which reports Experiments 1 and 2, it is possible to show statistically significant differences between networks by using this measure, but impossible to account for these differences, and hence impossible to predict new results. The second measure, time to solution, did not lead to significantly different results. Even if it had, as it does in Experiment 4, it would be difficult to separate the various effects entering into the time: if the group increases its efficiency by reducing the number of messages but keeps the time per message constant, then the total time will be reduced; however, reducing the number of messages requires some thought, which in turn increases the time per message. Finally, some time-saving individual learning occurs: the subjects become used to the apparatus, they are more sure of what they wish to write and so write more rapidly, and they learn to abbreviate their messages. Such learning cannot be considered in the same category as group learning.

As we have mentioned, these results have already been published; hence, we shall not repeat them here except insofar as we have carried out new analyses. Experiment 3 does contain some results on learning, but these are more appropriately discussed in Chapter V. Experiment 5 is similar to 4, but because only the time aspect of learning has been analyzed at present, these results will be discussed in Chapter IV.

This leaves Experiment 4 as the subject of discussion for this chapter and much of the next. The purpose of this experiment was to separate some of the effects that were mentioned in connection with Experiments 1 and 2, either by elimination or by a true separation. In Appendix 1 a detailed description will be found; here we shall state only the principal properties of this experiment and the reasons for certain selections. First of all, all organizational messages were eliminated, and the problem messages were restricted to the form "man A has the number x," the subject having only to fill in x.

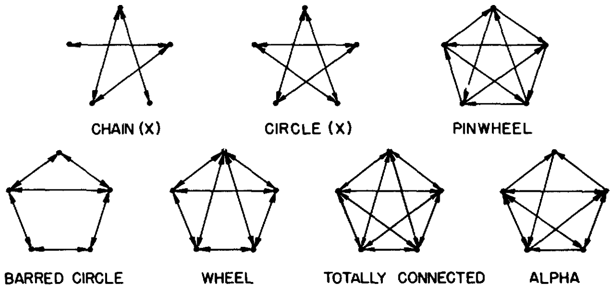


Fig. III. 1

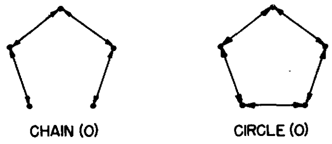


Fig. III. 2

This, of course, had the effect of reducing appreciably the number of different situations that might arise within the experiment, hence allowing the possibility of reasonable frequencies of occurrence of the different situations. Second, the subjects were no longer allowed to send messages at will. Each subject closed a relay when ready to send, and when all five relays were closed, a bell rang as the signal to send. At each such signal each subject sent one and only one message in which he was required to include all the problem information he knew at that time. The time between successive sendings of messages we call an act; the experiment thus consists of a series of trials and each trial of a series of acts. A trial was completed when each person knew what input information the others had. The subjects were told the minimum number of acts to solution (with one exception discussed below). The achievement of this minimum was their task.

The experimental program called for the examination of as many different networks as seemed feasible, subject to the condition that enough groups be run in each network so that there was some hope of obtaining the desired probability estimates. We elected to run 10 groups in each network for 25 trials, and to study the networks shown in Fig. III. 1.

Two minor side programs were added to this major program. First, it is apparent that many of the networks have several geometrical representations while preserving the same topology. Does this matter? It was decided that 5 cases each should be run on different geometrical versions of the circle and the chain (see Fig. III. 2). Second, each of the networks in Figs. III. 1 and III. 2 has a minimum solution of three acts, except the chain which has a five-act minimum solution. We decided to run 20 groups in the chain (x) network, and to tell 10 of the groups that the minimum would be 5, and the other 10 that it would be 3. These cases will be distinguished by writing chain (x-3) and chain (x-5).

These boundary conditions led to data from which it is possible to determine the information each man has when he sends a given message; thus we can relate his performance to what he knows and, to some extent, relate group performance as measured in acts per trial to individual performance. Separate from the decision problem is the data on time required for decision. Since in this experiment these two factors are independent, at least to a first approximation, we discuss them separately. At the end of the next chapter, which is devoted to time, we present some first mathematical steps toward re-introducing the dependence of one on the other.

2. Measurement of Learning

The performance of groups of subjects in our experiments can be called "group" performance because the groups have been intentionally made task-oriented and the tasks have been so designed that they require group effort for their solution. By the very means that group effort has been elicited, each member of the group has been required to perform. There is no group performance, in the sense we use this phrase, unless each member of the group performs in some one of the ways possible within the restrictions laid down by the group task, the initial conditions, and the boundary conditions.

The quality of the group performance can be specified in terms of the degree of fulfillment of the group task subject to the boundary conditions within which the group must work; the more efficiently the task problem is solved, the better the quality of the group performance. As a result of the experience of the group with the task, the efficiency of the group may increase. This improvement in the quality of group performance comes about because each member of the group has knowledge of the quality of the group performance at the end of each trial and this knowledge has an effect upon his behavior on subsequent trials. In other words, the group learns the problem because, and only because, the feedback of group results causes the members of the group to learn.

To express many of our experimental results we shall need terms descriptive of the group efficiency. Since no such measure is known which fits the general case, we have found it necessary to develop a measure for each experimental variant.

The number of acts actually observed on any particular trial is a sample of the group behavior possible at that time, rather than the only possible behavior. Therefore, our experimental data in Experiment 4 are in the form of act-per-trial distributions; that is, in the form of observed frequencies for each possible number of acts. The group acts-per-trial distribution is determined by a set of behavioral probabilities which characterize the individual performances at each node. The relation between the sets of individual probabilities and the group distribution function is many-one; that is, a set of individual node probabilities uniquely determines an act distribution, but a given act distribution may be the outcome of many different sets of individual probabilities.

It is easy to show that changes in individual behavioral probabilities may be produced by means we should not like to call learning. For example, the probability of a person's belching is higher after eating than prior to eating — we do not call such a change learning. In general, motivational and emotional states have as profound an influence on the probabilities of behavior as the state of relevant knowledge has. We can distinguish one effect from another, at least in principle, in terms of the operations which produce each. In practice, it is sometimes quite difficult to be sure that an operation such as the experiencing of a trial in our experiments does not have effects in more than one of these areas. Thus, although the running of the trials with the feedback provided is certainly a necessary and usually a sufficient operation to produce learning, it may also generate boredom and hence affect probabilities of response by an operation on the motivational state of the subject. However, having fixed probability as our indicator, we are still faced with the problem of finding an effective way to make comparisons between different values.

In principle there are at least four distinct ways to cope with the problem of measuring only the desired properties of a system:

a. Using a measure which is independent of any operation except those operations which affect the property being measured; namely, learning. The science of psychology has yet to provide such a measure for learning, and in particular, change in response probability is not such a measure.

b. Using a theory of behavior which includes a rich enough theory of measurement so that the influences of unwanted factors can be determined and corrected for. Again, this is not now possible because of the relatively poorly developed state of psychology.

c. Devising experiments to provide for statistical control of unwanted factors and employing the techniques of analysis of variance, particularly the covariance method, to get an approximation to the true picture of the independent influence of the variable being studied.

d. Attempting to make the groups comparable by the experimental control of variables other than those we wish to study.

The fourth method, one of the two open to us, is the one we have chosen. It is not wholly satisfactory since we have not completely controlled motivational conditions. However, the present state of knowledge in psychology makes possible our confidence that our efforts have been reasonably successful. Thus, although we cannot claim to have measured learning in the strictest sense, we can claim that any statistically significant differences in our learning indicator from group to group arise from real differences in learning.

Our data comes in the form of sets of frequencies or of frequency distributions, and it would be very convenient if these distributions could be reduced to single numbers. Miller and Frick (37) have recently proposed such a measure on the basis of information theory. This measure is obtained by treating the possible response choices of a subject as a set of symbols from which messages are formed, where the probability of a symbol being chosen is defined as the probability of the response to which it is associated. The selection of symbols is assumed independent. The average amount of information per symbol is then defined, for messages formed by random selections of symbols from the given ensemble of symbols, by the average number of binary decisions per symbol which are necessary to code the message. Where there are n choices to be made and p_i is the probability of choice of the i-th possibility, the amount of information is given by

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i.$$

As the set of probabilities diverges more and more from uniformity, the information-measure value becomes smaller and smaller. The summary number H has its maximum value, $\log_2 n$, when the probability is equi-partitioned among all possibilities, and it has a zero value when one possibility has the probability 1 and all others have the probability 0. Neglecting the information theory interpretation of the meaning of H and considering that a maximum of H corresponds to the zero of learning whereas the zero value of H corresponds to perfect learning, it is natural to propose the function

$$L = \log_2 n - H$$

$$= \log_2 n + \sum_{i=1}^{n} p_i \log_2 p_i$$

as a measure of learning. The measure L has the very desirable property that it measures the degree to which available choices are not chosen with equal frequency, and thus measures learning in a uniform way whether there is only one or more than one way of behaving upon which responses are concentrated. Moreover, it expresses what it measures as a single number.

There is one serious drawback to the use of the measure L. It measures the jaggedness of a distribution of probabilities but says nothing about where the peaks occur. Learning is always in terms of some problem, natural or imposed, and the relation of the problem to the distribution of response probabilities is that the problem specifies where the peaks must occur for perfect learning. With L we measure the degree of stereotypy of behavior but not its degree of conformity to the conditions of the problem. For example, suppose two rats are run in the same experimental situation in which there are only two terminal responses possible. Their behavior can be characterized by two probabilities: p_1 for one response, and p_2 for the other response. Let us suppose response 1 is rewarded with food and response 2 is not so rewarded. Thus a solution of the problem has been achieved to the degree to which response 1 is made in preference to response 2. Let us further suppose that rat A comes to make only response 1 and rat B comes to make only response 2. In this case (actually approximated in some experiments) $p_1 = 1$ and $p_2 = 0$ for rat A, and $p_1 = 0$ and $p_2 = 1$ for rat B. The measure L will have the same value for each rat, and in a sense, both have learned the same amount, but rat A has learned what the problem required whereas rat B has learned exactly the opposite. It may well be argued that for the situation of our example to arise there must have been some other combination of motivation and reward influencing the rat's behavior than that of hunger and food; for example, anxiety at, and escape from, the starting point. If we grant this, then it may be said that the situation was one involving two superimposed problems and all that is needed is to so arrange the conditions of the experiment that only the desired problem can have an influence on behavior. This is indeed a desiratum, but some problems are intrinsically of such a sort that one can not accomplish such a separation.

Any measure based solely on probabilities of response, whether a distribution or some statistic derived from a distribution, is bound to fail to describe learning completely. We may prove this by an example: consider two animals equal in every respect except for two responses. Suppose one animal has learned each response to the same high degree and the other has learned each to the same low degree. Now, since the two responses have been learned to the same degree, when either animal is put in a situation where it must choose one of the two, its choices will be equally distributed between the two; that is, the probability of response, for each response and for each animal, will be 1/2. Thus any measure based solely on these probabilities must describe the animals as having learned the same amount. This result is a source of difficulty only if the logical distinction we have made also makes an empirical difference; i.e. there must be some way to show that two organisms have learned responses to different degrees even where there are no differences in probability of response. This has been shown to be the case by testing the difference in response strength of the two animals by placing each in a situation where only one of the two responses they have learned can be made. There will be a difference in their behavior; to wit, the animal which has learned to the greater degree will respond with a shorter latency. Moreover, there would be a corresponding difference in the response times of the animals even if they were tested with both possibilities of behavior open. These facts show that there are at least two measures relevant to learning: probability of response and latency of response. Our realization of the importance of time measurements for our experiments came too late for us to have collected all the data of this sort we now wish we had. Fortunately, we have recorded some time measures and the results we can derive from them are presented and discussed in the next chapter.

The above argument shows that it is reasonable to hypothesize a construct which might be termed "response strength." It could be defined as a property of an entity which might be either an organism or an organized group. In either case, it would have the following characteristic: from a statement of the set of response strengths which characterize an entity at a given time and of the inputs to that entity at the same time, probabilities and latencies of these outputs can be derived. As so defined, response strength is not the ultimate parameter upon which learning has its effects, because it is also influenced by such factors as motivational level. However, if these other factors are held constant experimentally, changes in response strength will be monotonically related to learning effects.

There is one final problem to be mentioned which applies to whatever measures we use. We have written of probabilities and latencies as if their values were strictly determinable. Of course, all that can be measured are statistical estimates of these hypothesized values. Our procedure has been to take some small sample and treat the statistical value obtained from that sample as not too different from the parametric value of the variable being measured. This is a perfectly defensible procedure if the population from which the sample is drawn remains unchanged statistically as the sampling is going on. The problem that confronts us, as it does all behavioral sciences, is that the process of drawing one sample destroys the population from which the sample was drawn. The problem does not arise in the physical sciences, in general, for each electron is like every other, and hence an experimenter can sample ad libitum until he is satisfied with the stability of his measurement. Thus we are plagued by either changing population or the unreliability of small samples. Specifically, in studying learning one cannot take repeated samples of behavior in a situation that has not been previously experienced, since taking the sample unavoidably generates the experience. In practice we take a random sample from an ensemble of individuals and form several groups and then assume that these ensemble statistics are the same as the ideal time statistics we should like. It is not clear that this assumption is justified.

Our procedure differs not a whit from common practice in the behavioral sciences; it is neither better nor worse. What we have done is to compromise on both counts, as little as possible on each, and at the same time to get reasonably stable measures. That is, we have taken our statistics over both time and the ensemble. As a specific example we have run ten groups in an experiment for twenty-five trials, then divided the trials into five sets and taken statistics over each set of trials lumping all ten groups together. As will be seen, we have statistical justification that the time breakdown is sufficiently fine so that serious errors do not result. However, the grounds for the over-all procedure are heuristic.

3. Acts to Completion

The primary group data obtained in Experiment 4 were the number of acts to complete a trial. For each network configuration these have been broken down into four nonoverlapping blocks of trials and the percentages presented in Table III. 1. In addition, we have plotted in Fig. III. 3 through Fig. III. 12 blocks 1-6 and 11-25 and the equiprobable random distribution which is discussed below.

There is in each case an added distribution which is called the equiprobable random distribution for that network. This is essentially a mathematical property of the network which is obtained by assuming the group action is quantized and that each node is a random sender which treats each of the available outgoing links as equiprobable. Furthermore, the node is assumed to send all the problem information he has at the time. It is as if we gave each person in the experiment a balanced n-sided die, where n is the number of outgoing links, and caused him to decide on his sendings by throwing the die. We are then interested in the number of cases that will be completed in minimum acts, minimum + 1 acts, and so forth, i.e. the

Table III. 1a

		Proportion of Cases Completed in Minimum + i Acts					
Network	Trial Block	Minimum	+ 1	+ 2	+ 3	+ 4	+ 5
Chain (x-5)	Equiprobable Random	0.078	0.149	0.197	0.180	0.118	0.103
	1 - 6	0.383	0.250	0.183	0.050	0.033	-----
	7 - 12	0.467	0.317	0.133	0.033	0.017	-----
	13 - 18	0.717	0.217	0.017	0.033	-----	-----
	19 - 25	0.686	0.243	0.043	-----	-----	-----
Chain (x-3)	Equiprobable Random	0.078	0.149	0.197	0.180	0.118	0.103
	1 - 6	0.383	0.333	0.183	0.033	0.033	-----
	7 - 12	0.533	0.300	0.100	0.017	-----	-----
	13 - 18	0.633	0.200	0.050	-----	-----	0.017
	19 - 25	0.686	0.143	0.014	0.029	0.014	-----
Chain (0)	Equiprobable Random	0.078	0.149	0.197	0.180	0.118	0.103
	1 - 6	0.367	0.300	0.200	0.100	-----	0.033
	7 - 12	0.700	0.167	0.033	-----	-----	-----
	13 - 18	0.767	0.200	-----	-----	-----	-----
	19 - 25	0.857	0.086	0.029	0.029	-----	-----
Circle (x)	Equiprobable Random	0.002	0.168	0.377	0.268	0.113	0.042
	1 - 6	0.100	0.267	0.483	0.133	0.017	-----
	7 - 12	0.283	0.417	0.267	0.033	-----	-----
	13 - 18	0.483	0.367	0.133	0.017	-----	-----
	19 - 25	0.457	0.414	0.071	0.043	-----	-----
Circle (0)	Equiprobable Random	0.002	0.168	0.377	0.268	0.113	0.042
	1 - 6	0.233	0.467	0.267	0.017	-----	-----
	7 - 12	0.367	0.467	0.167	-----	-----	-----
	13 - 18	0.300	0.633	0.017	0.017	-----	-----
	19 - 25	0.314	0.629	0.057	-----	-----	-----

Table III. 1b

		Proportion of Cases Completed in Minimum + i Acts					
Network	Trial Block	Minimum	+ 1	+ 2	+ 3	+ 4	+ 5
Barred Circle	Equiprobable Random	0.008	0.209	0.312	0.235	0.111	0.069
	1 - 6	0.067	0.433	0.333	0.117	0.050	-----
	7 - 12	0.183	0.467	0.317	0.033	-----	-----
	13 - 18	0.217	0.533	0.233	-----	-----	0.017
	19 - 25	0.343	0.443	0.171	0.014	-----	0.029
Alpha	Equiprobable Random	0.011	0.273	0.306	0.208	0.100	0.059
	1 - 6	0.167	0.383	0.333	0.100	0.017	-----
	7 - 12	0.200	0.483	0.217	0.083	-----	0.017
	13 - 18	0.233	0.483	0.217	0.067	-----	-----
	19 - 25	0.371	0.357	0.229	0.043	-----	-----
Wheel	Equiprobable Random	0.026	0.312	0.357	0.174	0.084	0.036
	1 - 6	0.117	0.433	0.333	0.117	-----	-----
	7 - 12	0.167	0.617	0.167	0.050	-----	-----
	13 - 18	0.267	0.483	0.233	-----	-----	-----
	19 - 25	0.271	0.614	0.100	-----	-----	-----
Totally Connected	Equiprobable Random	0.041	0.397	0.326	0.159	0.053	0.019
	1 - 6	0.117	0.633	0.200	0.050	-----	-----
	7 - 12	0.167	0.583	0.167	0.067	-----	0.017
	13 - 18	0.217	0.567	0.217	-----	-----	-----
	19 - 25	0.143	0.686	0.157	0.014	-----	-----
Pinwheel	Equiprobable Random	0.038	0.530	0.299	0.101	0.025	0.007
	1 - 6	0.050	0.667	0.266	-----	-----	-----
	7 - 12	0.083	0.650	0.250	0.017	-----	-----
	13 - 18	0.050	0.683	0.250	0.017	-----	-----
	19 - 25	0.057	0.743	0.186	0.014	-----	-----

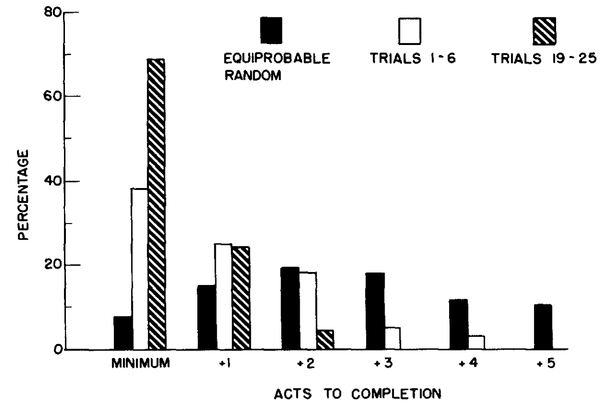


Fig. III.3
Chain (x-5).

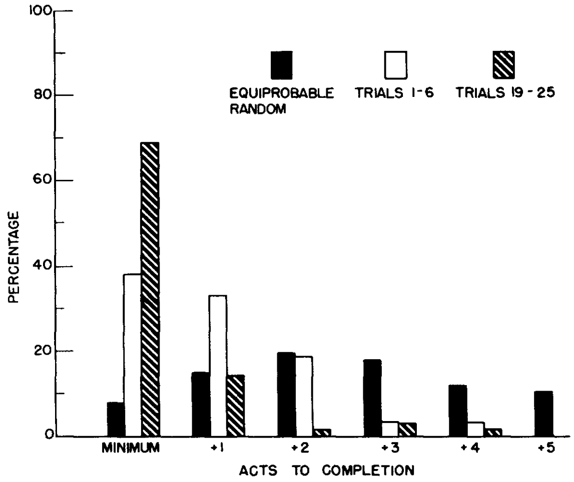


Fig. III.4
Chain (x-3).

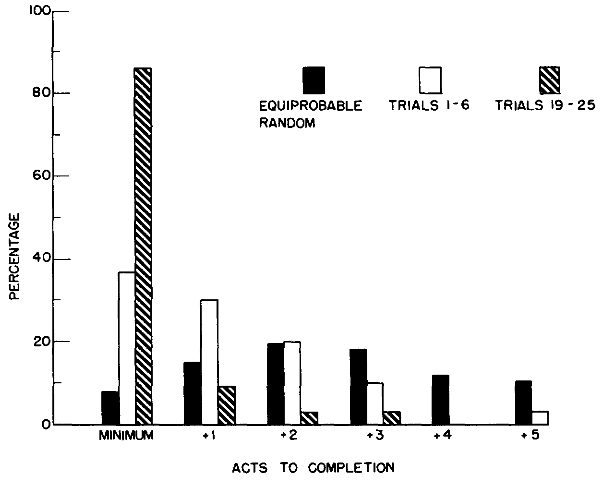


Fig. III.5
Chain (0).

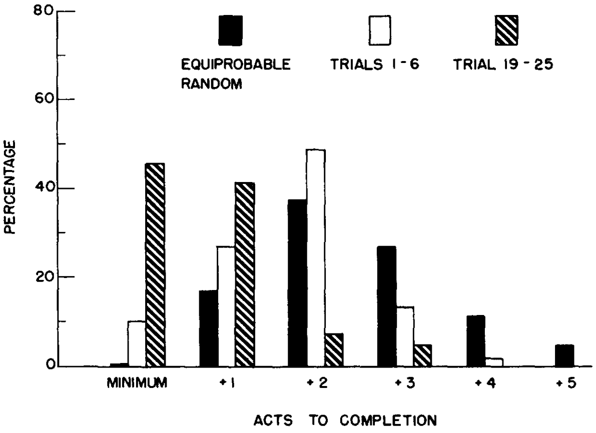


Fig. III.6
Circle (x).

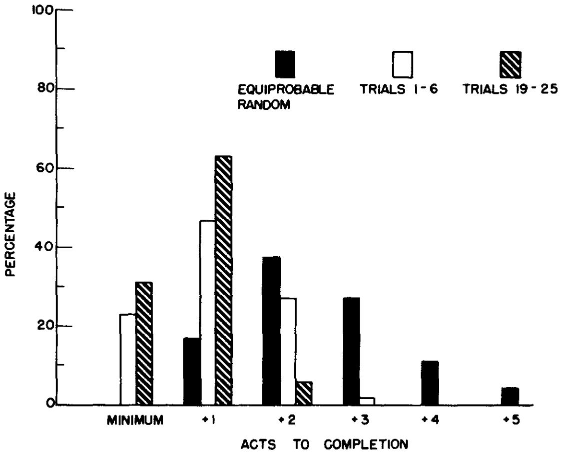


Fig. III.7
Circle (0).

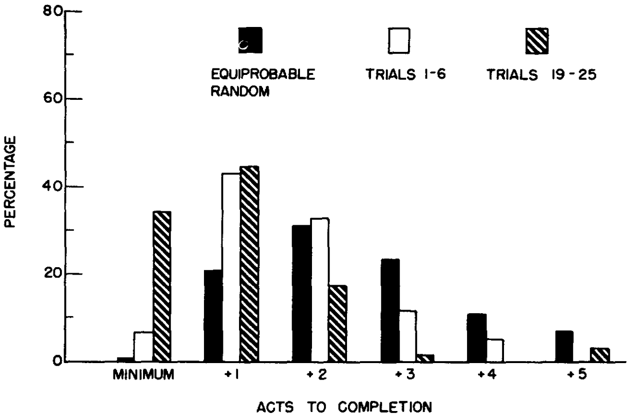


Fig. III.8
Barred circle.

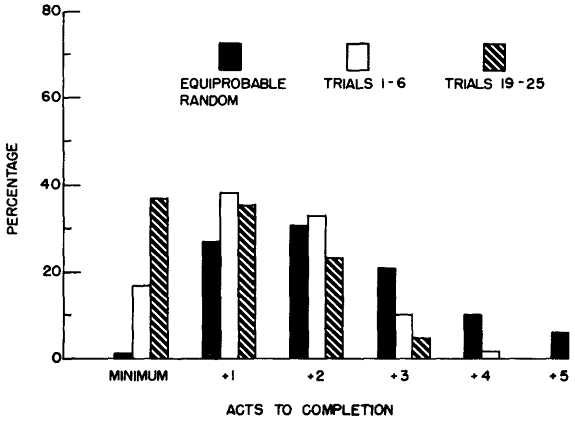


Fig. III.9
Alpha.

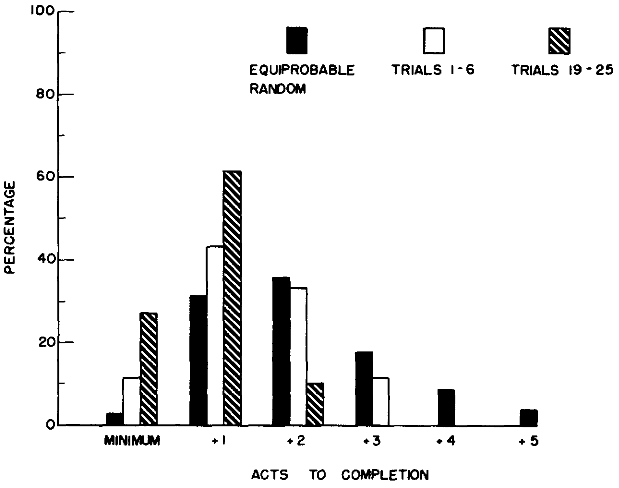


Fig. III.10
Wheel.

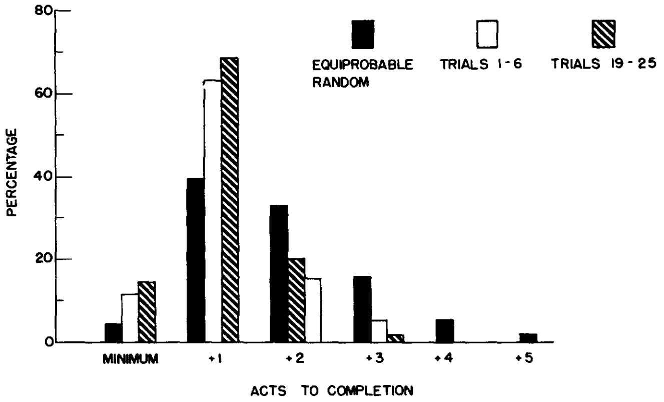


Fig. III. 11
Totally connected.

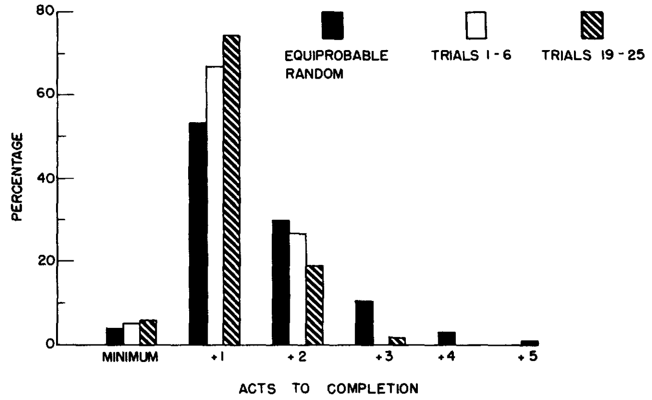


Fig. III. 12
Pinwheel.

distribution of acts to completion. (Mathematically, we may state the problem as follows: Let N be a matrix associated to the network;

$$N_{ij} = 1 \quad \text{if, and only if, there exists a link from } i \text{ to } j$$
$$= 0 \quad \text{otherwise.}$$

Let $\left\{N_\alpha\right\}$ be the set of all possible distinct matrices having the following properties: For any i there exists one and only one $j \neq i$, selected from the set such that $N_{ij} = 1$, such that

$$(N_\alpha)_{ij} = 1, \quad i \neq j$$

otherwise

$$(N_\alpha)_{ik} = 0, \quad j \neq k$$

and

$$(N_\alpha)_{ii} = 1.$$

If an equiprobable distribution is assigned to $\left\{N_\alpha\right\}$, what is the distribution of n such that

$$\left[\prod_{k=1}^{n-1} N_{\alpha(k)}\right]_{ij} = 0 \quad \text{for at least one pair i, j}$$

and

$$\left[\prod_{k=1}^{n} N_{\alpha(k)}\right]_{ij} > 0 \quad \text{for all i, j).}$$

It is believed that there is no known general solution to this problem, except in the case where the network is symmetric and has no closed loops of symmetric links, i.e. a tree in graph theory. As a consequence, we have resorted to Monte Carlo techniques using the high-speed digital computer Whirlwind I at M.I.T. to obtain statistical estimates of these distributions. We shall not go into the details of the coding problem here except to say that the source of random numbers were Kendall and Babington Smith's "Tables of Random Sampling Numbers" (86), a block of which were converted to I.B.M. punch cards by the RAND Corporation. From these cards the numbers were transferred to teletype tape. Finally, these decimal numbers were converted in Whirlwind I to binary digits by using the binary equivalents of 0 through 7 and discarding 8 and 9. Three thousand trials were carried out for each network, with a read-out at 1000 and 2000 as a stability check.

Looking only at the experimental results, we see first that with the exception of the pinwheel and totally connected networks, learning over trials occurs in all cases in the sense that initially there is a lower weight on solutions in minimum acts and a higher weight on minimum + 1 or + 2 than in the last trial block. Observe that if we were to apply the information theory measure without any reference to the reinforcement, we would not always detect this learning. For example, there is about the same peaking in the initial distribution for alpha, Fig. III.9, as for the final trial block distribution: it is just differently ordered. The H values are 1.9219 and 1.7434, respectively.

Without a well-defined measure of learning it is somewhat difficult to evaluate these results. It appears, for example, that circle (x), barred circle, and chain (x) form a sequence of decreasing learning, and that barred circle, alpha, wheel, and totally connected form another sequence of decreasing learning. A satisfactory measure would have to indicate these intuitive orderings.

We must sidestep this problem, but there are others of equal importance. Can we in any way predict these distributions? Looking at the latter sequence of decreasing learning suggests that increasing network complexity, in some yet to be defined sense, is correlated with a decrease of learning. This may indeed be true, but it cannot be the entire story for the pinwheel network shows the least learning of all. In many ways it is less complex than the barred circle; rather it is of the same order of complexity as the circle. It is our belief that it is fruitless to attempt to correlate such a complex resultant of individual learning as group learning with any single combinatorial or topological parameter of the network. An understanding of group learning in this type of situation will arise only through an understanding of the individual learning and the combinatorial properties of the network which compose this into a final group result.

Before turning to such an analysis of some of these networks, we shall discuss in a gross way the equiprobable distribution in relation to the experimental results. The most notable feature is the difference between the random case and the observed group results, even for the first trial block where one might suspect a large number of random decisions. It is interesting that there is a decreasing difference between the random and trial block 1-6 with decreasing learning in the network, with both pinwheel and totally connected having the least difference and circle and chain the most.

We may draw two principal conclusions: (a) Learning does in fact occur with increasing trials, the amount varying from network to network; (b) It is not correct to assume that the individuals begin to operate initially as equiprobable random senders.

4. The Quasi-Discontinuous Nature of Group Learning

The task set our subjects was to solve each of their series of problems in as few acts as possible. Each group (with one exception) was told the correct minimum number of acts for its network, and this knowledge gave them a goal and a standard for evaluation of their performance. While it is true that a performance not as efficient as a minimum performance may represent an improvement over performance on preceding trials and thus indicate to the group that progress toward the goal seems to be occurring, yet the achievement of minimum solutions is strongly called for by the instructions to the subjects. It therefore seems reasonable to dichotomize the efficiency of performance dimension into "success" for minimum solutions and "failure" for nonminimum solutions. Having done so we arrive immediately at proportion of successes as an indicator of degree of learning. The learning performance of the various experimental groups on this basis is shown in Figs. III.13 and III.14. Reference to Table III.4 in section III.6 will provide an example of the data from which these curves were drawn. A careful comparison of the table with the corresponding entries on the graphs will be sufficient to suggest that the form of the graph obscures important features of the learning process in these groups. Let it be noted that the points in trials where the several groups give evidence of having mastered the problem, that is, where they begin to get minimum solutions consistently, are widely dispersed. On the other hand, the number of trials it seems to take to make the sharp change in behavior from very few minimum solutions to almost all minimum solutions is small.

Since we have dichotomized the performance variable, the occurrence of a success or a failure can be looked on as the drawing of a sample of one from a pool of responses in which the proportion of successes is p, the value of p being a characteristic of the

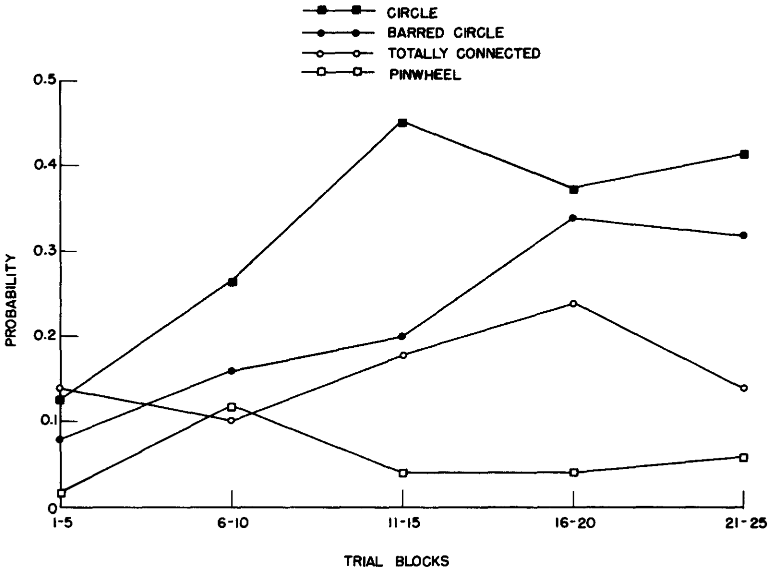


Fig. III. 13
Probability of minimum solution.

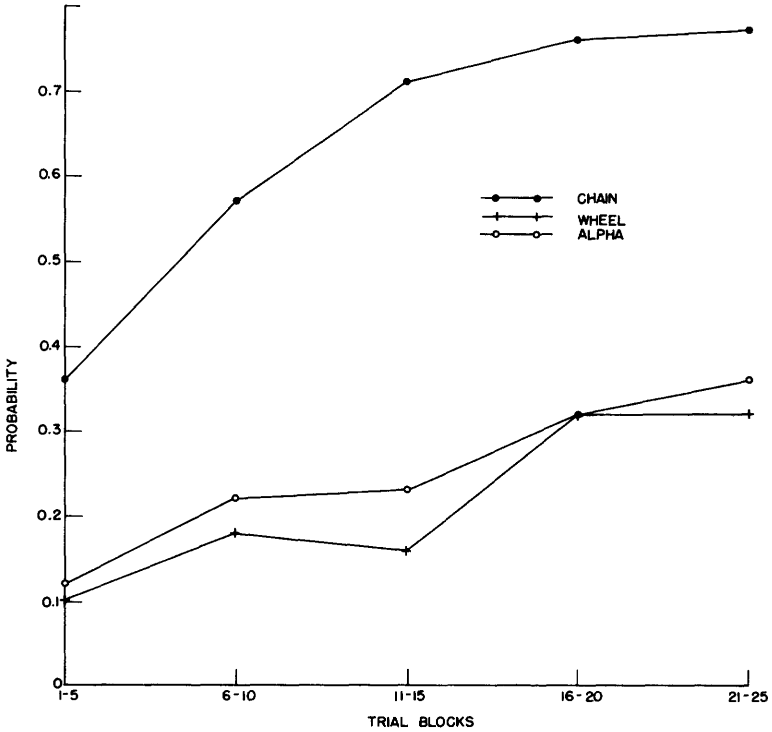


Fig. III. 14
Probability of minimum solution.

group at a particular stage in its process of learning. Over short blocks of trials the value of p for a particular group will be nearly constant, except possibly once in the life history of each group when it makes its major jump to problem mastery. We can, therefore, use the series of trials in such a short block to generate a sample from a binomial population of parameter p and use statistics on this sample to estimate p. It is reasonable further to assume that before learning for any group the p's are nearly equal. Now, if some one or more of the groups effectively masters the problem over a few trials and the balance of the groups does not, the mean value of p taken over the whole set of groups will be a compromise between the high p for the mastery groups and the low p for the nonmastery groups and will be representative of no actual group. If this hypothetical situation of quasi jumpwise learning does occur, then the variance of p will be (except for sampling fluctuation) nearly zero before any group learns and after all groups have learned, but it will be considerably different from zero when some of the groups have learned and others not. If, on the contrary, the learning were a smooth process similar in each group, the variance of p would remain near zero throughout the learning process.

Let n be the number of groups and let each be characterized by its p_i and $q_i = 1 - p_i$. If we then make up samples drawing once from each group, we shall have frequencies arrayed in the form

$$\prod_{i=1}^{n} (p_i + q_i).$$

If all the p_i's are equal, this reduces to the binomial, and, of course, equality of the p_i's is equivalent to a zero variance for p. It is readily shown that the second moment of this distribution is given by the expression (ref. 85, vol. I, p. 122):

$$\mu_2 = n\bar{p}\bar{q} - n \text{ var } p.$$

Solving for var p, we find

$$\text{var } p = \bar{p}\bar{q} - \frac{\mu_2}{n}.$$

If we show the hypothesis var p = 0 to be rejected when the data suggest that one of a set of groups has learned and not otherwise, we shall have the evidence we need. The variance for each group is plotted against trials in Figs. III.15 and III.16. Table III.2 gives the random sampling probabilities of the observed values of var p by networks and trial blocks. These results are in agreement with the belief in a point of discontinuity in the learning except in the case of chain (x). In this case there may also be some degree of quasi discontinuity, but the data do not suffice to demonstrate it. That the pinwheel shows values of var p which do not differ significantly from zero in any trial block accords perfectly with the complete failure of this network to learn a minimum solution. Similarly for totally connected, the variance stays very close to zero except

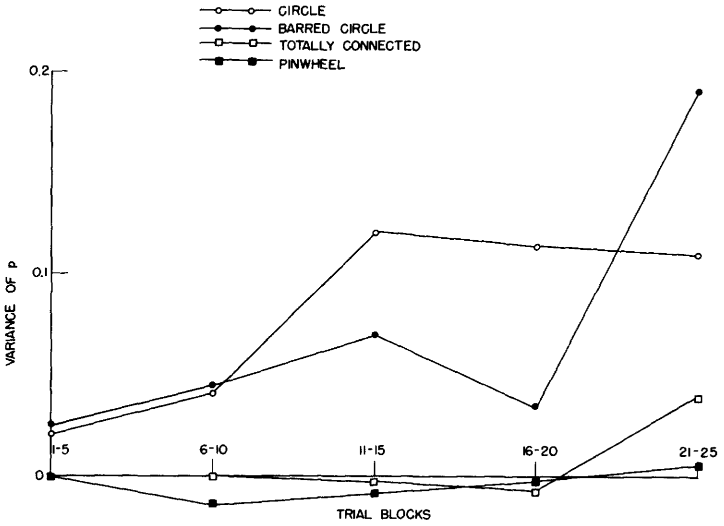


Fig. III. 15
Variability of degree of learning.

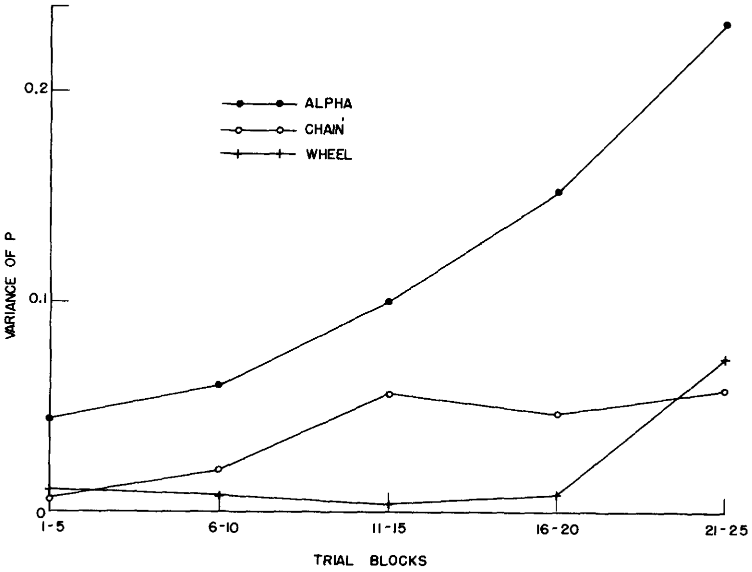


Fig. III. 16
Variability of degree of learning.

Table III.2

Sampling Probabilities for var p

Network			Trial		
	1 - 5	6 - 10	11 - 15	16 - 20	21 - 25
Alpha	<0.01	<0.01	<0.00001	$<0.0^{8}1$	$<0.0^{8}1$
Circle (x)	0.45	0.05	<0.00001	$<0.0^{5}1$	$<0.0^{7}1$
Barred circle	0.03	<0.01	<0.001	0.07	$<0.0^{8}1$
Wheel	0.43	0.68	0.79	0.75	$<0.0^{8}1$
Totally connected	0.98	1.00	0.99	0.96	<0.01
Pinwheel	0.99	0.30	0.26	0.98	0.94
Chain (x)	0.10	0.38	0.08	0.38	0.52

for the final trial block in which one group appears to be mastering the problem.

The features we have just demonstrated in the experimental data, that individual groups do not follow the same course as the mean of a set of groups, must be taken into account in any reasonable theory constructed to cover the type of group behavior our experiments have dealt with. Such a theory must allow for slow changes in behavioral tendencies except that under certain conditions (which the theory must specify) the occurrence of at least one and probably two or more successive minimum solutions gives rise to a very rapid alteration of behavioral tendencies.

5. Discussion of Learning in Circle and Pinwheel

If we agree that we have established the existence of group learning to varying degrees, the amount being some function of the network, then the next question to be asked is how the group learning is related to the learning of individuals. In principle this is straightforward for the type of communication occurring in Experiment 4, but in actual practice there is considerable difficulty. If we first discuss the type of model we have in mind, the difficulties will be clear.

Essentially, we are looking for nodal statistics which, by combinatorial techniques, will lead to the group statistics. Evidently, the nodal statistics will have to be in the form of conditional probabilities which state that if certain conditions have been met in the past, then the probability that node i will send his message to node j is p. By the very meaning of these statistics, it follows that if sufficient conditions are included, the group results will be given, the trivial case being when each of the groups examined is treated as unique, and prediction is given only for those situations. Such an approach will be important only if the number of conditions that have to be considered is relatively small, so that situations which have not actually arisen in our experiments can be predicted. This is to say, we must so select the conditions that (a) when computations are carried out combining the nodal statistics for a given network, the result differs from the observed group results by an amount that can be confidently attributed to chance fluctuations in the data, and (b) the conditions are sufficiently simple that stable nodal statistics are obtained from a reasonable amount of experimental data.

It seems to be in the nature of the situation that we assume a learning model having a structure similar to that given by Bush and Mosteller (16, 17). Such models take the form of an operator which operates on the i-1, i-2, ..., i-k stages of the process to give the i-th stage. This, mathematically, leads to recursion expressions which, when appreciably simplified, can be solved in terms of some initial conditions raised to the i-th power. A model of this type, coupled with the first condition we mentioned, requires a strengthening of the second condition, for if there are small errors in our estimates of parameters these errors will be so seriously magnified when i is, say, 20 that it is very unlikely that the model will numerically fit the data even though it does conceptually. This type of cumulative error seems characteristic of learning models, and it suggests that there is an area of work on the stability of such models.

A second major difficulty is a mathematical one. Supposing that we have sets of conditional probabilities and a particular network, can one obtain the group results mathematically? As we pointed out earlier, the simpler problem of determining the group act statistics, when equiprobable random node distributions were assumed, required the use of a Monte Carlo technique on a digital computer. It therefore follows that for any more complex conditional probabilities we shall again have to use this technique, unless there are some particularly simple questions for particularly simple cases that are subject to analytic treatment. We shall be able to do some work with the circle. The remainder of the networks have not been analyzed in such detail, and probably can only be by means of a high-speed computer. However, because of the cumulative-error difficulty, it remains to be decided whether such a program is worthwhile at this time.

Let us consider what factors we wish to include. A primary condition for the probabilities in an analysis of learning must be the reward which conditions the learning. In our experiments each trial had associated with it a set of bell rings. This acted as a reward if it was the minimum possible set for the network. So, we shall distinquish whether the previous trial was completed in minimum acts or not. Of course, if we have an experiment in which the minimum occurred so infrequently that it can be neglected, the occurrence of minimum solution is not a condition. As was shown in sections III. 3 and III. 4, this is true for pinwheel in which the minimum was obtained about 5 percent of the time.

It also seems clear that the state of information at a node, in relation to what that node knows or believes to exist at nodes to which he communicates, is important. Determination of this, of course, is very complex. First of all, without a much more complex design we can have no information as to what a person believes another person to know about the initial data given the group. In principle, we can determine what is the most he could logically know. This is extremely difficult, and in most cases there simply is not a sufficient amount of data to give decent frequencies for all the different possibilities. The next, simpler, step is to consider whether the node under consideration could or could not know if his message would add any new information to the receiver. If we do this, we ignore situations where node i knows that if he sends his information to j he will add one new piece and if he sends it to k he will add 3 new pieces. This is probably not a serious error for more often than not the subject could not have this much detailed information. We shall give this type of analysis for the chain, but it is computationally still too difficult for any other network.

The next, and final, step in simplification is to consider only the source and destination of past messages as the condition of future messages. In this we completely ignore what the content of the message was. We shall further restrict our consideration to the messages at one node on the previous act. So our probabilities will depend on: (a) whether the previous trial was completed in minimum acts, (b) to whom the node under consideration sent his message in the previous act, and (c) from whom the node under consideration received messages in the previous act.

These three types of conditions result in, for the totally connected network, 32 different probabilities to be computed and, of course, fewer for less complex nodal arrangements. This is still too many for most networks considering the amount of data we have, so we have found it necessary to restrict our attention to some of the simpler networks. We shall present in detail an examination of the circle, chain, and pinwheel. The circle and pinwheel have the advantage that each node is like each other node in its linkage relations; so, assuming statistically identical people, we may lump the data for all the nodes in each of the networks. In the chain the end nodes are topologically identical, as are the adjoining ones, called the middle nodes, but the center node is unique in being related to the end nodes in a manner different from the way the middle nodes are. Of course, the end nodes will be ignored, for their behavior is completely stereotyped by the conditions of the experiment. It may well be possible to carry out the analysis for the barred circle, but this has not yet been done.

We first examine the pinwheel. As we mentioned, we do not have to make the probabilities conditional on the previous trial since trials in minimum acts occurred so infrequently as to make them negligible. The topologically distinct input-output conditions that are possible on the previous act are given in Fig. III.17 where the solid arrows indicate the sending of a message, and the dotted ones an open channel. The cases are numbered and distinguished according to the number of inputs. In Fig. III.18 are plotted, for acts II and III combined, the conditional probabilities that a person will send to the person he did not send to in the previous act—the probabilities of alternation.

These results make sense from a "locally rational" point of view if we assume (a) that a person can remember only for the previous act, and (b) that a person does distinguish between giving one person more information than another. For if this is the case and if p(A,X) is the probability of alternation under the condition X, then

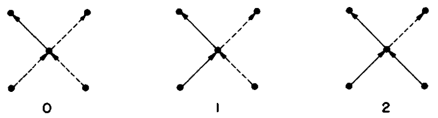


Fig. III.17

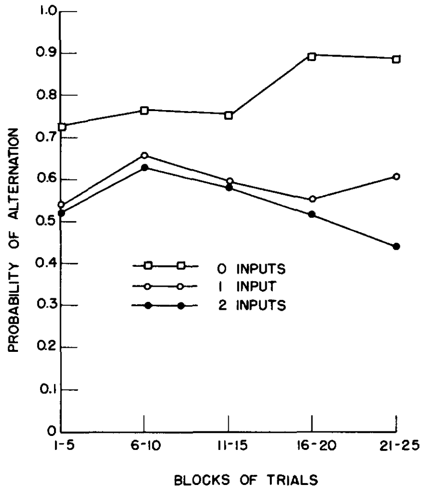


Fig. III.18
Conditional nodal probabilities for pinwheel.
Acts II and III combined.

what we may term locally rational behavior would dictate

$$p(A, 0) = 1$$

$$p(A, 1) = p(A, 2) = 1/2$$

If we weaken the second assumption and say that he will send to the person to whom he can give the most information, but with his decision based only on the input-output relations of the immediately proceeding act, then we should expect

$$p(A, X) = 1, \quad X = 0, 1, 2.$$

If we admit the person has a better memory, then all we can say is

$$p(A, 0) = 1$$

$$p(A, 1) \geqslant 1/2$$

$$p(A, 2) \geqslant 1/2$$

The data closely approximate 1/2 for conditions 1 and 2, and there is a definite trend toward 1 for condition 0.

For the circle, the possible topologically different nodal configurations are given in Fig. III.19. These cases have been called, respectively, none, same, both, and opposite, and denoted by N, S, B, and O. However, in the case of the circle we must

Fig. III.19

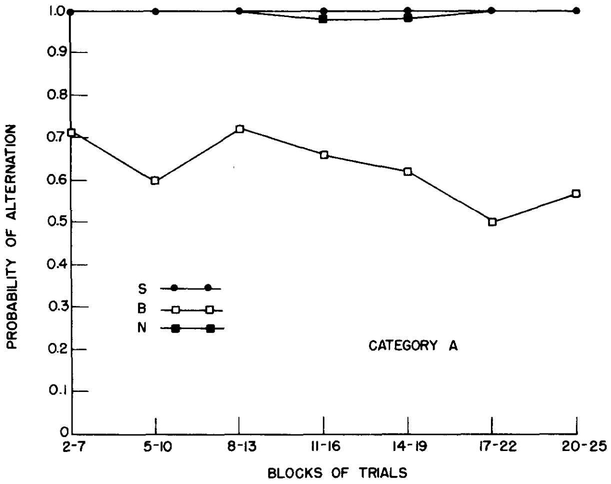


Fig. III.20
Conditional nodal probability for circle (x). Acts II and III combined.

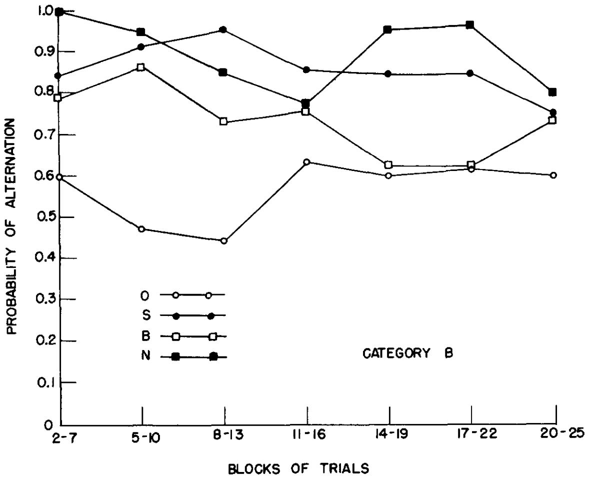


Fig. III.21
Conditional nodal probabilities for circle (x). Acts II and III combined.

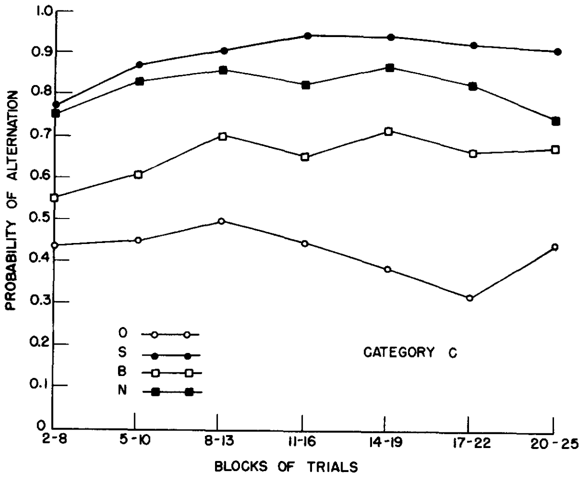


Fig. III.22

Conditional nodal probabilities for circle (x). Acts II and III combined.

take into account the feedback to the group as to their performance on the previous trial. We may consider the following exhaustive categories:

A. The previous trial was completed in minimum acts, and the present trial has been locally the same in the sense that the node under consideration has received messages from and sent messages to the same nodes as in the previous trial up to the point under consideration.

B. The previous trial was completed in minimum acts, but the present trial has not been locally the same.

C. The previous trial was not completed in minimum acts.

As in the pinwheel we have combined the statistics for acts II and III and have ignored any remaining acts. The reason for this is that we expect the statistics to be dependent on the act number, with greater aberrations occurring for a large act number. Having separated act II from act III we found that the differences seemed to be well within the variability of the data, and so for the final plots they have been lumped together. The results are plotted in Figs. III.20, 21, and 22. Observe that there is no plot of the "opposite" category in case A; this is due to the fact that no such case can occur. The number of cases in the B category was much smaller than in the other two, so much so that the data are highly variable. Later, we shall show why this is the case.

It is interesting to note that most of these curves seem to be to some extent independent of the trial block. Since learning is a function of experience in the experiment, we might expect all the curves to change monotonically with trials. There seem to be two reasons why a statistic may be independent of trials: (a) It may simply be independent of experience in the experiment. (b) It may be that the breakdown of the conditions includes the relevant experience, leaving the residue essentially independent of trials. The latter means a change in the sampling population because of our categorization.

For example, we see in the plot for the C condition that in the last trial blocks there is a drop in the "same" and "none" categories when, naively, we should expect an increase, or at least no change. This may be explained as follows: as the trials progress, more and more groups achieve a minimum solution and stick to it; this, it turns out, most often occurs in groups for which the alternation probability for the "same" and "none" categories is high. But having achieved a minimum solution they are removed from the C category and placed in the A and B categories, leaving on the average people in the C category who do not have as high a tendency to alternate. Thus, the sample from which the curves are obtained is changing.

If by P_N, P_B, P_S, P_O we mean the probability of a node alternating when in the state N, B, S or O, then looking at the circle from a "locally rational" view, as we did the pinwheel, we should expect

$$P_N = 1$$

$$P_B = 1$$

$$P_S = 1$$

$$P_0 = 1/2$$

Thus, comparing the individual statistics for both the pinwheel and the circle when the previous trial was not completed in minimum to the "locally rational" behavior, we see there are some differences in the degree of rationality, but they are not great. Equally well, the equiprobable random statistics for the frequency of minimum solutions indicate but little difference between these two networks. However, the observed group statistics yield markedly different results:

Table III.3

Percentage of Trials Completed in Minimum Number of Acts

Trials	Pinwheel	Circle (x)
1 - 6	0.050	0.100
7 - 12	0.083	0.283
13 - 18	0.050	0.483
19 - 25	0.057	0.457

We cannot yet say that the individual statistics coupled with the network topology will account for these differences; this awaits the computer solution. However, a further examination of the results does give support to the belief that they may.

Consider first the circle. By exhaustively writing all 32 possible ways in which each node may select one and only one outgoing link over which to send a message, i.e. all possible communication structures, it may be shown that there are exactly

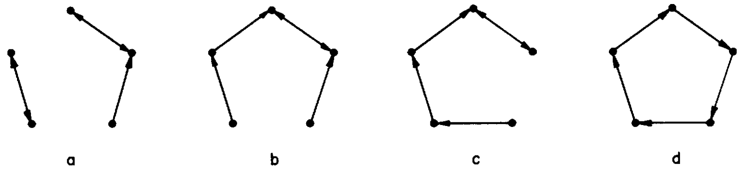


Fig. III.23

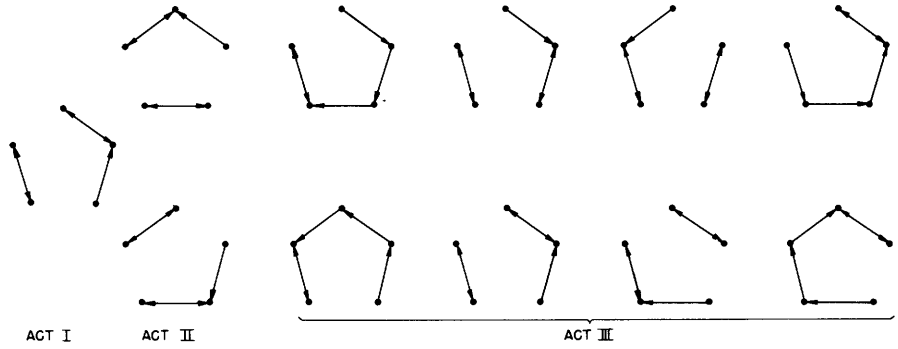


Fig. III.24

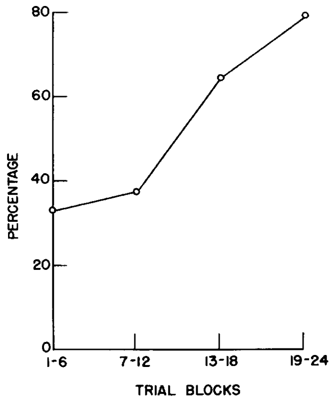


Fig. III.25
Perseveration of same initial-act structure for minimum trials in circle (x).

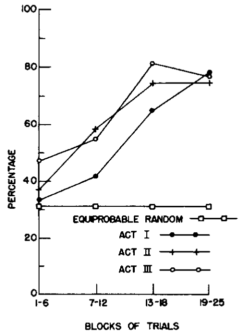


Fig. III.26
Percentage of type a communication structure in circle (x).

four topologically different cases (see Fig. III.23). For a given labeling of the nodes there are 10 each of the first three and two of the last type.

It can be shown that a minimum solution of the task, which requires that each person receive a message (possibly indirect) from each of the other nodes, can be achieved if and only if the act I communication structure is of type a. And, in fact, for a given type a structure there are 8 minimum solutions (see Fig. III.24). The proofs of these results are simple, but because they are lengthy, they will not be presented here. It thus follows that the probability of a minimum solution in the circle when the nodes have equiprobable random distributions is 0.24 percent, which compares with the value of 0.17 percent obtained from the digital computer.

If by P_N, P_B, P_O, and P_S we mean the probability of a node alternating when it is in each of the conditions N, B, O, and S, and if we suppose that a particular trial is begun with a type a communication structure, it is not difficult to show that the probability of a minimum solution is

$$P_N^2 \, P_S^5.$$

As can be seen in the curves of Figs. III.20, III.21, and III.22, these numbers are fairly large in all cases; to all intents and the purposes, the number is 1 if the preceding trial was completed in minimum and the present trial is locally the same as the last. Furthermore, from a local point of view, rational behavior would dictate that

$$P_N = P_S = 1.$$

In this case, how well the circle does depends only on the probability that the initial act structure is of type a. The random probability of obtaining a type a initial act structure is 5/16 = 0.3125. However, after it has been obtained, the probability of

Fig. III.27

obtaining a minimum solution is high, and if a minimum solution is achieved, Fig. III.25 indicates that the probability of obtaining the same type a initial structure on the next trial is also high. In fact, Fig. III.26 shows the increase of type a structures with trials. Thus there is an increasing tendency for the group to obtain minimum solutions simply because of the increasing probability of a type a initial act. It is reasonable to expect, and it is substantiated by the argument of section III.4, that once a group has traversed the same path to success two or, at the most, three times, the entire process will be so ingrained in the subjects' memories that they can continue obtaining minimum solutions until the experiment is concluded or until they become bored. Before presenting a more detailed mathematical analysis of learning in the circle, we shall discuss the pinwheel in a more or less analogous way.

For the pinwheel it can easily be shown that each of the 32 possible initial structures begins at least one sequence which is completed in minimum acts. Of these, only the two structures illustrated in Fig. III.27 allow a minimum solution by pure alternation. This may be neglected since each node is in the condition of 1 input, which, according to Fig. III.18, has a conditional alternation probability of approximately 1/2, yielding a probability of

$$\frac{1}{16}\left(\frac{1}{32}\right)^2 = \frac{1}{16,384}$$

for an alternating sequence. If we begin with any of the other possible initial structures and let the second act be obtained by alternation, there is at least one node, and sometimes more, in the third act that must not alternate and others that must alternate. There is no way within the context of the experiment for these people to know which of these they must do. It appears true, but has not been definitely shown, that in every minimum solution for the pinwheel there is at least one node which has one decision which cannot be based on the logic of his local environment nor on some such rule as, "I will do the opposite of what I did the last time." Rather, there is always a node which must perform in a fashion dictated by knowledge about the activities of other nodes which is very difficult, and sometimes impossible, for him to acquire.

Note, however, from the equiprobable random results, that if a network is desired which does very well without the individual nodes doing very well, the pinwheel is the most satisfactory of all the networks studied. More than half the cases are completed in 3 or 4 acts.

We may distinguish the two cases as follows: In the circle, locally rational behavior will lead to rational solutions (i.e. minimum solutions) for the group. For the pinwheel, any rational group behavior requires at least one person to perform in a fashion for which he has no basis.

6. Probability Model for Learning in the Circle

This section is devoted to making the discussion of the last section computationally more precise with respect to the circle network. The theory evolved will be only applicable to that network, though in principle the outlines could be applied to other cases if enough of the abstract properties of these networks were understood.

From the distributions of acts to completion given in section III. 3, it is clear that for the circle an appreciable increase in minimum solutions occurs over trials, and from the discussion of section III. 4, it is clear that this increase is not uniformly distributed over all groups. Some circle groups learn and others do not. We may see this even more explicitly in Table III. 4 which tabulates groups vs trials with the entries the number of acts to completion recorded. It is quite evident the groups 3, 6, 7, and 8 learn. It is our task here to make a theory which predicts reasonably well this learning on the basis of certain nodal transfer functions. To carry this out we shall need two things: a definition of what we shall mean by learning and an assumption as to what the transfer function shall take into account.

To define learning (in this situation) we shall, at any stage of the process, dichotomize the class of groups into those that are learned and those that are nonlearned in such a fashion that once a group is in the learned category it never leaves it. Definition: A group is nonlearned until it has obtained three consecutive minimum solutions; it is then placed in the learned category. Thus, we see from Table III. 4 that groups 1, 3, 6, 7, and 8 are in the learned category from trials 16, 7, 16, 15, and 9, respectively. This includes the four groups we mentioned above and one more, 1, which never, after it is placed in the learned category, achieves a minimum solution. This is an unhappy circumstance, and we do not at present see how to get around it. The belief is that this is a rare phenomenon, but we do not have the data to be sure. We shall want our theory to predict the probability that a group is in the learned category on the i-th trial. In addition, we should like to know the frequency of minimum solutions in the nonlearned category as a function of trials.

Let R_i be the probability a group is in the learned category on the i-th trial, and let L_i be the probability a group in the nonlearned category has gotten the trials i-2, i-1, and i in minimum. Then

$$R_i = R_{i-1} + (1-R_{i-1})L_{i-1} \qquad (1)$$

Our problem then is to evaluate L_i. To do this we shall introduce certain auxiliary variables:

$p_i(a)$ is the probability that the initial structure on the i-th trial of a group in the nonlearned category is of type a.

U_i is the probability that if on the i-th trial the group was in the nonlearned category, it obtained a solution in minimum acts.

We shall now obtain three recursive formulas, one each for the variables $p_i(a)$, U_i,

Table III. 4

Number of Acts to Completion: Circle (x)

Groups	Trials																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	6	7	3	4	4	4	5	4	3	4	4	5	3	3	3 |	4	5	4	4	4	4	4	6	4	4
2	4	5	6	6	6	5	4	5	5	4	5	4	4	4	4	4	4	4	4	4	5	6	4	4	6
3	5	5	5	3	3	3 |	3	3	4	3	3	3	3	3	3	3	5	3	3	3	3	3	3	3	3
4	5	6	5	4	5	6	5	5	4	6	5	4	4	6	4	5	5	5	4	5	4	4	5	4	5
5	5	5	5	6	5	5	4	4	5	6	5	4	3	4	5	4	5	3	3	4	3	4	5	4	4
6	5	5	5	4	5	4	4	5	3	5	3	4	3	3	3 |	3	4	3	3	3	3	3	3	4	4
7	4	6	4	4	4	4	5	4	3	3	4	3	3	3 |	4	3	3	3	3	3	3	3	3	3	3
8	5	5	4	5	4	3	3	3 |	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
9	5	5	5	4	5	3	4	5	3	4	4	4	4	5	4	4	4	4	4	3	4	4	3	4	4
10	5	5	4	5	5	5	5	4	4	5	4	3	3	4	3	3	4	4	4	3	4	3	4	3	4

*The vertical lines indicate transition to "learned" category.

and L_i. To do this, we shall have to make assumptions as to how far back the conditional probabilities of behavior extend. With the exception of the learning function L_i which is defined to extend over three trials, the remainder of the conditions will extend only to the previous trial, and then only consider whether it was done in minimum, and if so, whether the initial act of the present trial is the same as the initial act of the previous trial. To do this we must define certain conditional probabilities. These will be computed ultimately from the network and the observed values of the nodal transfer function (which is in the form of probabilities conditional on the structural configurations N, B, S, and O). For the moment we shall not be concerned with the means by which they are obtained; assume them known for the purposes of discussion.

$p_i(a, a)$ is the probability that the initial structure of the (i+1)st trial is of type a if the initial structure of the i-th trial was of type a, and that trial was not completed in minimum acts.

$p_i(a, \bar{a})$ is the probability that the initial structure of the (i+1)st trial is of type a if the initial structure of the i-th trial was not of type a. We do not distinguish whether the initial structure of trial i was type b, c, or d. It is probable that their occurrence is different from equiprobable random and that the transition probability to a is a function of the structure type; however, a few calculations have indicated that $p_i(a, \bar{a})$ does not vary greatly from its chance value of 0.3125, so in our calculations we assume this value. As we shall see, the task of calculating $p_i(a, \bar{a})$ from the nodal conditional probabilities is very time-consuming.

$q_i(a, a)$ is the probability that the initial structure of the (i+1)st trial is of type a if the i-th trial was completed in minimum acts.

$r_i(a, a)$ is the probability that the initial structure of the (i+1)st trial is the same type a structure as in trial i, when that trial was completed in minimum acts.

P_i is the probability of completing trial i in minimum provided the initial structure is of type a and the previous trial was not completed in minimum.

Q_i is the probability of completing trial i in minimum provided the previous trial was completed in minimum and the present trial has the same initial structure as the previous trial.

The third case, the probability that the i-th trial is completed in minimum, when the previous trial was completed in minimum and the initial structure is not the same (though of type a), is difficult to compute. For, as we saw in the last section, those people in the network who do not see any change from the previous trial will act as if the i-th trial had begun with the same structure as the previous trial. Our decision is to use the value P_i whenever this case arises, knowing that this introduces an error.

Now, the probability that initial structure is of type a on trial i, $p_i(a)$, is a sum of three components. First we may consider whether the previous trial was completed in minimum or not, and if not, whether the initial structure was of type a or not. If the previous trial was done in minimum, then there is a probability L_{i-1} that the group will go into the learned category. So,

$$\begin{aligned} p_i(a) &= U_{i-1}\, q_{i-1}(a,a) \left[1 - L_{i-1}\right] + \left[p_{i-1}(a) - U_{i-1}\right] p_{i-1}(a,a) \\ &\quad + \left[1 - p_{i-1}(a)\right] p_{i-1}(a,\bar{a}) \\ &= U_{i-1} \left[q_{i-1}(a,a)\,(1 - L_{i-1}) - p_{i-1}(a,a)\right] \\ &\quad + p_{i-1}(a) \left[p_{i-1}(a,a) - p_{i-1}(a,\bar{a})\right] + p_{i-1}(a,\bar{a}). \end{aligned} \tag{2}$$

The probability that a nonlearned group completes the i-th trial in minimum also falls into three cases. The last two, which are essentially the same breakdown as the nonminimum case above, simply require multiplying the last two terms of the first expression for $p_i(a)$ by P_i. The minimum case is more complex for it matters whether the initial structure is the same as in the (i-1)st trial, or a different type a structure. Taking this into account we have

$$\begin{aligned} U_i &= \left\{U_{i-1}\, r_{i-1}(a,a)\, Q_i + U_{i-1} \left[q_{i-1}(a,a) - r_{i-1}(a,a)\right] P_i\right\} (1 - L_{i-1}) \\ &\quad + \left[p_{i-1}(a) - U_{i-1}\right] p_{i-1}(a,a)\, P_i + \left[1 - p_{i-1}(a)\right] p_{i-1}(a,\bar{a})\, P_i \\ &= U_{i-1}\, r_{i-1}(a,a)\,(Q_i - P_i)\,(1 - L_{i-1}) + p_i(a)\, P_i. \end{aligned} \tag{3}$$

To get an expression for L_i we simply write down that the group is a nonlearned group which on the (i-3)rd trial failed to solve it in minimum, and that it did get it in minimum on the following three trials. Of course, one must take into account that having gotten it in minimum affects the probability of getting it in minimum on the succeeding trial. Thus,

$$\begin{aligned} L_i &= \left\{\left[p_{i-3}(a) - U_{i-3}\right] p_{i-3}(a,a) + \left[1 - p_{i-3}(a)\right] p_{i-3}(a,\bar{a})\right\} \cdot P_{i-2} \\ &\quad \cdot \left\{r_{i-2}(a,a)\, Q_{i-1} + \left[q_{i-2}(a,a) - r_{i-2}(a,a)\right] P_{i-1}\right\} \\ &\quad \cdot \left\{r_{i-1}(a,a)\, Q_i + \left[q_{i-1}(a,a) - r_{i-1}(a,a)\right] P_i\right\}. \end{aligned} \tag{4}$$

Observe that these three equations form a set of simultaneous, nonlinear difference equations with nonconstant coefficients. It is commonly accepted that nonlinear difference equations are more difficult than nonlinear differential equations, hence it is most unlikely that a closed solution to this system is possible. Thus we may expect to carry out numerical computations. The results of such a calculation will be presented after we discuss the evaluation of the parameters in the equations.

As has been our policy throughout, we shall assume the nodes are identical: hence we need only obtain one transfer function. We shall use in this context four sets of four conditional probabilities: one probability of each set to each of the four possible configurations N, B, S, and O.

Two of the sets are concerned with the probability of alternation from act to act within a trial (these will be distinguished by capital P's and Q's), and the other two

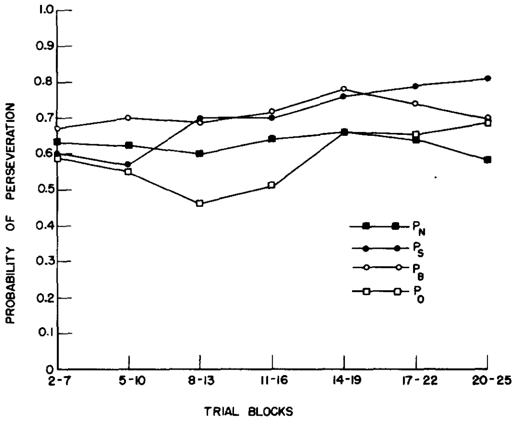


Fig. III.28

Probability of perseveration of initial-act choice for nonlearned groups when previous trial was not completed in minimum acts. Circle (x).

with initial act perseveration from trial to trial (small p's and q's will be used).

$$P_N^i, P_B^i, P_S^i, P_O^i$$

will be the conditional nodal probabilities in a nonlearned group of alternation from act to act on the i-th trial, if the previous trial was not in minimum or if it was in minimun and the present trial does not have the same initial structure. Such data are presented in Fig. III.22 for the set of all circle networks, not just the nonlearned ones. However, these cases do not exist with high frequency in the learned groups, so as an approximation to the value we want we shall use those given in Fig. III.22.

$$Q_N^i, Q_B^i, Q_S^i, Q_O^i$$

are the nodal probabilities of alternation from act to act by a member of a nonlearned group on the i-th trial if the previous trial was in minimum and the present trial has the same initial structure. Again, estimates of these values can be found in Fig. III.20 for combined learned and nonlearned groups. We shall need only Q_N and Q_S which are both practically 1; hence they may be approximated by 1 in the nonlearned groups.

If a member of a nonlearned group is in the condition N, B, S, or O on the initial act of trial i and that trial was not completed in minimum acts, then

$$p_N^i, p_B^i, p_S^i, p_O^i$$

are the probabilities that that node makes the same initial act selection on the (i+1)st trial as he did on the i-th. These values are plotted in Fig. III.28.

$$q_N^i, q_B^i, q_S^i, q_O^i$$

are the same type of probabilities when the i-th trial was completed in minimum.

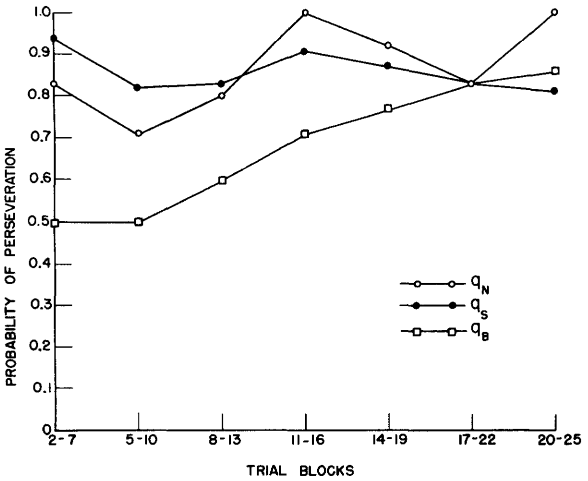


Fig. III.29

Probability of perseveration of initial-act choice for non-learned groups when previous trial was completed in minimum acts. Circle (x).

The results are given in Fig. III.29. Notice the rather large variability in these values, particularly for the N and S configurations. This is due to the fact that in any of these trial blocks the frequency of occurrence was small, of the order of 10 and 30, respectively, of which 90 percent was perseveration and 10 percent change. Clearly the probability estimates will be highly sensitive to random variations. This has led us to lump all that data and assume that these values were constant throughout the experiment.

$$q_N^i = 0.89$$

$$q_B^i = 0.69$$

$$q_S^i = 0.86$$

This assumption seems fair for N and S, but not as good for B where there appears to be a definite trend up. However, the frequency of occurrence of the B case is about 10, so we cannot be sure that this trend is not an artifact. It also appears that there may be a slight hump in the N and S curves at about trial 13, but this is even less certain than the upward trend of the B curve.

As we showed in section III.5, the probability of a trial being completed in minimum acts, given a type a initial structure, is $P_N^2 P_S^5$. Thus

$$P_i = (P_N^i)^2 (P_S^i)^5.$$

(It is by no means clear that this is correct if the previous trial was completed in minimum, and certain local variations occur in the present trial. There is a range of possibilities which, if taken into account, would lead to a problem of incredible complexity, so we have made the assumption that P_i is the correct value for this case.) Equally well

$$Q_i = (Q_N^i)^2 (Q_S^i)^5.$$

When we come to initial-act perseveration probabilities, it is perfectly clear that

$$r_i(a, a) = q_B^i q_N^i (q_S^i)^3$$

and we assumed that

$$p_i(a, \bar{a}) = 0.3125.$$

By examining all possible cases and carrying out the algebra, it is easy to show that

$$p_i(a, a) = \left[(p_S^i)^2 - p_S^i + 1\right] \left[p_B^i p_S^i + p_N^i P_S^i - p_B^i p_N^i\right] + \left[1 - p_S^i\right]^3$$

and

$$q_i(a, a) = \left[(q_S^i)^2 - q_S^i + 1\right] \left[q_B^i q_S^i + q_N^i q_S^i - q_B^i q_N^i\right] + \left[1 - q_S^i\right]^3.$$

Thus, given observed values for the nodal transfer function as characterized by

$$P_N, P_S, Q_N, Q_S$$

$$p_N, p_B, p_S, q_N, q_B, q_S$$

we may compute

$$P_i, Q_i, r_i(a, a), p_i(a, a), q_i(a, a).$$

These, with an assumed but reasonable value of $p_i(a, \bar{a})$, allow us to solve the three simultaneous difference equations, Eqs. 2, 3, and 4. Finally, the computed value of L_i from these equations allows the solution of Eq. 1 for R_i which may be compared with the observed group results.

To do this we have used observed values of some of the transfer function averaged over overlapping trial blocks. This, if it was not to cause a great deal of extra computation, required a decision as to which of two values to use for each i. Our choice has been given in Table III. 5.

In addition we have $Q_i = 1$, $q_i(a, a) = 0.6578$, and $r_i(a, a) = 0.3906$.

The results of the computation are given in Fig. III. 30. Note that the lack of smoothness is due to the assumption of parameters constant over several trials, and then a sudden change in values. Some of this detail will be explained more fully later. In that form no comparison can be made with the observed group results. If, however, we average blocks of 5 trials we obtain in Table III. 6 the observed probability of a group being in the learned category as compared with the predicted.

From the average value of U_i and the number of nonlearned groups we may obtain

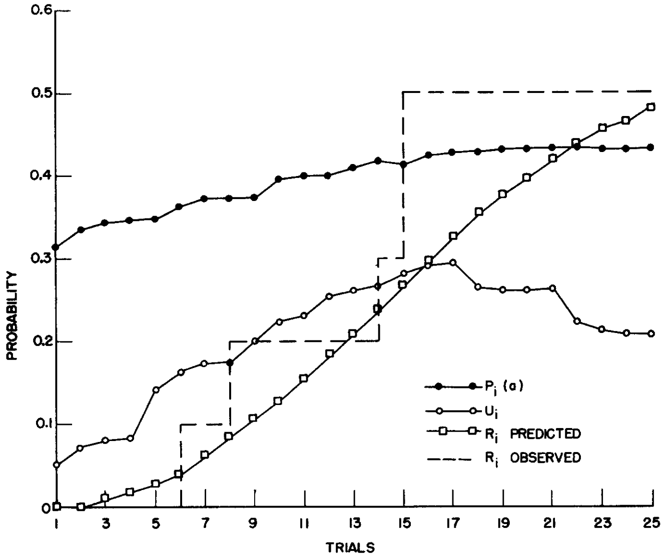


Fig. III.30
Predicted learning in the circle.

Table III.5

Trial Blocks	2 - 7	5 - 10	8 - 13	11 - 16	14 - 19	17 - 22	20 - 25
Corresponding Range for Computation	1 - 4	5 - 8	9 - 11	12 - 14	15 - 17	18 - 21	22 - 25
$p_i(a, a)$	0.3361	0.3199	0.4133	0.4151	0.4877	0.5236	0.5406
P_i	0.1648	0.3576	0.4601	0.5260	0.5818	0.4882	0.3667

Table III.6

Trials	1 - 5	6 - 10	11 - 15	16 - 20	21 - 25
R_i predicted	0.0116	0.0832	0.2110	0.3510	0.4536
R_i observed	0.0000	0.1200	0.2200	0.5000	0.5000

Table III.7

Trials	1 - 5	6 - 10	11 - 15	16 - 20	21 - 25
Predicted	4.315	8.334	10.132	6.89	5.57
Observed	3	10	13	5	4

an estimated frequency of minimum solutions among nonlearned groups. This is compared with the observed frequency in Table III. 7.

The theoretical predictions accord reasonably well with the observed values. The sources of possible error are many, as we have pointed out. In particular, the probability of a minimum solution on the (i+1)st trial, when the i-th trial was completed in minimum and the initial act structure for only some nodes is the same on the (i+1)st trial as the i-th, was assumed to be that observed when the i-th trial was not completed in minimum acts. For lack of data, $q_i(a, a)$ and $r_i(a, a)$ were assumed to be constant over trials — this is very unlikely. The recursive process tends to compound small errors in the values for early trials. The probability $p_i(a, \bar{a})$ was assumed constant, and though it probably does not deviate greatly from 0.3125, it surely does deviate somewhat.

It may be worthwhile pointing out that if certain restrictive assumptions are made, it is possible to give an explicit solution to the difference equations, Eqs. 2, 3, and 4. Assume that the parameters $p_i(a, a)$, P_i, Q_i, $q_i(a, a)$, $r_i(a, a)$, and $p_i(a, \bar{a})$ are constants independent of i, and that L_i is also independent of i.

To simplify notation we shall omit the subscript i on these quantities and we shall write $p_i(a) = p_i$. Then we have two equations of the form

$$p_i = AU_{i-1} + Bp_{i-1} + p(a, \bar{a})$$

$$U_i = CU_{i-1} + Pp_i \tag{5}$$

where

$$A = q(a, a)\,(1-L) - p(a, a)$$

$$B = p(a, a) - p(a, \bar{a})$$

$$C = r(a, a)\,(Q-P)\,(1-L). \tag{6}$$

By substituting

$$y_i = U_i - p_iP \tag{7}$$

Eq. 5 may be changed to

$$p_i = Ay_{i-1} + p_{i-1}\,(AP + B) + p(a, \bar{a})$$

$$y_i = Cy_{i-1} + p_{i-1}\,PC. \tag{8}$$

We wish to eliminate the constant term $p(a, \bar{a})$. This is effected by the change of variable

$$p_i^1 = p_i - \frac{p(a, \bar{a})}{1 - \frac{APC}{1-C} - AP - B}$$

$$y_i^1 = y_i - \frac{PC}{1-C}\left[\frac{p(a, \bar{a})}{1 - \frac{APC}{1-C} - AP - B}\right] \tag{9}$$

then

$$p_i^1 = Ay_{i-1}^1 + (AP + B)\, p_{i-1}^1 \tag{10}$$

$$y_i^1 = Cy_{i-1}^1 + PC\, p_{i-1}^1. \tag{11}$$

If we substitute Eq. 11 in Eq. 10, we obtain*

$$p_i^1 = ACy_{i-2}^1 + APCp_{i-2}^1 + (AP + B)\, p_{i-1}^1 \tag{12}$$

and if we lower the index of Eq. 10 by 1 and multiply by C, we obtain

$$Cp_{i-1}^1 = ACy_{i-2}^1 + (AP + B)\, Cp_{i-2}^1. \tag{13}$$

By subtracting Eq. 13 from Eq. 12,

$$p_i^1 = (AP + B + C)\, p_{i-1}^1 - BCp_{i-2}^1. \tag{14}$$

If we assume that

$$\frac{p_i^1}{p_{i-1}^1} = \alpha \tag{15}$$

then substitute Eq. 15 in Eq. 14

$$\alpha^2 - \alpha\,(AP + B + C) + BC = 0$$

so we obtain

$$\alpha_1,\ \alpha_2 = \frac{AP + B + C}{2} \pm \sqrt{\left(\frac{AP + B + C}{2}\right)^2 - BC}$$

and

$$p_i^1 = (C_1\, \alpha_1^i + C_2\, \alpha_2^i)\, p_1^1$$

where C_1, C_2 are constants which may be determined by two given initial conditions, one each for the equations in Eq. 5. They are of the form

$$p_1 = p^*$$

$$y_1 = y^*.$$

From the former we obtain

$$C_1\alpha_1 + C_2\alpha_2 = 1$$

---

*The following method of solution was suggested to us by Dr. Alan J. Perlis.

and from the latter

$$p_2 = Ay^* + (AP + B)p^* + p(a, \bar{a}) = \frac{p(a, \bar{a})}{1 - \frac{APC}{1-C} - AP - B}$$

$$+ (C_1 \alpha_1^2 + C_2 \alpha_2^2) \cdot \left[p^* - \frac{p(a, \bar{a})}{1 - \frac{APC}{1-C} - AP - B} \right]$$

which yields a second linear equation for C_1 and C_2. The important thing to observe is that both α_1 and α_2 lie in the interval 0 to 1, and in fact, in any examples we have considered, near 0; so we see from Eq. 9 that after a very short transient phase

$$p_i \approx \frac{p(a, \bar{a})}{1 - \frac{APC}{1-C} - AP - B}$$

$$U_i \approx \frac{p(a, \bar{a})}{1 - \frac{APC}{1-C} - AP - B} \cdot \frac{P}{1-C}$$

Thus, under these assumptions, learning is very much an equilibrium phenomenon. Now, in the approximate solution to the learning in the circle which led to Fig. III.30, we said the unevenness of the resulting curves was due to the assumption that parameters were constant over three or four trials. For the values of the parameters, it turns out that after i = 3 the transient phase is concluded; hence, the curves are composed of a series of transients. Inspection of them shows this to be approximately the case. The shape is somewhat distorted, for L_i is not strictly constant over any block of trials, only approximately so.

7. Discussion of Learning in the Chain

Since the chain is simply a circle with one symmetric link removed, it would appear that an analysis similar to that given for the circle could be carried out with, if anything, less difficulty, since the behavior of the end nodes was by experimental design completely stereotyped. This is not true. First, what is gained by the reduction to three behaving nodes is lost in the fact that it requires five acts for a minimum solution as compared with three in the circle. Second, and far more important, the structure of all possible minimum solutions is far more complex. For a given numbering of the three behaving nodes there are eight possible different communication structures (Fig. III.31). It may be shown by an exhaustive argument, which is simple but rather lengthy, that the set of all possible minimum solutions is given by the flow diagram in Fig. III.32. Thus there are 2304 possible ways of obtaining a minimum solution, yielding an equiprobable random probability of a minimum solution of 7.05 percent as compared with the value of 7.56 percent estimated by the computer. Now, unless there is a very strong added structure to this set of solutions, it is quite unreasonable to attempt to compute the probability of a minimum solution from the initial-act individual perseveration probabilities and the individual probabilities of alternation from act to

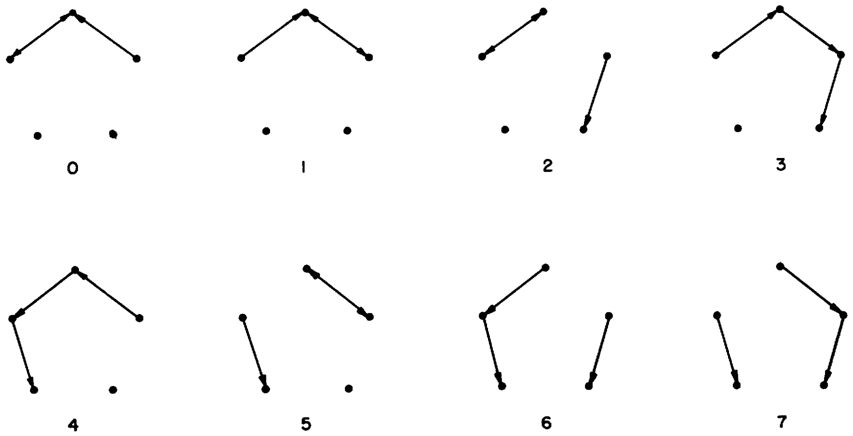


Fig. III.31

act. At the time of writing the authors have not discovered such an added structural condition which would make this feasible; hence no attempt has been made to reconstruct the group results from the individual node results.

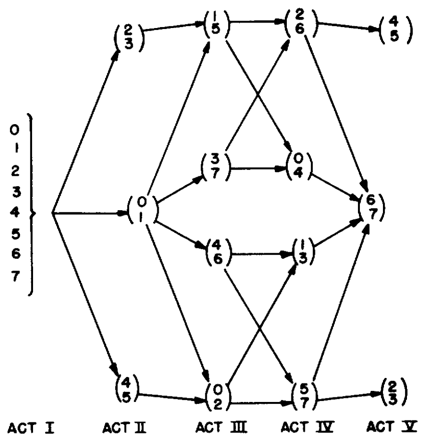


Fig. III.32

This difficulty is typical of the problems that beset the analysis of almost all the possible networks, and even such networks as the circle if nonminimum solutions are taken into consideration. To predict the group results from the individual conditional probabilities, it is necessary to have at least an estimate of the probability of a solution in k acts in terms of the individual probabilities. If such an estimate is to be exact, then a complete combinatorial analysis of the network possibilities must be undertaken, which, as we have just pointed out, is in almost all cases a tremendous undertaking. Without the development of appreciably different mathematical techniques, it is most unlikely that the exact solution technique will prove fruitful. There may be some cases in which the probability structure at the individual nodes is such that large numbers of solutions may be ignored without appreciable error, but this will only be a lucky situation. In the realm of statistical estimates there seem to be two techniques: to run experimentally groups of human subjects in each new network you wish to consider, and to use high-speed digital computers to obtain estimates of the results. The former will, of course, be necessary whenever a new area is being explored for which there is no prediction of the individual transfer functions. If, however, such knowledge is available it is certainly easier to carry out the numerical estimate than it is to run the experiments. This may be important, for it will be possible to study many more networks on a machine than it will ever be possible to study in the laboratory. If any of these show strikingly different results from the others (as, for example, the pinwheel and the circle), then it is time to return to the laboratory to see whether subjects react as predicted.

To return to the chain, since an analysis as complete as the circle is not at this time possible, we have not prepared as complete a study of the individual nodes. We have not, for example, obtained the individual-node initial-act perseveration probabilities. A gross indication of what these values must be is given in section III.8 where the perseveration of initial-act structures is discussed. On the other hand, it was computationally feasible to make a somewhat more precise analysis of the transitions from act to act within a given trial. Rather than simply estimating the alternation probability as a function of the input structure to the node in the previous act, we have considered what the person could logically know, on the basis of what he had sent in the past, as to where he would add information. We have separated the nodes into three categories—end, middle, and center—because of their topologically different relation to the rest of the network (see Fig. III.33). Furthermore, we have considered whether the previous trial was completed in minimum acts or not. The final separation is into chain (x-3) and chain (x-5); that is, whether or not the group was told that the minimum solution was three acts or five. The results are presented in Figs. III.34 through III.39, where the following notation has been used: The symbols come in pairs, identical except as to whether they are built around a Q or a P. As before, the Q shall mean that the previous trial was completed in minimum acts, and P shall mean that it was not. It will, therefore, be only necessary to give the definitions for, say, the Q symbols.

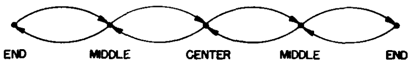


Fig. III.33

For a middle node:

$Q(C|I_c)$ is the probability that a middle node will send a message to the center node C when the message will only add information to the center node (I_c).

$Q(C|I_E)$ is the probability that a middle node will send a message to the center node when the message will only add information to the end node (I_E).

$Q(C|C, I_N + I_B)$ is the probability that a middle node will send a message to the center node when the last message he sent was to the center node, and either the present message will add no information at the end node or the center node (I_N) or it will add to both (I_B).

$Q(E|E, I_N + I_B)$ is the probability that the middle node will send a message to the end node when the last message he sent was to the end node and either the present message will add no information or it will add to both.

For the center node:

$Q(M|I_M)$ is the probability that the center node will send to the middle node M, to which the message is an addition of information when there is just one such middle node.

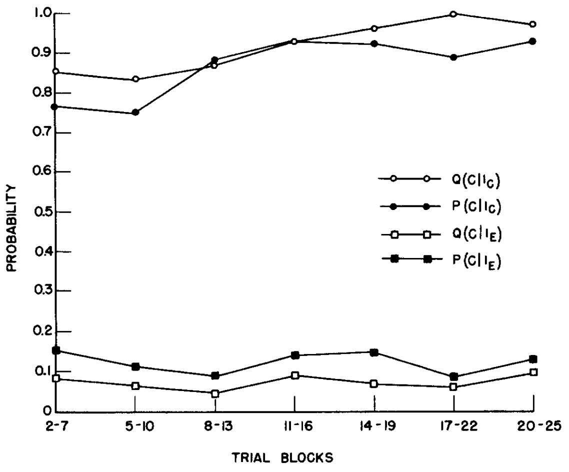


Fig. III.34
Middle node, chain (x-5), acts II-V.

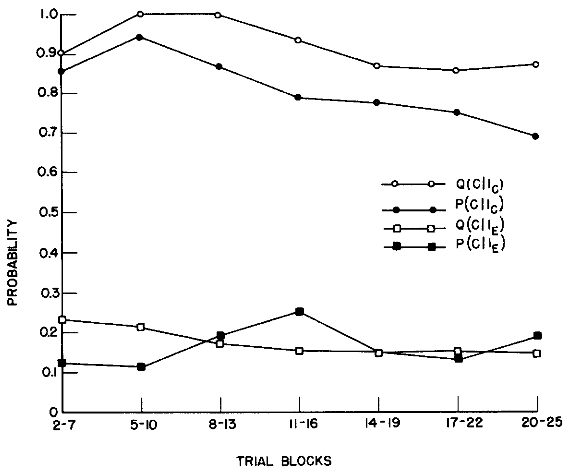


Fig. III.35
Middle node, chain (x-3), acts II-V.

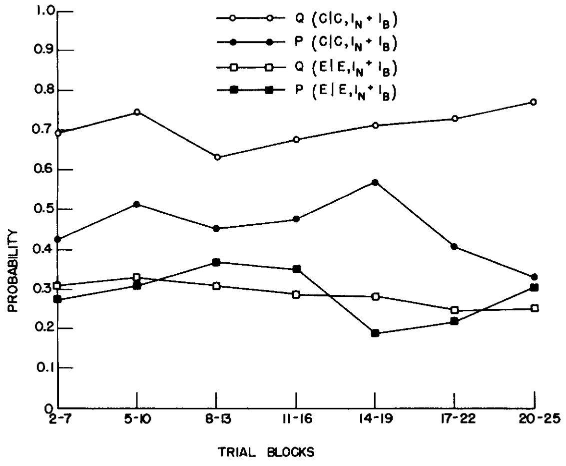


Fig. III.36
Middle node, chain (x-5), acts II-V.

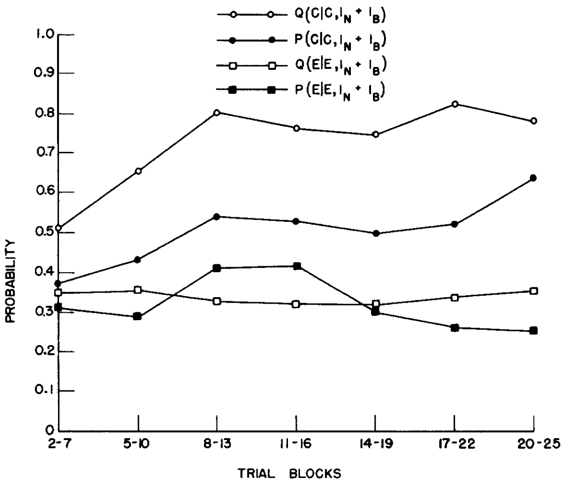


Fig. III.37
Middle node, chain (x-3), acts II-V.

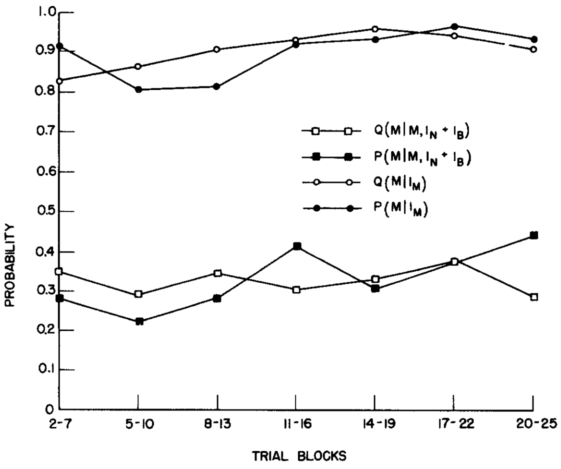


Fig. III.38
Center node, chain (x-5), acts II-V.

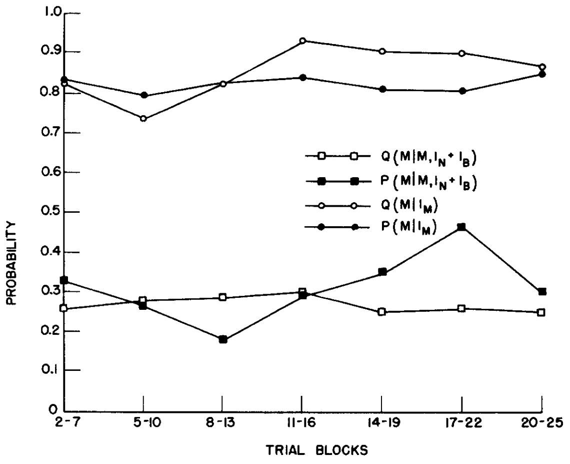


Fig. III.39
Center node, chain (x-3), acts II-V.

$Q(M|M, I_N + I_B)$ is the probability that the center node will send a message to the same middle node M as on the previous sending, when either it will add no information to either middle node (I_N), or it will add to both (I_B).

Before examining these curves in any detail, let us consider these probabilities from an a priori "locally logical" viewpoint. If we take "locally logical" to mean spreading information as rapidly as possible, it is immediately evident that

$$Q(C|I_c) = P(C|I_c) = 1$$

$$Q(C|I_E) = P(C|I_E) = 0$$

$$Q(M|I_M) = P(M|I_M) = 1.$$

Behavior in the $I_N + I_B$ case is somewhat more complex. Naively, it simply does not matter what we take these probabilities to be; however, there is a strong argument to indicate that $Q(C|C, I_N + I_B)$ should be greater than its value would be in the absence of any forces at all. For example, suppose that in act I a middle node and the center node interchange information. Then the middle node has information for both the end node and the center node; however, on act II it is much more useful to send this information to the center than to the end node, and it probably does not take very long to make this discovery. Thus, we must expect a definite tendency toward the center on act II, which will bias these probabilities which are averaged over four acts. It would be interesting to have these probabilities for each act, but this would require at least four times as much data as we now have.

With regard to the other two $I_N + I_B$ cases there do not appear to be logical reasons for them to deviate from the rather natural tendency to alternate. It is difficult to estimate how strong this is, but as we shall see there is some reason to believe it is about 0.7.

In looking at the curves corresponding to these probabilities we certainly cannot expect the logical values of 0 and 1 to occur; human fallibility and experimentation would not allow this. In addition, there is in the case of one, $Q(C| I_E)$, a possible psychological reason which would predict some deviation from purely (local) rational behavior. Suppose on act III a middle node has only information for the end node and that it does not have all the group information. It then has the choice of sending to the end node (which will do no real good for the problem does not end until each member has all the information) or to the center node. There, his message would indicate to the center node that this middle node has incomplete information, and thus suggest to the center node to send the answer to him. That is, in this situation a "locally irrational" choice may serve as a cue to the center node, and the message is not a problem message at all, but an indirect organizational message. Thus an analysis of it as a problem message is not pertinent. If such organizational messages did occur, and we have no way of knowing if this was the motivation in any of the decisions, then our experimental design was imperfect, for there were messages other than problem messages.

To analyze these curves we shall first look at the chain (x-5) case and then compare the chain (x-3) case to these. For the middle node, we see that $Q(C| I_c)$ and $P(C| I_c)$ begin near 0.8, and both increase with increasing trials, the minimum case ending very near the rational value of 1, the nonminimum case nearer 0.9. On the other hand, $Q(C |I_E)$ and $P(C| I_c)$ begin in nearly corresponding positions, but they do not tend toward zero. This may be due, as we mentioned, to using these messages in an organizational way. In both categories the separation between the curves is such that the more rational behavior follows a trial that was completed in minimum acts.

In the $I_N + I_B$ cases for the middle node, $Q(E |E, I_N + I_B)$ and $P(E |E, I_N + I_B)$ have values about 0.3 and are interlaced with no definite trend. This seems to indicate simply a preference for alternation with no appreciable distinction between the minimum and nonminimum cases, except greater variability in the latter from trial block to trial block. This may be due to a hunting or searching phenomena when the group is unsuccessful. On the other hand, the curves of $Q(C |C, I_N + I_B)$ and $P(C |C, I_N + I_B)$ show a marked separation and, in the minimum case, an appreciable tendency to repeat behavior. This is in accord with our discussion above which stated that it is important, very often, to send to the center in the $I_N + I_B$ condition. Again, the variability from trial to trial appears somewhat greater in the nonminimum case, suggesting always a searching for the correct solution.

For the center node, the values of $Q(M |I_M)$ and $P(M| I_M)$ begin at a point somewhat less than 1 and both curves tend toward 1. Furthermore, they are quite interlaced, suggesting little change in behavior from the nonminimum to the minimum case. This seems to be rational behavior. $Q(M| M, I_B + I_N)$ and $P(M| M, I_B + I_N)$ have nearly constant values of about 0.3, which accords with those found for $Q(E |E, I_N + I_B)$ and

$P(E|E, I_N + I_B)$ suggesting that this is a rough measure of the tendency to preserve behavior when there is no reason to behave one way or the other. Again, note that the curves are interlaced, very strongly suggesting that no learning from the nonminimum to minimum case occurs at the center node, and that all learning of that type is centered at the middle nodes. Of course, there is some trend for both types of nodes with trial blocks, but this is not very strong.

Let us now turn to the chain (x-3), the groups which were told that it was possible to complete the task in three acts, when in fact it was not possible. Looking at the group results (Figs. III. 3 and III. 4), we see there is no striking difference. Nonetheless, we should expect that a failure to achieve a "minimum" solution will cause variations in behavior, the searching phenomena that was mentioned above. If the curves are looked at by pairs with the chain (x-5) case, the following will be noted: For the middle node in the $I_N + I_B$ condition there is no difference that can be considered significant, except possibly that $Q(C|C, I_N + I_B)$ curve reaches its maximum value more rapidly. Evidently, any difference there is does not show up in these cases. The curves of $Q(E|E, I_N + I_B)$ and $P(E|E, I_N + I_B)$ are nearly the same, except that they are interlaced in the chain (x-3) case and the latter seems more variable. The most striking difference appears in the curves of $Q(C|I_E)$ and $P(C|I_E)$ which are well separated but, more important, have a significant trend down after an early peak. This is completely in contrast to what occurred in the chain (x-5) and what one would expect from a priori considerations.

For the center node there are no differences that are clear cut. Possibly the $P(M|M, I_N + I_B)$ curve is more variable from trial to trial, and $Q(M|I_N)$ and $P(M|I_N)$ reach lower asymptotes. There is nothing quite as striking as with the middle node, which again suggests that the predominant learning in the chain occurs at the middle nodes rather than at the center. Furthermore, the searching appears most striking when a node has only information for one other node, at least as far as he can tell. It is possible that when a chain group apparently fails to solve the problem in minimum acts, the subjects begin to doubt that the network is as simple as they had originally thought. The center and middle nodes may come to a belief that there are paths between the end nodes which do not pass through them. It is possible that such doubts will lead to the trends that have been observed in the chain (x-3) case. However, this is pure speculation.

In summary, it cannot be said that the individual and group results for these two conditions are strikingly different, but there are some differences. The weakness of the observed phenomena may be due to insufficient motivation to achieve the solution to the task; it can hardly be argued that 3 bell rings is a markedly stronger motivation than 5. An experiment designed to explore the resulting frustration of this type of situation with high motivation and its effects on group behavior should be of considerable interest.

8. Perseveration of Initial-Act Structures

No analysis of the type applied to the circle and the chain has been carried out for the other networks studied experimentally. This is primarily due to the greater richness of possible cases stemming from the greater complexity of these networks. If we attempt to examine the conditional probabilities of choices on the i-th act dependent on the (i-1)st act, there are so many possibilities that with but 10 experimental groups our estimates of these probabilities will be very poor.

However, the other phase of learning, the perseveration of the initial-act structure when the previous trial has been completed in minimum can be readily obtained, at least in those cases when a significant fraction of the trials was completed in minimum acts. It is not possible, as it was in the case of the circle, to make this subdivision finer to obtain the probability of perseveration for a particular node in a particular configuration on the previous act. Again this is due to the numerous cases that have to be considered, and the fact that in many cases different nodes cannot be lumped together because of their different topological relation to the rest of the network.

The initial-act perseverations are given in Table III.8 for each of the networks, broken down in trials 2 to 7, 8 to 13, 14 to 19, and 20 to 25. A four-fold table has been formed for each of the cases. The columns indicate whether the previous trial was completed in minimum (denoted Min) or greater than minimum (> Min), and the rows indicate whether the same initial-act structure was obtained, that is, zero change (denoted > 0), or a different one (denoted < 0).

The primary conclusion that can be drawn from these tables is that the initial-act perseveration tends to correlate with the learning as shown in the acts-to-completion distribution. Both the pinwheel and the totally connected networks show no learning in the sense that if the previous trial was completed in minimum, there is no strong tendency for the same initial-act structure to be selected. This must be tempered by the fact that in both situations there are very few cases of a trial completed in minimum acts, so the statistics cannot be considered very stable.

9. Remarks on Geometrical Effects

We shall be able to say very little about the geometrical effects of different representations of a given network. The only cases that were examined experimentally were the circle (x), circle (0), chain (x), and chain (0). As far as nodes with any choice are concerned, all the x cases appeared physically as shown in Fig. III.40a, and the 0 cases appeared as shown in Fig. III.40b. From an a priori point of view, a right-handed person should have a tendency to select slots on the left, as the motion out and across the body is more natural than one to the right. This should be stronger in the 0 cases than in the x, for the extreme right slot is somewhat awkward to reach. The argument is exactly the opposite for left-handed people, but since there are on the average considerably fewer left-handed people than right, the predominant tendency, if any, should

Table III. 8a

					Network				
Trials		Chain (x-3)			Chain (x-5)			Chain (0)	
		Min	> Min		Min	> Min		Min	> Min
2 - 7	0	0.103	0.121	0	0.125	0.125	0	0.233	0.100
	> 0	0.293	0.483	> 0	0.286	0.464	> 0	0.133	0.533
		Min	> Min		Min	> Min		Min	> Min
8 - 13	0	0.281	0.088	0	0.207	0.138	0	0.333	0.037
	> 0	0.281	0.351	> 0	0.276	0.379	> 0	0.444	0.185
		Min	> Min		Min	> Min		Min	> Min
14 - 19	0	0.315	0.056	0	0.356	0.102	0	0.448	0.069
	> 0	0.389	0.241	> 0	0.373	0.169	> 0	0.345	0.138
		Min	> Min		Min	> Min		Min	> Min
20 - 25	0	0.423	0.038	0	0.379	0.086	0	0.767	0.033
	> 0	0.385	0.154	> 0	0.345	0.190	> 0	0.133	0.067

Table III. 8b

	Network								
Trials		Circle (x)			Circle (0)			Pinwheel	
		Min	> Min		Min	> Min		Min	> Min
2 - 7	0	0.033	0.050	0	0.100	0.200	0	0.000	0.100
	> 0	0.067	0.850	> 0	0.133	0.567	> 0	0.050	0.850
		Min	> Min		Min	> Min		Min	> Min
8 - 13	0	0.117	0.033	0	0.200	0.100	0	0.017	0.067
	> 0	0.167	0.683	> 0	0.167	0.533	> 0	0.067	0.850
		Min	> Min		Min	> Min		Min	> Min
14 - 19	0	0.317	0.133	0	0.233	0.067	0	0.017	0.133
	> 0	0.167	0.383	> 0	0.067	0.633	> 0	0.033	0.817
		Min	> Min		Min	> Min		Min	> Min
20 - 25	0	0.390	0.136	0	0.233	0.200	0	0.000	0.067
	> 0	0.102	0.373	> 0	0.067	0.500	> 0	0.067	0.867

Table III.8c

		Network											
Trials		Barred Circle			Wheel			Totally Connected			Alpha		
		Min	> Min		Min	> Min		Min	> Min		Min	> Min	
2 - 7	0	0.017	0.067	0	0.000	0.000	0	0.017	0.033	0	0.083	0.033	
	> 0	0.050	0.867	> 0	0.117	0.883	> 0	0.100	0.850	> 0	0.083	0.800	
		Min	> Min		Min	> Min		Min	> Min		Min	> Min	
8 - 13	0	0.067	0.067	0	0.000	0.033	0	0.000	0.000	0	0.150	0.083	
	> 0	0.117	0.750	> 0	0.167	0.800	> 0	0.167	0.833	> 0	0.050	0.717	
		Min	> Min		Min	> Min		Min	> Min		Min	> Min	
14 - 19	0	0.133	0.100	0	0.068	0.068	0	0.050	0.000	0	0.200	0.067	
	> 0	0.083	0.683	> 0	0.203	0.661	> 0	0.167	0.783	> 0	0.033	0.700	
		Min	> Min		Min	> Min		Min	> Min		Min	> Min	
20 - 25	0	0.267	0.100	0	0.100	0.100	0	0.033	0.017	0	0.367	0.117	
	> 0	0.067	0.567	> 0	0.167	0.633	> 0	0.117	0.833	> 0	0.017	0.500	

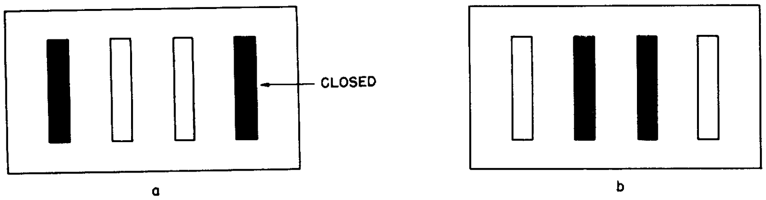


Fig. III.40

be to the left.

Any such effect after the first act will be thoroughly confused by the effects of the information interchange on the first act, so we shall consider only data obtained in the first act. However, since there is a considerable perseveration tendency in the first act, the effect may also be destroyed in that data. In the case of the circle, we have shown there is a marked increase in type a structures. In any type a structure there must be two nodes which send messages through the left open channel, and two to the right; it does not matter what the fifth one does (see Fig. III.41). So, even if there is a strong tendency to repeat behavior, we should expect any preference for the left to show up. However, the number of samples is effectively reduced from 50 in circle (x) and 25 in circle (0) to 10 and 5, respectively, which of course are very small samples. The obtained frequencies for 50 and 25 people, respectively (for the type a structure does not always appear), are given in Fig. III.42. We see that there seems to be no preference.

Fig. III.41

If we turn to the chain there is, even in act I, only one person who is in a topologically symmetric position: the center node. The observed frequencies for the center node of the chain (x-3) and chain (x-5) are given in Fig. III.43. Keeping in mind again that this is a small sample, there appears to be no preference. The case of chain (0) is a sample of five people, and as we can see from Fig. III.44, there seems to be a strong preference for the left of the order of 0.8. However, taking as a population of 10 the center nodes of the chain (x-5), it is possible to select a subpopulation of 5 which yields the second curve in Fig. III.44. This curve is indistinguishable from the chain (0) curve; hence, we cannot say whether there is in fact a tendency to the left, or whether what is observed is simply a sampling difficulty. But the fact that we may select 5 chains (x-5) having a high tendency to the left, and the chain (x-5) curve in Fig. III.43 imply the other 5 have a tendency to the right. Thus it appears that the perseveration effect from trial to trial rather than a left tendency is being observed.

We may make a tentative hypothesis as to the geometrical effect of a particular representation of a network: The effect on the average is slight if it exists at all; however, it may be that it is quite strong in particular individuals. Our evidence is not clear for it does not distinguish between perseveration and choice tendencies. It is an

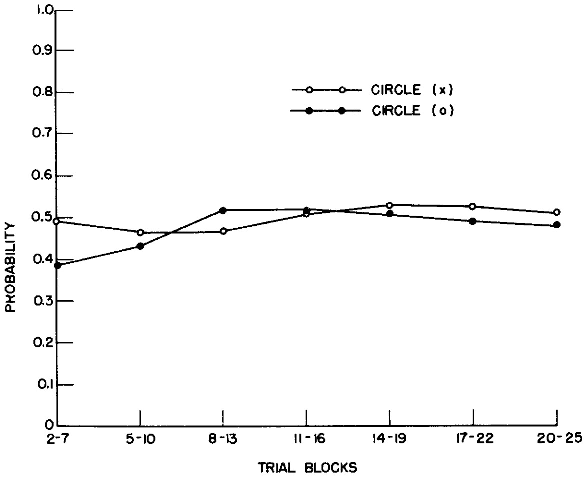


Fig. III. 42
Probability of sending message to left, act I, circle.

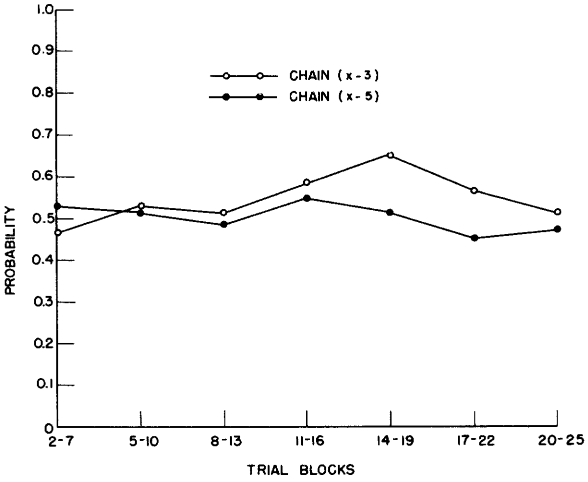


Fig. III. 43
Probability of center node sending message to left, act I, chain .

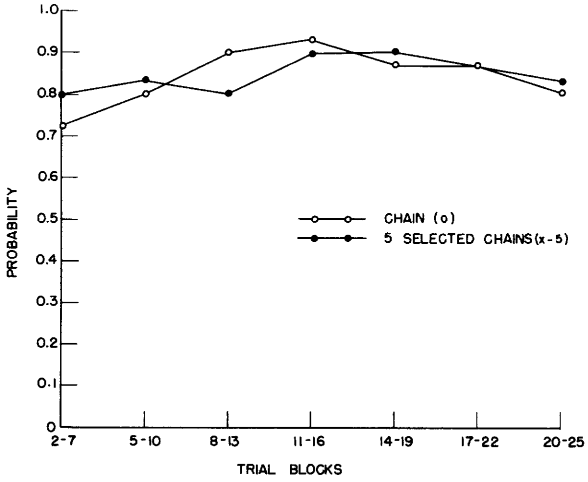


Fig. III.44
Probability of center node sending message to left, act I, chain.

effect which may influence learning, but it is probably negligible once the learning is under way, and decisions are being based more on the local information state and past behavior. If a network could be found in which a left tendency would prevent or hinder learning and another in which it would augment it, then differences might be observed. If this is felt to be an important aspect of the group process, then experiments will have to be explicitly designed to detect it.

10. Remarks on Complexity and Size

Earlier we pointed out that group learning seems to be correlated in some way with an undefined notion of complexity, but we suggested that it is most unlikely that there exists a single parameter of the network to characterize this. We are in a position now to elaborate these thoughts somewhat, but we shall not be able to say anything definitive, except that this is a complex problem.

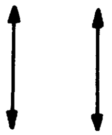


Fig. III.45

First of all, size is one aspect of complexity. It seems reasonable to say that a four-node circle is simpler than a five-node circle. But it is far from clear just what this means with respect to learning. For example, if we carry out an analysis of the four-node circle similar to that carried out earlier for the five-node case, it is easy to show that the minimum solution is two acts and that there is one class of structures, shown in Fig. III.45, having two members, which are the initial structures for all minimum solutions. We shall call it type a. Then

$$q(a, a) = q_S^4 + (1 - q_S)^4$$

$$r(a, a) = q_S^4$$

$$p(a, a) = p_S^4 + (1 - P_S)^4$$

$$p(a, \bar{a}) = \frac{2}{7} \Big\{(1 - p_B)^2 p_N^2 + p_B^2 (1 - p_N)^2 = 2 p_N p_S p_B (1 - p_B)$$

$$+ 2(1 - p_N)(1 - p_S)(1 - p_B) p_O + p_O^2 (1 - p_O)^2 \Big\}$$

$$P = P_S^4$$

Using the same values for the transfer function as we obtained experimentally in the five-node case (we shall discuss this assumption) and carrying out the analysis for 15 trials, we obtain Table III. 9.

Table III. 9

	Trial Block		
	1 - 5	6 - 10	11 - 15
4 Node R_i	0.0114	0.0839	0.1641
5 Node R_i	0.0116	0.0832	0.2110

Thus, if the conditional probabilities remain the same, the "simpler" four-node case does not learn to obtain minimum solutions as well as the five-node case does. However if we compare three-act solutions in the two cases, the four-node cases will do considerably better. Furthermore, if actual experiments were run, it seems likely that percentage-wise the four-node case will do better on minimum solutions than the five-node case, because the subjects will rapidly see through the network and see that a minimum solution is given by Fig. III. 46.

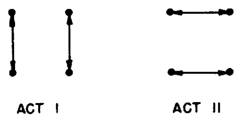


Fig. III. 46

However, if a similar phenomenon occurred between, say six- and seven-node networks, it is very unlikely that subject perception of the situation will differ appreciably, and the assumption of the same conditional probabilities would be good. In that case, should we say there is more or less learning? How much?

If we return to situations having a fixed number of nodes, and attempt to relate learning to complexity, we have shown that there are at least two aspects of complexity that must be considered. In the pinwheel network, the aspect of the situation which apparently prevented learning was the complexity in the structure of all possible solutions relative to the individual information state. The topology of the network itself is similar in many respects to the circle, and as to memory of past decisions, exactly the same, that is, one bit for each decision. However, in any minimum solution there arise decisions for which the local information state is not adequate to prescribe rationally what the subject should do.

A second aspect of complexity is illustrated in, say, the totally connected network. Here there are, for example, all the possible minimum solutions of the circle, and many more, but each decision a person makes involves two bits of information, or a total of 10 per act for the group. The memory problem of previous successful trials is twice as great as in the circle or the pinwheel, and as in the pinwheel, many of the informational cues are lacking that exist in the circle. It is unlikely that the recall is only a half of what it was in the circle, but it is probably somewhat less. This factor is compounded five times; hence the probability of a correct recollection of a previous minimum solution will be considerably smaller than in the circle. It does not seem possible at present to carry out a precise analysis for the totally connected network, since we do not know the structure of minimum solutions and there is not enough data to estimate all the possible conditional probabilities, but it does seem clear that the simple memory problem is a primary deterrent to learning in this network. These same effects appear to be present in the barred circle, alpha, and wheel, in that order. The number of bits to be remembered per act are 6, 6, and 7, respectively.

This argument could be made more substantial if some work were done on individual memory of the sequence of decisions, with and without a structure to the sequence, the results being presented in terms of recall as a function of the number of bits in a decision. In all likelihood it will be important to distinguish recall in various parts of the sequence.

11. Summary

The observed distributions of acts to completion indicated that for some networks considerable changes over trials occurred, and for others fewer, and for the pinwheel practically none. This was attributed to learning. In addition, it was shown that these distributions are considerably different from those that would have been obtained from groups of subjects who threw a die to decide where to send. Having established this learning, the next step was to account for it in terms of the learning of the individuals.

Learning in the circle and the pinwheel were discussed in some detail, and the effect of their topological differences on the decisions of the subjects were pointed out. It was shown, at least verbally, that combining the apparently locally rational behavior of the subjects with the topological properties of the network leads to the wide differences between these cases that were actually observed.

It was possible, by a more careful mathematical examination of the circle and an appropriate definition of learning for that case, to account for the observed group results in terms of 10 conditional probabilities for the nodes which were estimated from the observed data. A number of approximations were made that can be eliminated, if a more precise mathematical analysis is made, and if more data are obtained in order to yield more stable frequencies.

We then turned to an analysis of the chain. It was shown that there is a considerable mathematical difficulty in this case, and in almost all others, which, at least at present, makes it impossible to carry out an analysis analogous to that given for the circle. The conditional probabilities of the nodal transfer function were presented and examined in some detail.

In the last three sections rather incomplete remarks were made on several other topics. It was shown that perseveration of initial-act structure occurs in all networks that learned, as was to be expected from the analysis of the circle and the chain. The discussion of the effects of the size of a network on the obtained results and of the geometric representation of the topology of the network were both inconclusive.

CHAPTER IV – ACTION TIME

1. Experimental Design and Time Measurement

In the early exploratory phase of our experimental program (Experiments 1, 2, and 3), the times when the subjects sent messages were unconstrained; that is, the subjects chose freely not only the content but also the occasion of their messages. The only times that were recorded in these situations were the total times taken by the group to reach a solution in each trial. Because the message content was free, the time recorded included time spent in the processes of constructing, sending, and interpreting messages not directly concerned with reaching the solution for a specific trial. For example, there were organizational messages relevant to the improvement of the efficiency of group action over a series of trials and various more or less irrelevant messages expressing attitudes toward different features of the group situation and the group performance. These latter two categories of messages contain many features of intrinsic interest with respect to the investigation of the interrelation of individual personality and group action, but from the standpoint of the consideration of the temporal relations in group action, they serve mainly to becloud the picture. It is not that such processes do not occur in nonlaboratory groups and are therefore unworthy of attention (quite the contrary!), rather it is that the sources of variability in solution time under these conditions are so multifarious as, for the present, to defy detailed explanation. The hand-in-hand consideration of both the treatment of data and the construction of theory force upon us a policy of "one thing at a time."

Our more recent experiments have been run with the message content restricted and the group action quantized. For time data obtained under these conditions, we have developed a theory of group-action time which provides a well-fitting explanation of the data. The exposition and application of this theory forms the major portion of the present chapter. We have developed nothing closely approaching a satisfactory theory for free-content, nonquantized cases. In section 8 of this chapter an attempt to reduce the data of one nonquantized case to an equivalent quantized case is presented; this can be accepted only as an approach to the data when not enough was recorded, and when a wholly adequate theory is not available. In section 9 the outline of a possible theory for the nonquantized case is given; however, as it stands, the mathematics is so complex that it is very unlikely that an explicit solution can be expected for more than two nodes. Whether such simple solutions can be pieced together for cases of greater complexity remains to be seen. The work in Appendix 3 bears on this problem and does indeed solve the n node communication problem for a certain narrowly delimited class of group processes. For this theory to apply, the message flow must be restricted to the transmission of one single content which may be repeated over and over. The price of such a stringent restriction is to leave our problem essentially unsolved.

2. Quantized Action

We shall restrict our attention, for a while, to groups which are action-quantized. Let τ_i be the time at act i for the group to signal the environment that it is ready to take action, and let δ_i be the time required by the environment, following act i, to feedback a signal to the group to take action. If we assume that the task is completed in a finite number of steps k, then the time to complete the task is obviously

$$\sum_{i=1}^{k} \tau_i + \sum_{i=1}^{k} \delta_i .$$

We shall assume that the completion of the task is dependent only on the actions taken by the group and is independent of the time τ_i. Thus, we may treat the problem of the time distribution to complete the task as being composed, in an appropriate fashion, of the independent distribution of the number of acts to complete the task, the distributions τ_i of the time for the i-th act, and the distributions of the times δ_i for the response from the environment. Without a rather complete specification of the environment and the group, it seems impossible to discuss the distributions δ_i, so we shall not consider this problem here. It will suffice, for our purposes, to note that in Experiments 4 and 5, $\delta_i = 2$ sec. The previous chapter presented a discussion of the act distribution for the experiments in which action-quantization occurred. We shall, in this chapter, assume that the act distribution is known.

Our aim is two-fold: to obtain a reasonable form for the individual time distributions, and to show how these are composed to give the empirical group distributions τ_i. The differences in time per act which arise among the performances of different persons at the nodes of the various networks in our experiments may be attributed to two classes of causes: (a) pre-experimental differences in the individuals themselves; (b) differences in the relation of the node the individual occupies to the network of which it is a part. We are not studying the first class, and we have attempted to eliminate the effects of individual differences by randomization. The study of the analytic form of the individual time distribution is the topic of the next section. The composition of these distributions to yield the group distribution in terms of the given network is the topic of sections IV.4 and IV.5.

Before turning to these problems, we shall introduce some notation which will be used throughout the chapter. On the i-th act and for the j-th person in a group it is possible, in principle, to determine a probability distribution

$${}_i f_j \, (t - t_{i-1})$$

where

$$t_{i-1} = \sum_{\sigma=1}^{i-1} \tau_\sigma + \sum_{\sigma=1}^{i-1} \delta_\sigma$$

of the time for man j to signal the environment. It is obvious that

$$ {}_{i}f_{j}\ (t) = 0 \quad \text{for } t \leqslant 0 $$

$$ \int_{-\infty}^{\infty} {}_{i}f_{j}\ (t)\ dt = 1. $$

In general, ${}_{i}f_{j}\ (t)$ will be a function of the task, the network, and the boundary conditions. In fact, in some situations, it may also be a function of the particular sequence of acts which has led the group to where it is. In order to simplify notation, we shall not make these dependencies explicit. In addition, we shall omit the subscript i, assuming the discussion is for a specific act, and replace the variable $t - t_i$ by simply t. Thus

$$ f_j\ (t) $$

will be the distribution of the j-th man, and we shall always write the cumulative distribution as

$$ F_j\ (t) = \int_{-\infty}^{t} f_j\ (x)\ dx. $$

3. The Form of the Individual Time Distributions*

The actions demanded of our subjects require that they apprehend the information available to them before they take the action of sending a message and (in most cases) that they make a choice of the destination of their message. Under the condition that the information available (the relevant stimulus situation) is a steady state, the form of the action time distribution has been shown on both theoretical and empirical grounds to be well described by an exponential decay function (18, 43). For two reasons, we cannot assume that the stimulus situation is a steady state from the start of an act until a message is sent. First, the signal which initiates an act is a signal to send messages, whereas the relevant stimulus situation exists only when the messages have been received. This results in a short and fairly constant time lag. Second, the message content is not instantaneously effective as a stimulus, but builds up to a steady state of effective stimulation over a short period of time. For these reasons we must give a somewhat more general derivation of the form of the action time distribution. Rather than assuming the strength of the message-sending response tendency has a constant proportionality to the time interval during which it acts for short time intervals, we must assume that the proportionality is a function of time. Thus we will not write $\lambda\ \Delta t$ but $\lambda(t)\ \Delta t$.

---

*The following treatment is formally the same as that given for the discharge of a condenser.

Let us suppose that if the organism is stimulated at time 0 and does not respond in the interval from 0 to t, then the probability of a response in the interval $(t, t + \Delta t)$ is

$$\lambda(t)\ \Delta t. \tag{1}$$

Let $Q(t)$ be the probability that no response has occurred in the interval from 0 to t, then the probability of a response, $f(t)\ \Delta t$, in the interval t to $t + \Delta t$ is

$$\lambda(t)\ Q(t)\ \Delta t.$$

But clearly

$$Q(t + \Delta t) = Q(t) \left[1 - \lambda(t)\ \Delta t\right]$$

or

$$\frac{Q(t + \Delta t) - Q(t)}{\Delta t} = - \lambda(t)\ Q(t)$$

so

$$\frac{dQ(t)}{dt} = - \lambda(t)\ Q(t)$$

or

$$\ln Q(t) - \ln Q(0) = - \int_0^t \lambda(t)\ dt.$$

But $Q(0) = 1$ for any real situation, so

$$Q(t) = \exp\left[- \int_0^t \lambda(t)\ dt\right].$$

Thus the frequency distribution of a response occurring at time t, subject to the condition that no response has previously occurred, is

$$f(t) = \lambda(t) \exp\left[- \int_0^t \lambda(t)\ dt\right]. \tag{2}$$

In our work it will be appropriate to think of functions $\lambda(t)$ which are displaced so as to have an origin at $t_o > 0$. t_o is the least dead time between the time of the signal to begin a message and the actual time it is sent. In this case we shall have expressions of the form

$$\begin{aligned} f(t - t_o) &= \lambda(t - t_o) \exp\left[- \int_0^{t-t_o} \lambda(t)\ dt\right] \\ &= \lambda(t - t_o) \exp\left[- \int_{t_o}^{t} \lambda(t - t_o)\ dt\right], \quad \text{for } t > t_o \\ &= 0, \quad \text{for } t \leqslant t_o. \end{aligned} \tag{3}$$

Integrating by parts it is easily shown that

$$\int_{-\infty}^{\infty} f(t)\ dt = 1 - \exp\left[-\int_{0}^{\infty} \lambda(t)\ dt\right].$$

Thus f(t) is a probability density if, and only if

$$\int_{0}^{\infty} \lambda(t)\ dt = \infty$$

otherwise there is a finite probability that no decision will be reached.

The utility of the distribution Eq. 2 or Eq. 3 depends on what function we take $\lambda(t - t_o)$ to be. Since we are using it to characterize the build-up of the message sending tendency as determined by the stimulus information, it is reasonable to suppose that this process occurs rapidly compared to the total mean times taken for an act, as in Fig. IV.1. This will lead to a distribution of the kind shown in Fig. IV.2. We may approximate $\lambda(t)$ by a step function beginning slightly after t_o (the amount depending on the steepness of rise); see Fig. IV.3. In this case the distribution is an exponential decay curve as shown in Fig. IV.4. This approximation will be good if the slope of $\lambda(t - t_o)$ is very sharp near $t = t_o$. Moreover, for the use we shall make of Eq. 3 the goodness of fit of the rising limb is not at all critical. To modify Eq. 3 for the approximation, we take a new t_o which includes the dead time plus the initial portion of the rise time of $\lambda(t - t_o)$ during which $\lambda(t - t_o)$ is nearly zero and from $t = t_o$ (the new t_o) on we take $\lambda(t - t_o) = \lambda$ (a constant). Thus

$$f(t)\ dt = \lambda e^{-\lambda(t-t_o)}\ dt. \tag{4}$$

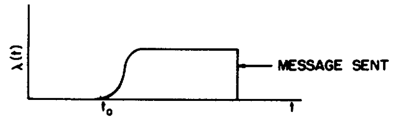


Fig. IV.1

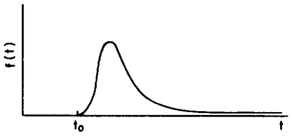


Fig. IV.2

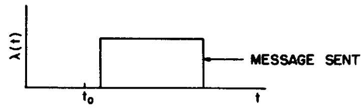


Fig. IV.3

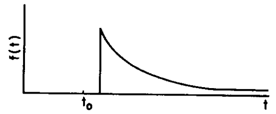


Fig. IV.4

4. Theory of Group Act Time

Since an act is completed only when each member of the group has signaled the environment, the time distribution for the act, $\tau(t)$, is given by the distribution of the largest value of time when one and only one value is selected from each of the $f_j(t)$, $j = 1, 2, \ldots, n$. The probability that the largest value is between t and $t + \Delta t$ and that it is selected by man k is

$$f_k(t)\ \Delta t \prod_{\substack{j=1\\ j\neq k}}^{n} F_j(t).$$

Thus the probability that the group selects a largest value between t and $t + \Delta t$ is given by

$$\tau(t)\ \Delta t = \sum_{k=1}^{n} f_k(t) \prod_{\substack{j=1\\ j\neq k}}^{n} F_j(t)\ \Delta t = \tau(f_1, f_2, \ldots, f_n)\ \Delta t.$$

Observe that by carrying out the indicated differentiation we have

$$\tau(f_1, f_2, \ldots, f_n) = \sum_{k=1}^{n} f_k(t) \prod_{\substack{j=1\\ j\neq k}}^{n} F_j(t) = \frac{d}{dt}\left[\prod_{j=1}^{n} F_j(t)\right]. \tag{5}$$

Thus $\tau(t)$ is a distribution, for

$$\int_{-\infty}^{\infty} \tau(f_1, f_2, \ldots, f_n)\ dt = \int_{-\infty}^{\infty} \frac{d}{dt}\left[\prod_{j=1}^{n} F_j(t)\right] dt = \prod_{j=1}^{n} F_j(\infty) - \prod_{j=1}^{n} F(-\infty) = 1.$$

In the case that all the $f_j(t)$ are equal, say, to $f(t)$, Eq. 5 becomes

$$\tau(t) = nf(t)\left[F(t)\right]^{n-1} = \frac{d}{dt}\left[F(t)^n\right] \tag{6}$$

which is the well-known distribution of the largest of n selections from a given distribution $f(t)$ (see ref. 85, vol. I, p. 218).

Though our intermediate argument requires Eq. 5, actual computations will be carried out only for Eq. 6. Thus we shall need moments of the distribution of the largest of n selections from a distribution of the form

$$f(t) = \lambda e^{-\lambda t}$$

that is

$$\mu_i'(n) = n \int_{-\infty}^{\infty} t^i \lambda e^{-\lambda t} \left(1 - e^{-\lambda t}\right)^{n-1} dt$$

$$= n \int_{-\infty}^{\infty} t^i \lambda e^{-\lambda t} \sum_{j=0}^{n-1} (-1)^j \binom{n-1}{j} e^{-j\lambda t} dt$$

$$= \frac{n}{\lambda^i} \sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j \frac{i!}{(j+1)^{i+1}} = \frac{g_i(n)}{\lambda^i} \qquad (7)$$

This series has been evaluated for several cases that are of interest to us in Table IV. 1

Table IV. 1

Values of $g_i(n)$

n	$\mu_1'\lambda$	$\mu_2'\lambda^2$	$\sigma\lambda=\sigma'$	$\mu_3'\lambda^3$	$\mu_4'\lambda^4$	$\mu_5'\lambda^5$
2	1.50000	3.50000	1.11803			
3	1.83333	4.72222	1.16667			
4	2.08333	5.76389	1.19316	20.29514	88.09142	458.30352
5	2.28333	6.67722	1.20980			

5. Reduction of Number of Parameters

Let us consider Eq. 5 in relation to the data we have. If we suppose that each of the $f_j(t)$ is different and each is characterized by only one constant, then, since the location of the origin of each of the distributions is arbitrary, it is clear that $\tau(f_1, f_2, \ldots, f_n)$ is dependent on a minimum of 2n parameters, unless certain simplifying assumptions are made. The values of these parameters must be determined from empirical data. If we can obtain data about the individual f's directly, then we can readily accept this number of parameters, since it requires only the first two moments of the time distribution for the individuals. If, however, there is only data about $\tau(f_1, f_2, \ldots, f_n)$, as is the case at present, the first 2n (= 10 in our case) moments are needed to determine these. Not only is this impractical from a computational point of view, but the inefficiency of estimation from high order moments and the moderate sample size do not permit us to use more than the first two moments with any reasonable degree of confidence in the reliability of our result. We must, therefore, find simplifying assumptions which reduce the number of parameters to two. Since each individual distribution is dependent on two parameters, we must find situations in which all the distributions can be taken to be the same, or situations in which some of the individual distributions can be treated as the same, and the remainder neglected without causing a large error.

Consider first the circle network: In act I each of the people in this network are in essentially equivalent situations; each man receives one piece of input data and must make a decision between two choices. Thus, if we assume they are statistically identical people, it is reasonable to suppose that each person has the same distribution f(t). On act II, information has been transferred and some of the people (those who received information) have a new decision problem as to where to send it. However, it is very unlikely that each of the people will have received a new piece of information; in fact, if the sending is equiprobable over the two links, the chance is 1/16. The equiprobable chance that one person receives no inputs and the others at least one input is 5/8. However, as we have seen in the previous chapter (see Fig. III-26), the probability of selection of a structure of type a (Fig. IV-5) increases appreciably above its chance value of 5/16, hence increasing the probability of exactly one node failing to receive a new piece of information. Thus, it is reasonable to suppose that four of the nodes are operating on a distribution

$$f(t) = \lambda e^{-\lambda(t-t_1)}$$

and one node on

$$f_o(t) = \lambda_o e^{-\lambda_o(t-t_o)}$$

and $\lambda_o > \lambda$ and $t_o < t_1$. By act III it is very probable that each of the nodes has received at least one piece of information in addition to its own, so that each has a decision problem. Thus we return to the approximation by five identical nodes. It is clear that the assumption of statistically identical situations in the third act is poor, for some of the nodes will have considerably more information than the others; hence, their decision problem will be more difficult and will tend to cause a greater spread in the distribution.

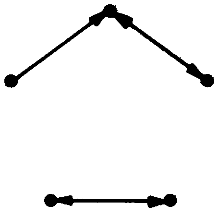


Fig. IV-5

The argument for the pinwheel is essentially the same, except that in act II we do not have as great confidence in assuming one node rather than two has received no additional information. This follows from the results of Chapter III which showed a notable lack of learning in the pinwheel.

For the chain there are in all acts two nodes (the end ones) which have no decision problem, and there is an exceedingly high probability after only a short period of learning that each of the other three will receive additional information (at least for the first two acts). It thus seems reasonable to suppose that three of the nodes have a distribution

$$f(t) = \lambda e^{-\lambda(t-t_1)}$$

and two

$$f_o(t) = \lambda_o e^{-\lambda_o(t-t_o)}$$

with $\lambda_o > \lambda$ and $t_o < t_1$.

The same argument applies to act I of the totally connected network as applies to act I of the circle, with possibly a different distribution since, in this case, the decision is one out of four links. For act II there is a pure chance of 125/128 of having at least one node which has received no information, and a probability of 19/64 that at least two nodes have received no added information. The assumptions to be made here are less clear than those which we made for the circle, but it appears that the assumption of four identical nodes is appropriate. Again in act III we shall assume five identical nodes.

The remainder of the networks studied in Experiment 4 do not have such simple structures, and so the argument is both less easy and less convincing. We have therefore decided simply to present the data for the four cases mentioned. Table IV.2 summarizes our assumptions.

Table IV.2

Number of Statistically Identical Nodes Having
Larger Standard Deviation and Mean

Network	Act I	Act II	Act III
Circle	5	4	5
Pinwheel	5	4	5
Chain	3	3	3
Totally Connected	5	4	5

In general, then, we have a group with n nodes of which k has a distribution f and n-k a distribution f_o. It is reasonable to suppose that when k is greater or equal to n-k, λ_o is sufficiently greater than λ, and t_1 is greater than t_o, the error in replacing $\tau(f_o, \ldots, f_o, f, \ldots, f)$ by $\tau(f, \ldots, f)$ is small relative to some characteristic size of f, say, relative to the maximum value of f, f_{max}. We make this more precise.

Define

$$\Delta(t,k) = \tau(f_o, \ldots, f_o, f, \ldots, f) - \tau(f, \ldots, f) \qquad (8)$$

then

$$\frac{\Delta(t,k)}{f_{max}} = F^{k-1} \left\{ k \left[F_o^{n-k} - 1 \right] \frac{f}{f_{max}} + (n-k)\, F\, F_o^{n-k-1} \frac{f_o}{f_{max}} \right\}. \qquad (9)$$

To show this, observe first that

$$\tau\left[\tau(f_1, \ldots, f_k),\ \tau(f_{k+1}, \ldots, f_n)\right] = \tau(f_1, \ldots, f_k) \int_{-\infty}^{t} \tau(f_{k+1}, \ldots, f_n)\, dx$$

$$+ \tau(f_{k+1}, \ldots, f_n) \int_{-\infty}^{t} \tau(f_1, \ldots, f_k)\, dx$$

$$= \frac{d}{dt}\left(\prod_{j=1}^{k} F_j\right) \int_{-\infty}^{t} \frac{d}{dx}\left(\prod_{j=k+1}^{n} F_j\right) dx$$

$$+ \frac{d}{dt}\left(\prod_{j=k+1}^{n} F_j\right) \int_{-\infty}^{t} \frac{d}{dx}\left(\prod_{j=1}^{k} F_j\right) dx$$

$$= \frac{d}{dt}\left(\prod_{j=1}^{k} F_j\right)\left(\prod_{j=k+1}^{n} F_j\right)$$

$$+ \frac{d}{dt}\left(\prod_{j=k+1}^{n} F_j\right)\left(\prod_{j=1}^{k} F_j\right)$$

$$= \frac{d}{dt}\left(\prod_{j=1}^{k} F_j \prod_{\sigma=k+1}^{n} F_\sigma\right) = \frac{d}{dt}\left(\prod_{j=1}^{n} F_j\right)$$

$$= \tau(f_1, f_2, \ldots, f_n).$$

Then

$$\tau(f_o, \ldots, f_o;\ f, \ldots, f) = \tau\left[\tau(f_o, \ldots, f_o),\ \tau(f, \ldots, f)\right]$$

$$= \tau(f_o, \ldots, f_o) \int_{-\infty}^{t} \tau(f, \ldots, f)\, dx$$

$$+ \tau(f, \ldots, f) \int_{-\infty}^{t} \tau(f_o, \ldots, f_o)\, dx$$

$$= (n-k)\, f_o\, F_o^{n-k-1}\, F^k + kfF^{k-1}\, F_o^{n-k}$$

$$= F^{k-1}\, F_o^{n-k-1}\left[(n-k)\, Ff_o + kF_of\right].$$

From this one readily obtains Eq. 9.

For

$$f = \lambda e^{-\lambda t}$$

$$f_o = \lambda_o e^{-\lambda_o t}$$

$$\frac{\lambda_o}{\lambda} = 1.5$$

and

$$t = \frac{\alpha}{\lambda}$$

we compute the numbers shown in Table IV. 3.

Table IV. 3

Percentage Error in Replacing $\tau(f_o, f_o, f, f, f)$ by $\tau(f, f, f)$

α	0	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
$\frac{\Delta}{f_{max}}$	0.0	-9.7	-15.8	-11.8	-4.4	+5.2	+6.2	+4.4	+2.5	+0.7
$\frac{f}{f_{max}}$	100	77.8	60.7	47.2	36.8	22.3	13.5	8.2	5.0	1.8

Table IV.4

Percentage Error in Replacing $\tau(f_o, f_o, f, f, f)$ by $\tau(f, f, f)$

α	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	4.00
$\frac{\Delta}{f_{max}}$	0.0	-7.4	-11.3	-8.0	+1.5	+4.6	+3.8	+2.3	+0.7
$\frac{f}{f_{max}}$	100	77.8	60.7	47.2	28.7	17.4	10.6	6.4	2.4

If, however,

$$f = \lambda e^{-\lambda(t-\delta)}$$

$$\delta = \frac{0.25}{\lambda}$$

and

$$\frac{\lambda_o}{\lambda} = 1.5$$

the numbers shown in Table IV.4 are obtained.

We are unable to show that the conditions for Table IV.4 are met, but we can show that they seem to be reasonable, and certainly in any future experiments data may be obtained to show whether the assumption that $\tau(f_o, f, f, f, f_o) \approx \tau(f, f, f)$ is appropriate. Let

$$f_o = \lambda_o e^{-\lambda_o(t-t_o)}$$

and

$$f = \lambda e^{-\lambda(t-t_1)}$$

where

$$t_o = t_1 - \delta \cdot$$

Now the means of these two distributions are easily seen to be

$$\mu_o = t_1 - \delta + \frac{1}{\lambda_o}$$

$$\mu = t_1 + \frac{1}{\lambda}.$$

Let

$$q = \frac{\mu_o}{\mu}$$

then

$$\mu_o \lambda = t_1 \lambda - \delta \lambda + \frac{\lambda}{\lambda_o} = q\mu\lambda = q(t_1\lambda + 1).$$

So

$$q = \frac{t_1\lambda - \delta\lambda + \frac{\lambda}{\lambda_o}}{t_1\lambda + 1}.$$

Now suppose the error is negligible; that is, that

$$\delta \geqslant \frac{0.25}{\lambda}$$

$$\frac{\lambda_o}{\lambda} \geqslant 1.5$$

then, as we shall see from the next section, the approximate values of the parameters that fit the data are $\lambda \approx 0.1$ and $t_1 \approx 10$
so

$$t_1 \lambda \approx 1$$

and

$$q \leqslant \frac{1 - 0.25 + 0.66}{1 + 1} = 0.708.$$

That is to say, we are assuming only that the mean of the neglected distribution is 70 percent of the mean of the distribution f; this certainly has a ring of reasonableness.

6. Quantized Data: Experiment 4

We shall use the data in this section to show that two assumptions we have made are reasonable for the experiments we have run. The first assumption is that individuals can be treated as statistically identical so that group differences depend on the network rather than on the particular people in the network. We shall show this to be reasonable by demonstrating that significant differences among the networks occur, but that the inferred individual distributions are nonetheless not significantly different. The second assumption is that the individual distributions can be well approximated by exponential decay curves. We shall test the reasonableness of this by fitting curves to the group-time data and using the χ^2 goodness-of-fit test. Of course, these two assumptions are not independently tested, but what we are concerned with is the joint adequacy of our assumptions to describe the data.

There is a trend in the action time data with trials as can be seen in Figs. IV. 6, 7, 8, and 9. (We have presented the time data for only those cases we are able to analyze. Suffice it to say the rest are similar in nature to those given.) Fortunately, after a sharp initial drop, both the mean time and the variability about the mean are nearly constant. Because of this fact we can use all the data on a given network from the flat portion of the curve in a single combined frequency distribution to test goodness of fit. In particular, the trials eliminated were: the first three trials of circle, pinwheel, and chain; the first five trials of totally connected.

The group time per act was recorded on an Esterline-Angus pen recorder and the data read off the tape to the nearest second with the aid of a ruler. The sample size was 220 (ten groups for 22 trials) for circle, chain, and pinwheel; and 200 (ten groups for 20 trials) for totally connected. With these sample sizes and the obtained spread of the distributions the smallest class intervals which would yield acceptable values for expected frequencies were three seconds. It was found that in the process of measuring

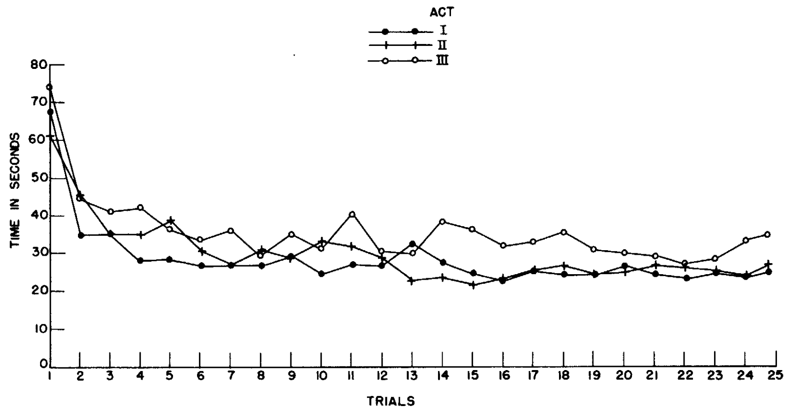


Fig. IV. 6

Mean time per act, circle (x).

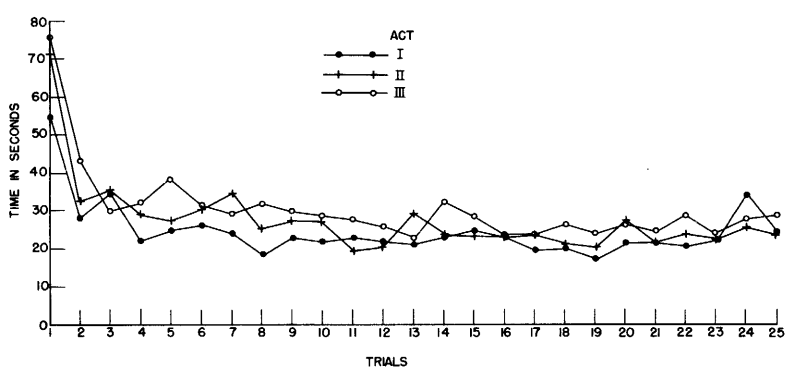


Fig. IV. 7

Mean time per act, chain (x-5).

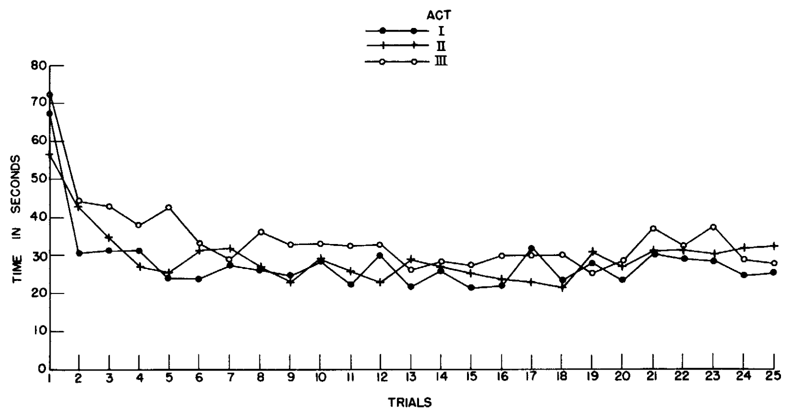


Fig. IV. 8

Mean time per act, pinwheel.

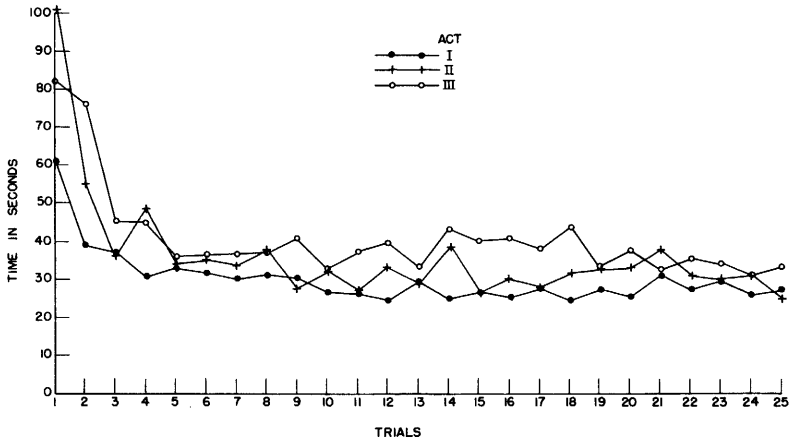


Fig. IV. 9

Mean time per act, totally connected.

time intervals from the tape, there was a fairly strong bias in favor of even numbers, and therefore only groupings into class intervals of an even number of seconds would adequately represent the shape of the distribution. For these reasons the circle, chain, and pinwheel data were grouped in class intervals of four seconds. In the case of the totally connected networks it was found that the data were very erratic with a four-second grouping, so a six-second grouping was employed.

The parameters of the distributions can be estimated from the moments by the equations from section IV. 4

$$\mu_1'(0) = t_o + \frac{g_1(n)}{\lambda}$$

$$\mu_2(0) = \frac{\sigma'}{\lambda} .$$

The values of $g_i(n)$ and σ' are given in Table IV. 1 and the rule for selection of the proper n is given in Table IV. 2. The distributions with which we have to deal are very skewed and, as is well known, moments are inefficient estimators for the parameters of skewed distributions. From a computational standpoint it was impractical to use a maximum likelihood estimator or a successive approximation to an efficient estimator since these methods lead, for the functions we have to fit, to complicated transcendental equations. Also the normal equations for a least-squares fit are of the same type. It was therefore decided to use a minimum χ^2 criterion to find the values of λ and t_o to fit the data, and this was done by numerical successive approximation. Since the sample size is fixed, and there are two parameters to be fitted, the degrees of freedom to be used in the goodness-of-fit test will in every case be the number of class intervals less three. The data on the fit of the theoretical curves to the empirical are given in Table IV. 5.

The theoretical fitted curves and the corresponding data points are shown in Fig. IV. 10 through Fig. IV. 21.

Using the values of λ and t_o found by the minimum χ^2 fitting criterion, we compute inferred values for the means and variances of individual action time distributions. These values and the group mean and variance data from which they were derived are shown in Table IV. 6.

In circle, chain, and pinwheel our hypothesis supposes that there will be differences in group mean action time from one network to another but, insofar as the theory is correct, no differences in individual mean action time (other than sampling variation). We expect differences from act to act in both group and individual means. The way that the nodes are assumed to contribute to the determination of group time in each network involves the consequence that each node which so contributes has two possible inputs and two possible outputs. Thus, whenever a person in any such nodal position has received information, he has a two-way choice of where to send that information. The nodes in the totally connected network differ in this respect: they each have four

Table IV.5

Goodness-of-Fit Test, Time Data, Experiment 4

Act I

Network	t_o	λ	χ^2	df	p
Circle (x)	8.1	0.120	6.29	7	0.51
Pinwheel	10.0	0.140	9.05	6	0.18
Chain (x)	9.6	0.135	4.46	5	0.49
Totally Connected	7.8	0.115	5.30	5	0.39

Act II

Network	t_o	λ	χ^2	df	p
Circle (x)	10.0	0.115	13.20	8	0.12
Pinwheel	10.0	0.115	12.31	7	0.10
Chain (x)	12.5	0.140	9.05	6	0.20
Totally Connected	12.5	0.113	5.50	5	0.37

Act III

Network	t_o	λ	χ^2	df	p
Circle (x)	11.8	0.106	14.36	9	0.11
Pinwheel	13.7	0.128	10.03	7	0.19
Chain (x)	14.5	0.120	7.23	7	0.41
Totally Connected	14.0	0.100	4.28	4	0.38

Table IV.6

Inferred and Observed Moments: Time Data, Experiment 4

	Inferred Individual μ'_1			Observed Group μ'_1		
Act	I	II	III	I	II	III
Circle	16.4	18.7	21.2	26.13	27.55	33.45
Pinwheel	17.1	18.7	21.5	26.21	27.35	32.00
Chain	17.0	19.6	22.8	22.89	25.16	28.21
Totally Connected	16.5	21.3	24.0	27.98	31.49	36.09
	Inferred Individual σ			Observed Group σ		
Circle	8.3	8.7	9.4	9.68	11.54	11.46
Pinwheel	7.1	8.7	7.8	10.95	10.46	12.28
Chain	7.4	7.1	8.3	9.96	9.42	10.09
Totally Connected	8.7	8.8	10.0	11.71	12.21	14.81

Table IV.7

Analysis of Variance: Time Data, Experiment 4

I. Raw Means

Source	df	ss
Total	8	86.22
Acts	2	60.88
Networks	2	23.01
Error	4	2.32

Acts

$F = 52.47$

$p \sim 0.001$

Networks

$F = 19.83$

$p \sim 0.001$

II. Transformed Means

Source	df	ss
Total	8	40.00
Acts	2	37.73
Networks	2	1.76
Error	4	0.51

Acts

$F = 147.96$

$p << 0.001$

Networks

$F = 6.90$

$p > 0.05$

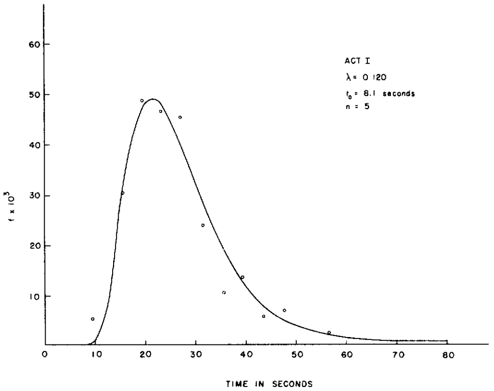


Fig. IV.10

Theoretical and observed time distribution, circle (x).

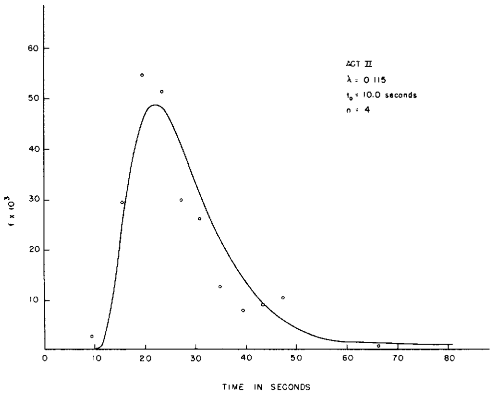


Fig. IV.11

Theoretical and observed time distribution, circle (x).

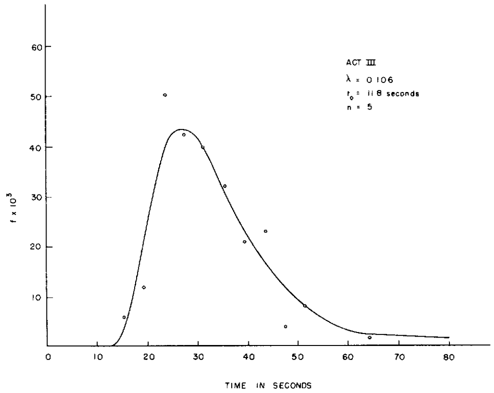


Fig. IV.12

Theoretical and observed time distribution, circle (x).

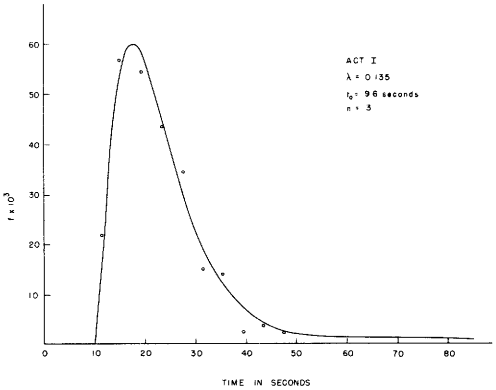


Fig. IV.13

Theoretical and observed time distribution, chain (x).

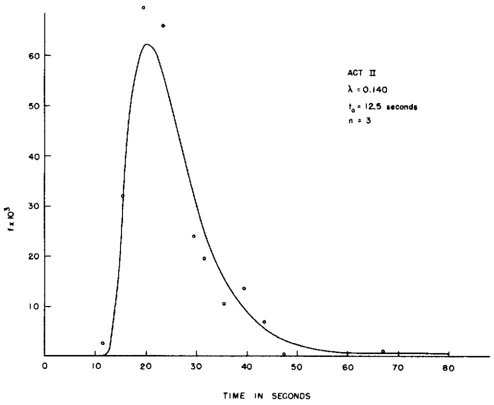


Fig. IV.14

Theoretical and observed time distribution, chain (x).

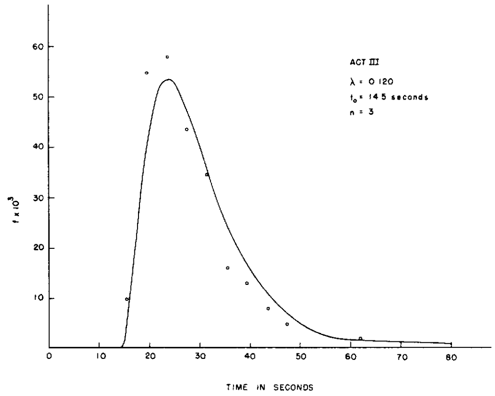


Fig. IV.15

Theoretical and observed time distribution, chain (x),

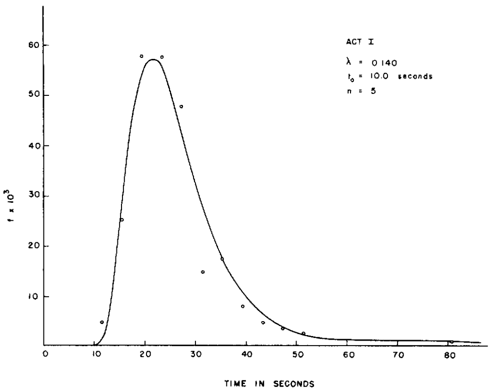


Fig. IV.16

Theoretical and observed time distribution, pinwheel.

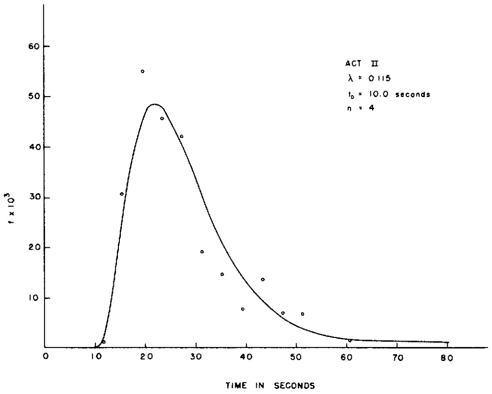


Fig. IV.17

Theoretical and observed time distribution, pinwheel.

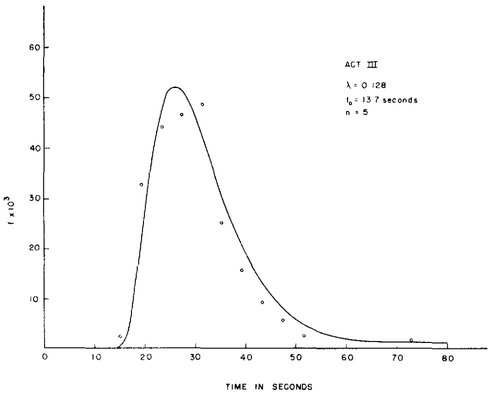


Fig. IV.18

Theoretical and observed time distribution, pinwheel.

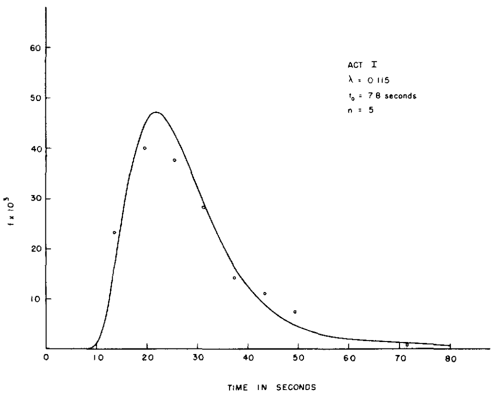


Fig. IV.19

Theoretical and observed time distribution, totally connected.

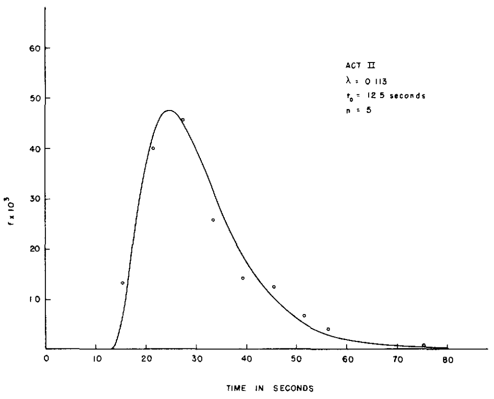


Fig. IV.20

Theoretical and observed time distribution, totally connected.

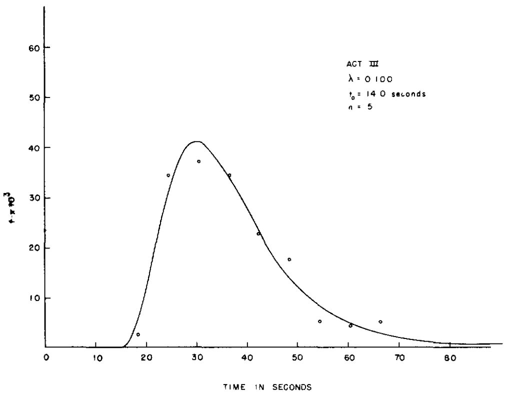


Fig. IV.21

Theoretical and observed time distribution, totally connected.

possible inputs and four possible outputs. On this account we do not expect individuals in totally connected networks to have the same time parameters as individuals in the other three networks.

Table IV.7 presents an analysis of variance for the group mean data and for the inferred individual means for circle, chain, and pinwheel on acts I, II, and III.

These results bear out the contention of our previous theoretical argument that to a first approximation at least, it is possible to treat individuals as statistically identical. The significant difference among networks in raw mean action time disappears when we find transformed mean action times for the individuals in the respective networks. Thus the differences among groups using different networks are shown to depend on differences in the networks, as they must, and it is shown that the approximate theoretical treatment we have given is adequate to account for these group differences to the extent that they exist in our data.

7. Quantized Data: Experiment 5

Experiment 5, a common marble experiment, was an action-quantized experiment similar in apparatus and procedure to Experiment 4 whose time results have just been discussed. The former, however, differed from Experiment 4 in several important respects:

(a) Task. Each subject was given a set of five differently colored marbles, with one color common to all five sets. The task for each subject was to determine that common color. This is to be contrasted with the task in Experiment 4 in which each subject was given a number and a trial was concluded when each subject had obtained all five numbers.

(b) Constraints. The subjects were constrained to action on a quantized time scale, determined in exactly the same fashion as in Experiment 4, and they were obliged to send one, and only one, message card in each act. However, the content of their messages was in no way restricted.

(c) Experimental conditions. Three networks (circle, chain, and star) were studied. Six experimental groups were examined in the circle network, six groups in the chain, and seven groups in the star, each group participating in a series of fifteen trials. The number of acceptable trials had to be reduced to 15 from the desired 30 to encompass the experiment in a reasonable length of time.

We may draw an immediate conclusion: From such small samples it is unlikely that the data will prove to be very stable, and this variability will be even worse than a random selection of six or seven samples from Experiment 4, since the message content was unconstrained.

We shall proceed in a fashion analogous to the last section. In Fig. IV.22 through Fig. IV.25, we have plotted mean time per act vs trials for the different networks. We see from these that only after a period of learning do we obtain relatively homogenous populations; consequently we have omitted early trials and lumped the remaining ones.

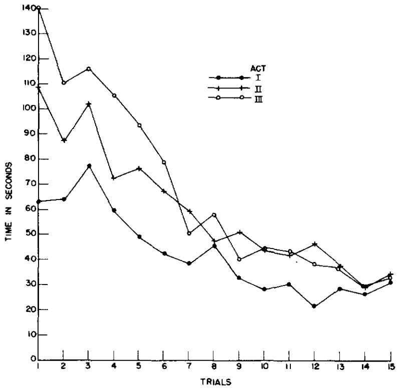


Fig. IV.22

Mean time per act vs trials, circle (x), 6 groups.

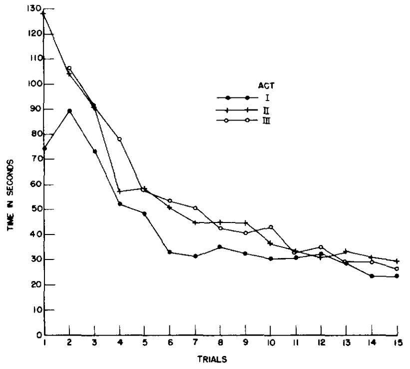


Fig. IV.23

Mean time per act vs trials, chain (x), 6 groups.

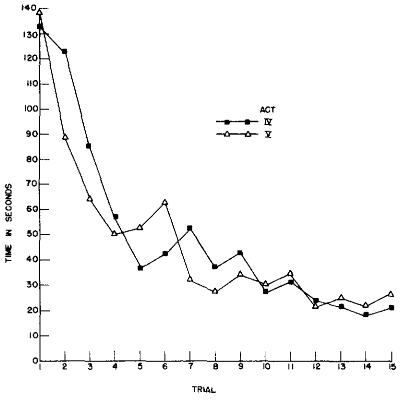


Fig. IV.24

Mean time per act vs trials, chain (x), 6 groups.

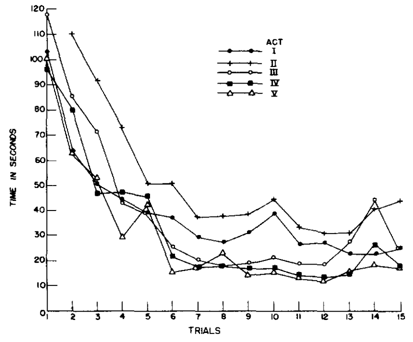


Fig. IV.25

Mean time per act vs trials, star, 7 groups.

Table IV. 8 indicates which trials have been used.

Table IV. 8

Trials Used from Experiment 5

	Act				
Network	I	II	III	IV	V
Circle	6 - 15	7 - 15	7 - 15	—	—
Chain	6 - 15	5 - 15	5 - 15	8 - 15	8 - 15
Star	6 - 15	7 - 15	6 - 15	6 - 15	6 - 15

Within this range of trials the data may be lumped on the basis that the populations from which the samples occurred are nearly homogenous. Even with such lumping of the data, we obtain sample sizes of the order of 40 points to be distributed on a time scale from, roughly, 10 sec to 50 sec. It is thus apparent that even when we group this data into, say, 6- or 8-sec intervals, the sample size is too small to expect stable frequency distributions. Thus, it would be desirable whenever possible to lump the data further; this must mean combining several acts. The lumping of acts can be argued on both an a priori basis and by reference to the curves of Fig. IV.22 through Fig. IV.25.

Circle: Acts I and II are certainly from different populations, since in the former, no filtering of the information can occur, whereas it can occur in the latter act. By filtering we mean the action of reducing the number of marbles that may be the common one, when a node has his and another set of information. In addition, there exists a greater problem for decision in the second act. If filtering of the information does occur in the second act, then each node has more information to send in the third act, but it has been reduced to a more compact form by the filtering process. Thus, act III is very similar in process to act II, which suggests combining them. The curves of Fig. IV.22 substantiate the contention that acts II and III are nearly the same, and that act I is quite different.

Chain: Act I, as in the case of the circle, is unique. By a similar argument, acts II and III can be lumped together, for in the chain, act II results in a filtering of the information by the middle men, who then pass it to the center man. Much of the decision time for the third act is occupied by the occurrence of the filtering process in the center, which can, when the group is operating efficiently, obtain the answer at this point. Acts IV and V consist in little more than a relay of this information from the center man to the end men, so that the entire group knows the result. There may, of course, be more acts than this to obtain the answer; however, by the fifth or sixth trial the organization of the chain is quite well worked out (see secs. III. 3 and III. 7) so that the groups are doing the problem in minimum acts nearly all the time. The argument is neither particularly well upheld nor destroyed by the curves of Figs. IV.23 and IV.24.

Star: Act I is unique. Since everyone must send a message, the center man must have all the information to obtain the answer at the end of act I. In act II, most of the decision time is probably taken up by his filtering the data to obtain the answer. So act II is unique also. Acts III, IV, and V are simply a relaying of this answer to the other three end men in the star, and so they must all be essentially the same. This, of course, is true only after the efficient organization of the star has been obtained; this occurs very early as we have seen in Fig. IV.25. Except for trial 14, this seems to be the case for the groups we have run.

We have lumped our data as indicated above, and we have grouped them in either 6- or 8-sec intervals. The choice depended on the spread of the data. We found that finer groupings were not satisfactory, since the sample size in an interval became so small that the variability of the data gave a very rough curve. The grouping used is indicated in Table IV.9.

We have not attempted to fit theoretical curves to this data as we did in Experiment 4, because we did not think that the computational effort was warranted by the nature and meagerness of this data. We shall try to show in the following discussion that a priori considerations would lead, at least in some cases, to curves having three or more parameters to fit this data. If this is the case, then we would need much finer groupings, and hence much more data, to procure the necessary degrees of freedom to obtain a test of our fit.

In the analysis of Experiment 4, we assumed a constant value of t_o, the absolute minimum time in which a decision could be reached. We may interpret t_o as the minimum writing time, which, for the case of messages allowing nothing but numbers, may well be nearly a constant. On the other hand, in Experiment 5 the subjects were allowed to write anything they wished and had to write the colors of several marbles to send a message. Here, individual differences in writing speed began to be a factor, and thus it is somewhat unreasonable to suppose t_o to be constant. If we suppose that t_o is distributed in some fashion depending on two parameters, the location of the mean and some measure of the shape, then the fit requires at least three parameters.

This would account for some of the difficulty at the initial point of these distributions, which is not very serious, but it would certainly not account for, say, the shape of the distribution in acts II and III combined of the chain or act II of the star. Let us suppose that t_o is a fixed number and not distributed, then we know that the individual time distribution is

$$\lambda(t-t_o) \exp\left[- \int_{t_o}^{t} \lambda(t)dt \right].$$

Throughout the previous discussion we have assumed $\lambda(t)$ to be a constant; this has

Table IV.9

Time Per Act Frequency Distribution, Experiment 5

Circle

Act											
I	Interval	11-16	17-22	23-28	29-34	35-40	41-46	47-52	53-58	59-∞	
	Frequency	2	13	11	8	7	5	2	4	5	
II + III	Interval	7-14	15-22	23-30	31-38	39-46	47-54	55-62	63-70	71-78	79-∞
	Frequency	2	10	24	21	17	11	8	4	6	4

Chain

Act											
I	Interval	14-17	18-21	22-25	26-29	30-33	34-37	38-41	42-45	46-49	50-∞
	Frequency	2	5	12	10	9	9	5	2	2	4
II + III	Interval	13-18	19-24	25-30	31-36	37-42	43-48	49-54	55-60	61-∞	
	Frequency	2	18	24	18	19	17	16	8	10	
IV + V	Interval	5-10	11-16	17-22	23-28	29-34	35-40	41-46	47-52	53-∞	
	Frequency	1	20	23	16	10	11	4	3	8	

Star

Act													
I	Interval	9-14	15-20	21-26	27-32	33-38	39-44	45-∞					
	Frequency	8	16	21	11	4	2	8					
II	Interval	10-17	18-25	26-33	34-41	42-49	50-57	58-65	66-∞				
	Frequency	2	11	21	9	6	3	6	5				
III + IV + V	Interval	4-7	8-11	12-15	16-19	20-23	24-27	28-31	32-35	36-39	40-43	44-47	48-∞
	Frequency	4	64	44	30	24	11	7	4	7	4	2	8

proved to be a good assumption when the decision problem is very simple indeed and is not confounded by other factors. We have no reason to suppose that this is the case when the decision problem becomes much more complex. In fact, in the filtering of information – the decision as to which colors are common to these two sets of colors – there are intuitive reasons to suppose that this is not the case, but that $\lambda(t)$ builds up from a 0 value to some asymptotic value λ. A particularly simple form of such a build-up is

$$\begin{aligned} \lambda(t) &= \lambda - \frac{k}{t} \quad t \geqslant t_o \\ &= 0 \qquad\quad t < t_o \end{aligned}$$

which yields the distribution

$$\left(\lambda - \frac{k}{t}\right)\left(\frac{t}{t_o}\right)^k e^{-\lambda(t-t_o)} \quad t \geqslant t_o \tag{11}$$

This distribution is dependent on three parameters and is, for k an integer, the difference between a type III, order k distribution, and an order k - 1 chopped off at the lower end at $t = t_o$. It may be reduced to a distribution depending on only two parameters if we make the reasonable assumption that $\lambda(t_o) = 0$; that is,

$$\lambda = \frac{k}{t_o}.$$

A third parameter n is introduced, as before, in the decision as to how many are the identical distributions from which the selection is made. Without further measurements this is most easily obtained by the same a priori arguments we previously made. For the circle and the chain we would use the same values: 5, 4, and 3, 3. For the star the argument is more complex. As in the chain, the end men will have distributions whose means are considerably less than that of the center man in all cases; however, when the decisions are simple as in, for example, the act I or acts III, IV, and V, it does not seem unreasonable to suppose that the distribution of the largest value from the selections from the end-man distributions has a mean comparable to that of the center man. Thus we cannot very well approximate the group distribution by the center-man distribution. This will entail assuming two distributions of, for instance, the exponential type, and thus fitting the data by 4 parameters, or 3, if we assume the same t_o for each. In act II the decision problem of the center man is so much more complex than that of the end men, one might very well assume that the group distribution and the center-man distribution are the same, but that the distribution is of the form of Eq. 11.

In all the cases mentioned and summarized in Table IV.10, we can estimate the parameters λ and t_o from the first two moments, except in the star cases that require two distributions. Suppose we have the distributions

$$f_o = \lambda_o e^{-\lambda_o t} \qquad \text{and} \qquad f_1 = \lambda_1 e^{-\lambda_1 t}$$

and make the assumption that they both have the same dead time t_o. Then a linear translation of the time axis will allow us to make $t_o = 0$. From Eq. 10 we know

$$\tau(f_o, f_o, \ldots, f_o, f_1) = (n-1)\, F_o^{n-2} F_1 f_o + F_o^{n-1} f_1$$

Thus the i-th raw moment for a star on n nodes is

$$\mu_i^1(n) = \int_{-\infty}^{\infty} t^i\, \tau(f_o, f_o, \ldots, f_o, f_1)\, dt$$

$$= (n-1) \int_0^{\infty} t^i \left(1-e^{\lambda_o t}\right)^{n-2} \left(1-e^{-\lambda_1 t}\right) \lambda_o e^{-\lambda_o t}\, dt$$

$$+ \int_0^{\infty} t^i \left(1-e^{-\lambda_o t}\right)^{n-1} \lambda_1 e^{-\lambda_1 t}\, dt\, .$$

Recall that in Eq. 7 we defined

$$g_i(n) = \lambda_o^i n \int_0^{\infty} t^i \lambda_o e^{-\lambda_o t} \left(1-e^{-\lambda_o t}\right)^{n-1} dt$$

which was shown to be

$$g_i(n) = n \sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j \frac{i!}{(j+1)^{i+1}}.$$

Thus,

$$(n-1)\, \lambda_o \int_0^{\infty} t^i \left(1-e^{-\lambda_o t}\right)^{n-2} e^{-\lambda_1 t}\, e^{-\lambda_o t}\, dt$$

$$= (n-1)\, \lambda_o \int_0^{\infty} t^i \sum_{j=0}^{n-2} (-1)^j\, e^{-\left[(j+1)\lambda_o + \lambda_1\right] t} \binom{n-2}{j} dt$$

$$= \frac{(n-1)}{\lambda_o^i} \sum_{j=0}^{n-2} (-1)^j \binom{n-2}{j} \frac{i!}{\left[j+1 + \frac{\lambda_1}{\lambda_o}\right]^{i+1}}$$

$$= \frac{g_i(n-1)}{\lambda_o^i} + \frac{(n-1)}{\lambda_o^i} \sum_{j=0}^{n-2} (-1)^j \binom{n-2}{j} \sum_{\rho=1}^{\infty} \left(\frac{\lambda_1}{\lambda_o}\right)^{\rho} \frac{(-1)^{\rho}}{\rho!} \frac{(i+\rho)!}{(j+1)^{i+\rho+1}}.$$

Similarly,

$$\int_0^{\infty} t^i \left(1-e^{-\lambda_o t}\right)^{n-1} \lambda_1 e^{-\lambda_1 t}\, dt$$

$$= \frac{(n-1)}{\lambda_o^i} \sum_{j=0}^{n-2} (-1)^j \left(\frac{n-2}{j}\right) \sum_{\rho=1}^{\infty} \left(\frac{\lambda_1}{\lambda_o}\right)^{\rho} \frac{(-1)^{\rho}}{\rho!} \frac{(i+\rho)!}{(j+1)^{i+\rho+1}} \left(\frac{\rho}{i+\rho}\right) + \frac{i!}{\lambda_1^i} .$$

Thus,

$$\mu_i^1(n) = \frac{(n-1)}{\lambda_o^i} \sum_{j=0}^{n-2} (-1)^j \left(\frac{n-2}{j}\right) \sum_{\rho=1}^{\infty} \left(\frac{\lambda_1}{\lambda_o}\right)^{\rho} \frac{(-1)^{\rho}}{\rho!} \frac{(i+\rho)!}{(j+1)^{i+\rho+1}} \left(\frac{\rho}{i+\rho} - 1\right) + \frac{i!}{\lambda_1^i}$$

$$= \frac{i!}{\lambda_1^i} - \frac{1}{\lambda_o^i} \sum_{\rho=1}^{\infty} \left(\frac{\lambda_1}{\lambda_o}\right)^{\rho} \frac{(-1)^{\rho}}{\rho!} \left(\frac{i}{i+\rho}\right) g_{i+\rho}\,(n-1) .$$

Approximating this series by, say, the first three terms, one can obtain estimates on the parameters quite rapidly; however, the data must be quite good for these estimates to be tolerable since we require not only the first two moments, but also the third. We do not believe that there is sufficient time data in Experiment 5 to warrant this computation.

Table IV.10

Summary of Assumptions for Analysis of Experiment 5 Time Data

		Circle	
Acts	I	II + III	
	$k = 0$	$k > 0$	
	$n = 5$	$n = 4$	
		Chain	
Acts	I	II + III	IV + V
	$k = 0$	$k > 0$	$k = 0$
	$n = 3$	$n = 3$	$n = 3$
		Star	
Acts	I	II	III + IV + V
	$k = 0$	$k > 0$	$k = 0$
	two distributions	$n = 1$	two distributions

8. Nonquantized Data: Experiment 3

In his common-symbol experiment (No. 1) Leavitt recorded the time taken for each trial and the number of messages sent on each trial (34). This data is shown in Fig. IV.26 and Fig. IV.27, respectively. From these plots it can be seen that both time per trial and messages per trial decrease as the groups experience more and more trials. The groups were motivated to complete each trial in as short a time as possible, and they made progress toward this goal. We should like to be able to say how such progress was made. From the gross impressions of the experimenter, it was evident that shorter times involved both a decrease in the time taken per message and a decrease in number of messages per trial. It would be desirable to assess the relative importance of these two modes of achieving time economy. However, the time per trial is not simply the product of time per message and message per trial, because in this nonquantized case, trials are not composed of messages in simple succession. On the contrary, since each man in the group was free to send at any time, the time to prepare and send a message at one node overlapped the time to prepare and send a message at other nodes in an irregular manner. Leavitt's data is not amenable to a rational treatment to circumvent this difficulty.

Smith's noisy marble experiment (No. 3) produced the same sort of time data as Leavitt's experiment. His results for time per trial and messages per trial are shown in Fig. IV.28 and Fig. IV.29, respectively. Smith also reconstructed the pattern of message sending for each trial of each group he ran. If we make the assumption that the sending pattern observed in Smith's nonquantized situation is approximately what would have occurred if the experiment had been quantized, it is possible to find for each trial the equivalent number of act quanta to do that trial with its given message-sending pattern. When this has been done, we can conceive of the trials as composed of a succession of act equivalents and therefore interpret the time per trial as the product of number of act equivalents per trial and time per act equivalent. We have given as data the time per trial, and from the message-sending pattern we have deduced the number of act equivalents per trial. Simply dividing the former by the latter gives us the time per act equivalent. We are therefore in a position to give at least an approximate evaluation to the roles of speed of action (time per act equivalent) and efficiency of action (act equivalents per trial) in producing trial time economy. The act interpretation of the message-sending patterns for Smith's star and chain groups was sufficiently clear, so that this task could be done with fair confidence in the reliability of the result. The data, plotted against trials, is shown in Figs. IV.30 and IV.31. The patterns for the circle groups were too confused to be usable.

Once we have reduced the nonquantized data roughly to acts, they are approximately in a form which enables us to give them a treatment similar to that which we have developed in detail for the quantized case, and hence we can, in principle, compare the two types of experiments in their temporal aspects. From the data on time per act

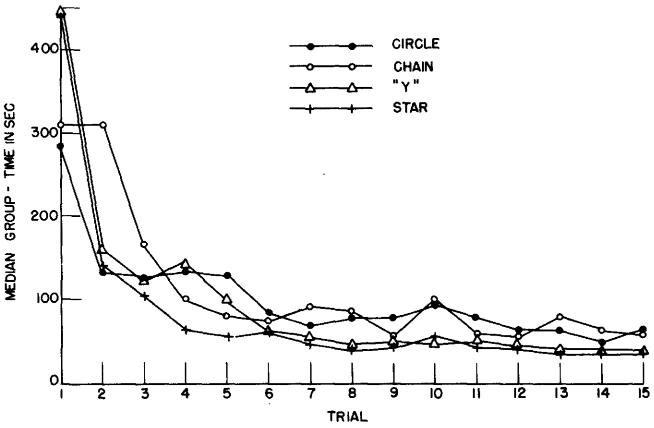


Fig. IV.26
Median group time per trial, Experiment 1 (correct trials only).

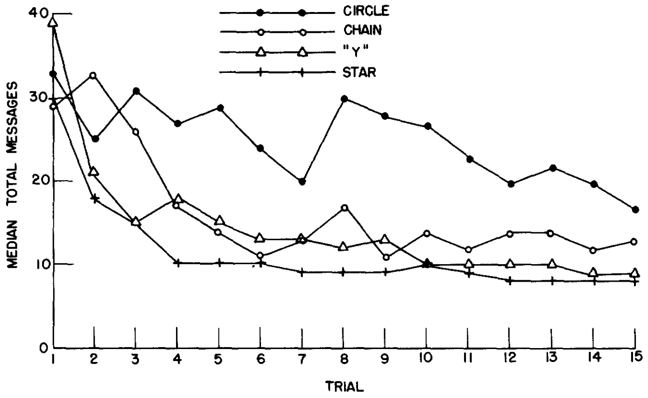


Fig. IV.27
Median messages per trial, Experiment 1 (correct trials only).

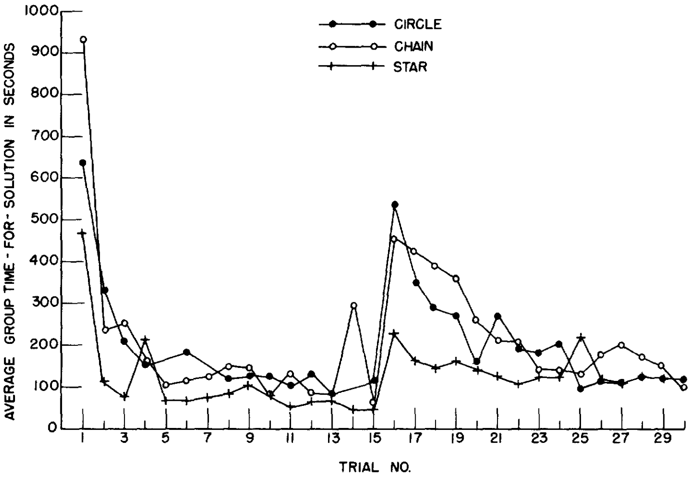


Fig. IV.28
Mean time per group, Experiment 3.

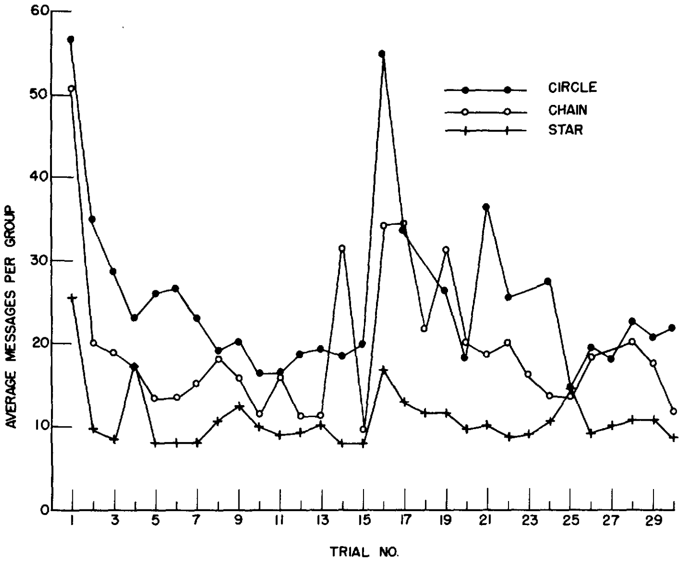


Fig. IV.29
Mean number of messages per group, Experiment 3.

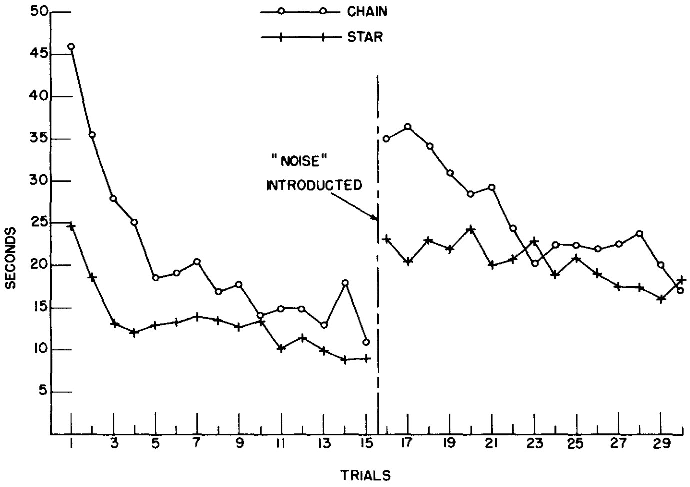


Fig. IV.30
Time per act equivalent.

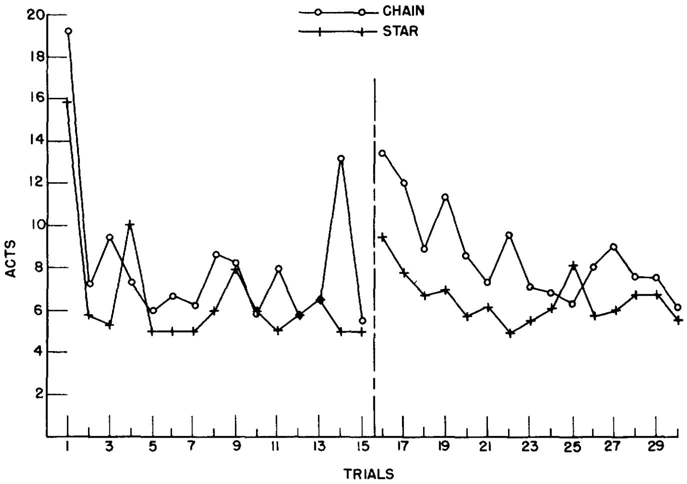


Fig. IV.31
Act equivalents per trial.

equivalent we can plot frequency distributions of action time. We shall develop below the theoretical form of this curve, making the same assumptions about the delay functions of individual nodes as we used in the sections on the quantized case. By fitting the appropriate theoretical distribution functions to the empirical distributions in each case, we can calculate nodal parameters and thereby compare the two experimental situations in terms of these parameters. The reason for not doing so is given below.

We shall neglect the fact that nodal parameters may vary somewhat from act to act and treat all acts within a given group's performance on the same footing. Let us denote the number of acts in a trial by k, the total time for a trial by t_k and the time per act by t. Hence

$$t = \frac{t_k}{k}$$

If the i-th act terminates at t_i, we can write

$$t_k = kt = t_1 + (t_2 - t_1) + \ldots + (t_k - t_{k-1})$$

and if we substitute

$$\tau_i = t_i - t_{i-1}$$

$$t_k = kt = \sum_{i=1}^{k} \tau_i .$$

We have assumed that we have identical nodes and also unchanging parameters from act to act; therefore the distribution function for τ_i is the same for all i, and we may find the distribution of t_k as the distribution of a sum of variables with known distributions. From the distribution for t_k the distribution for t is easily found.

We assume (see sec. IV. 3) that τ_i is distributed as

$$f(\tau_i) = \lambda e^{-\lambda(\tau_i - t_o)}$$

and find that t is distributed as

$$f(t) = \frac{k^k \lambda^k e^{-\lambda kt}}{1 - \lambda t_o} \left[\frac{t^{k-1}}{(k-1)!} - \frac{t_o t^{k-2}}{k(k-2)!} \right] .$$

Note that in this expression the parameter k represents number of acts per trial. The value of k is not fixed for any group but has a distribution. We have not investigated the problem of the form of the k distribution, but we can give a gross characterization on a priori and empirical grounds. The smallest value k can have is known a priori to be 5 for both star and chain, and we know empirically that the frequency of the larger values of k becomes rapidly negligible. From these facts we conclude that when the distribution of k is taken into account in the expression for f(t), the result will be a

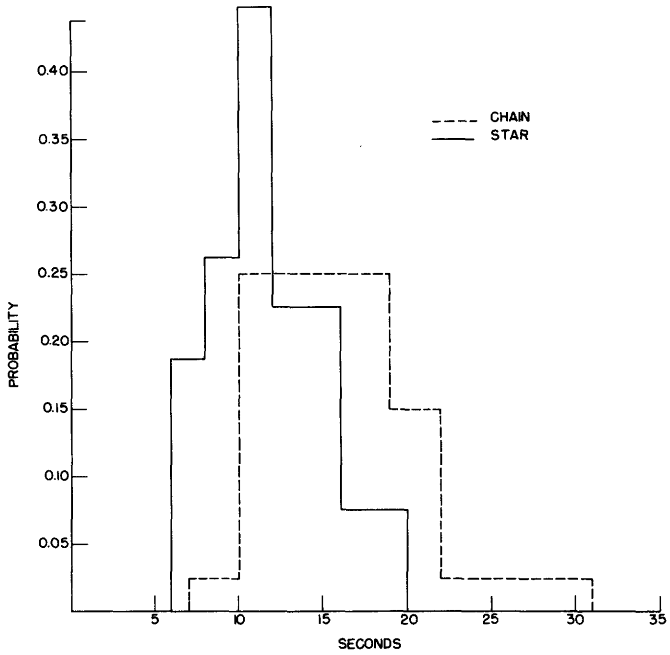


Fig. IV.32
Distributions of time per act (trials 6-15).

weighted sum of type III curves of various orders. The result is thus similar in general form to the results previously obtained for the quantized case. It is further to be noted that the theoretical frequencies which are given by f(t) are critically dependent on the value of the product, λt_o. The data available in Experiment 3 is too meager to warrant the use of curve-fitting techniques. Therefore, we must be content to present the empirical distributions shown in Fig. IV.32, from which it is apparent the data are at least of the general form required by our theoretical discussion.

Inspection of Fig. IV.32 shows the times per act in trials 6 through 15 to be shorter in star than in chain to a degree which makes a statistical test of significance superfluous. On the other hand, the act distributions during these same ten trials do not differ significantly, as may be seen in Table IV.11.

Table IV.11

Acts per Trial	Frequency		Significance
	Star	Chain	
5	17	19	df = 3
6	6	5	$\chi^2 = 2.673$
7, 8	11	6	$p \sim 0.40$
>8	6	10	

During the five initial trials the act distributions do differ; see Table IV.12.

Table IV.12

Acts per Trial	Frequency		Significance
	Star	Chain	
5	12	6	df = 1
>5	8	14	$\chi^2 = 3.78$
			$p \sim 0.05$

The time-per-act distributions also have different means in the first five trials. Wilcoxon's Matched Pairs Signed Ranks Test was used because of the strong departure from normality in the data. With $\mu = 20$ and τ (sum of ranks with less frequent sign) $= 18$, the test gives $p < 0.01$.

Thus we may describe the difference in performance of star and chain in the following way: Star has an initial advantage in time per act and continues to have this advantage. Star has an initial advantage in acts per trial, but chain improves more than star and obliterates this difference. The improvement in time per act is slower than the improvement in acts per trial, so that improvement in time per trial is at first due for the most part to an increase in efficiency of organization and is later due for the most part to an increase in speed of action for both networks. This phenomenon is stronger for star than for chain.

9. On The Interrelation of Decision Time and Decisions

Quantized action can be imposed upon a group in the context of laboratory experimentation, but it is not ordinarily found in the behavior of groups outside the laboratory. Typically, the time at which a message is sent, as well as the person to whom it is sent, is determined by the properties of the sending node and its inputs of messages from other nodes rather than by a strict rule imposed from without the group. As an example, consider that node i sends to j if he has not received from j, but sends to k if he has received from j. Suppose further that j is preparing a message for i at the same time that i is preparing his message. It then becomes inescapably important to compare the action times of i and j, namely t_i and t_j, since if $t_i > t_j$, i's message will go to j, whereas if $t_j > t_i$, it will go to k. It is clear then that the conditions which determine where a message is sent are inextricably bound up with action times.

The determination of conditional probabilities and of action time distributions for the nodes is a straightforward empirical problem. Given these measurements, the calculation of the resulting statistics of group performance is a purely mathematical problem, but one of great complexity. As we have pointed out in the previous chapter, even with a highly restricted information flow and with time and decision independent, it is not practical, except for one case, to calculate the group-performance statistics. In the future, high-speed computers may make other calculations possible. However, one might hope for a mathematical theory which, by employing probability distributions for the transfer functions of the nodes, leads to a simpler statement of the problem and possibly a solution in terms of tabulated functions. As an example of this type of theory, Appendix 3 presents a human group interpretation of the mathematical formalism of linear electrical network theory. This theory, it will be noted, allows both decision times and decision probabilities to be taken into account; that is, it is not restricted to an action-quantized situation. However, the interpretation there has been accomplished only at the expense of an enormous simplification of the group-network problem. Conditional probabilities are not admitted: their place is taken by fixed sending probabilities for each outgoing link. Even more drastically simplifying is the restriction of the information flow to one (repeated) item which is simply counted each time it is received at each node. This, of course, is not an adequate theory. The removal of these restrictions in order to have a theory which both applies to more general group situations and admits mathematical solution is a major theoretical undertaking.

In this section we propose to formulate the assumptions of a general problem which involves both decision and time for decision interaction and more than one elementary particle of information. Attempts, so far, to rephrase these assumptions as a system of integrodifferential equations have failed. It is hoped that by stating the problem exactly, others will succeed where we have failed. There is more point to the section that this, however, for we shall outline a method of prediction from the assumptions which is computationally simple and which will permit experimental verification of these assumptions (or ones similar to them) when data are available.

Suppose a communicating group of n nodes, $1, 2, \ldots, n$, is dealing with a set of information U. The elements $a \in U$ are essentially labels assigned to indivisible pieces of information (indivisible possibly only in the context of the problem). Assume as given for each node i, a time delay function $f_i(t)$, and for each pair of nodes i, j, a set of conditional sending probabilities $r_{ij}(V)$, where V ranges over all possible subsets of U.

It is assumed that if node i has received a message from or has sent a message to any other node at time τ, and no other message has been sent or received in the interval from τ to t, then the probability density that i sends a message at t is $f_i(t-\tau)$. If, however, an intermediate message arrives or is sent at time τ_1, $\tau < \tau_1 \leqslant t$, then the process begins anew at time τ_1. This is to say, each node is activated according to the given delay function by the immediately preceding incoming or outgoing message, and this activation is independent of what information it has at the time.

If at time t the node i does in fact send a message, and if at that time it has exactly the subset V of U, then the probability that the message is sent to j is assumed to be $r_{ij}(V)$. It is explicitly assumed that the conditional decision probabilities do not change directly with time.

Finally, assume that when a node does send a message, it sends all the information that it has and it does this without any loss of memory to itself. Thus, if at time t node i has $V \subset U$, and node j has $V' \subset U$, and j receives a message from only i, and i receives no message, then following the message exchange i has the information V and j the information $V \cup V'$. In summary, then, the assumed transfer function of the idealized person implies that he has a perfect memory of the content of the messages received, that any outgoing message is governed by a given delay function $f_i(t)$ which is initiated by the last message he sent or received, that his choice of destination is governed by the set of conditional probabilities $r_{ij}(V)$ which are independent of time, and that each message he sends contains all the information he has at the time of sending.

In addition, a driving function for the group is given in the following form: for each node i and each element $\alpha \in U$, there exists a frequency function $g_i(\alpha, t)$ which describes the input of the information α to the node i from without the group. Clearly, if i receives α N times, then

$$\int_{-\infty}^{\infty} g_i(\alpha, t)\, dt = N.$$

The nodes are assumed to react to an input from without the group in the same fashion as to a message from within the group.

Problem: given $f_i(t)$, $r_{ij}(V)$, and $g_i(\alpha, t)$, $i, j = 1, 2, \ldots, n$, $\alpha \in U$, $V \subset U$, determine the probability $P_i(\alpha, t)$ that node i has the element of information $\alpha \in U$ at time t.

This problem may be simplified so that the driving function of the group is simply an initial condition if there exists some t_o (which by a simple translation of the time scale may be taken to be 0) such that

$$g_i(\alpha, t) = 0, \quad \text{for all } i,\ \alpha, \text{ and } t \leqslant t_o.$$

If this is so, then for each i, α pair such that $g_i(\alpha, t) \not\equiv 0$ we shall introduce a new node, called $i(\alpha)$. This node is chosen to have the properties that it has only the information α at time $t = 0$, that it has a delay function $g_i(\alpha, t)$, that it can send only to the node i, and that it can receive messages from no other node. Formally

$$f_{i(\alpha)}(t) = g_i(\alpha, t)$$

$$r_{i(\alpha)i}(V) = 1, \quad \text{for all } V$$

$$r_{i(\alpha)j}(V) = 0, \quad \text{if } j \neq i$$

$$r_{ji(\alpha)}(V) = 0 \quad .$$

It follows immediately that

$$P_{i(\alpha)}(\beta, t) = 1, \quad \text{if } \alpha = \beta$$

$$= 0, \quad \text{otherwise}$$

using the same notation as before. We shall assume this property. In this augmented group the driving functions no longer exist; they are replaced by the following initial conditions: For the original nodes

$$P_i(\alpha, o) = 0, \quad \text{for all } \alpha \epsilon U$$

and for the added nodes

$$P_{i(\alpha)}(\beta, o) = 1, \quad \text{if } \alpha = \beta$$

$$= 0, \quad \text{otherwise.}$$

Thus, the problem is to determine $P_i(\alpha, t)$, $i = 1, 2, \ldots, n$, given $f_i(t)$, $r_{ij}(V)$, and the above initial conditions. This, as we mentioned, can in principle, and possibly in practice, be carried out if the problem is reduced to a system of integrodifferential equations. The source of mathematical difficulty appears to be the following: Let the state of the system be described by the information at each node; then the system is not a Markov process in which "the future development is completely determined by the present state and is independent of the way in which the present state developed" (see ref. 81, Chap. VI, p. 337). This is clear, for let two systems be in exactly the same information state; but in the one, let node i be most recently stimulated at time τ, and in the other at time τ', $\tau \neq \tau'$. The probabilities governing the behavior of i, and so of the system, are different in the two cases. Very little of a general nature is known about the solution of non-Markov systems. If the definition of state of the system is altered, a Markov formulation can be given, but a new difficulty arises. Let the state of node i at time t be the information $V(i, t)$ at the node, and the time $\tau(i, t)$ of the most recent stimulation of i. For small Δt, the state variable $\tau(i, t + \Delta t) = \tau(i, t)$ or τ' where $t \leqslant \tau' \leqslant t + \Delta t$. Thus, τ does not pass through a continuous change, and any equations we write must take into account a discontinuous variable. So, looking at the problem either way, our difficulties are not surprising; indeed, they may be expected to exist for any intuitively reasonable and nontrivial characterization of information flow in a group.

Fortunately, the practical problem of prediction about experimental group results is at least partially soluble. This will be best illustrated by an example. Suppose three nodes, 1, 2, and 3, are connected as a chain with the links [12], [21], [32], [23], and each has an initial piece of information. Since the problem is essentially finite, one may indicate all possible orderings of message sendings which will lead to a solution. A solution, in this case, is defined as each node knowing what initial information the others have. Of all these possible orderings, only a relatively few will actually arise with any frequency in practice, and we would feel confident that the assumptions applied if from them we could predict these frequencies.

Consider a particular solution of the problem for the three-node chain. Suppose the sending [12] occurs in $(\tau_1, \tau_1 + \Delta\tau_1)$, [32] in $(\tau_2, \tau_2 + \Delta\tau_2)$, [21] in $(\tau_3, \tau_3 + \Delta\tau_3)$, and

[23] in $(\tau_4, \tau_4 + \Delta\tau_4)$, where $0 \leqslant \tau_1 < \tau_1 + \Delta\tau_1 \leqslant \tau_2 < \tau_2 + \Delta\tau_2 \leqslant \tau_3 < \tau_3 + \Delta\tau_3 \leqslant \tau_4 < \tau_4 + \Delta\tau_4 \leqslant t$. We shall show below that it is possible from our assumptions to estimate the probability of this occurrence. Then, if we sum (integrate) over all possible ways this ordering might occur, we shall obtain the probability that a solution was obtained by time t, using the given ordering. For small groups with simple networks, the number of orderings which will have significant frequencies will be relatively small, so that prediction and comparision will be possible.

Let us calculate the probability of occurrence of the above ordering, where we assume $f_i(t) = f(t)$, $i = 1, 2, 3$, and $r_{12}(V) = r_{32}(V) = 1$ for all V and $r_{21}(V) = r_{23}(V) = 1/2$ for all V. These assumptions are not made with any thought of reality, but for simplicity in calculation. Note that the first decision made is the minimum of three selections from f(t), which has a probability of occurring in the interval $(\tau_1, \tau_1 + \Delta\tau_1)$ of

$$f(\tau_1)\,\Delta\tau_1 \left[\int_{\tau_1}^{\infty} f(x)\,dx\right]^2 .$$

The probability that node 1 made the selection (that is, sent the first message) is, of course, 1/3. Following this event, the distribution of decisions for nodes 1 and 2 are both $f(t-\tau_1)$, and for node 3 is f(t). The probability that a decision is made in $(\tau_2, \tau_2 + \Delta\tau_2)$, and that node 3 reaches a decision before either 1 or 2 is

$$f(\tau_2)\,\Delta\tau_2 \left[\int_{\tau_2}^{\infty} f(x - \tau_1)\,dx\right]^2 .$$

In the same manner, the probability that node 2 reaches a decision in the interval $(\tau_3, \tau_3 + \Delta\tau_3)$ before either 1 or 3 reaches a decision is

$$f(\tau_3 - \tau_2)\,\Delta\tau_3 \left[\int_{\tau_3}^{\infty} f(x - \tau_1)\,dx\right]\left[\int_{\tau_3}^{\infty} f(x - \tau_2)\,dx\right]$$

and the probability that 2 sends to 1 is 1/2. Similarly, the final probability is

$$\frac{1}{2}\, f(\tau_4 - \tau_3)\,\Delta\tau_4 \left[\int_{\tau_4}^{\infty} f(x - \tau_3)\,dx\right]\left[\int_{\tau_4}^{\infty} f(x - \tau_2)\,dx\right].$$

Clearly, the probability of the entire event is the product of these quantities. The probability that the ordering occurs in time t is the integral over the four τ's where $0 \leqslant \tau_1 \leqslant \tau_2 \leqslant \tau_3 \leqslant \tau_4 \leqslant t$,

$$\frac{1}{4}\int_0^t d\tau_4 \int_0^{\tau_4} d\tau_3 \int_0^{\tau_3} d\tau_2 \int_0^{\tau_2} d\tau_1\, f(\tau_1)\left[\int_{\tau_1}^{\infty} f(x)\, dx\right]^2$$

$$\cdot\, f(\tau_2)\left[\int_{\tau_2}^{\infty} f(x - \tau_1)\, dx\right]^2 f(\tau_3 - \tau_2)\left[\int_{\tau_3}^{\infty} f(x - \tau_1)\, dx\right]\left[\int_{\tau_3}^{\infty} f(x - \tau_2)\, dx\right]$$

$$\cdot\, f(\tau_4 - \tau_3)\left[\int_{\tau_4}^{\infty} f(x - \tau_3)\, dx\right]\left[\int_{\tau_4}^{\infty} f(x - \tau_2)\, dx\right].$$

Obtaining numerical values for a given f is, admittedly, a laborious task, but it is not beyond ordinarily available computational facilities. Of course, for $f(t) = \lambda e^{-\lambda t}$ which, as we have seen earlier, fits some data very well, an explicit evaluation is possible.

A positive advantage for this technique is that models of a generality and subtlety well beyond general formulation and solution can still be verified. For example, the conditional probabilities might be so modified that if i sends to j at τ, and if by τ', i still has the same information, then i will not send to j at τ'. This, in the above example, would lead to a coefficient of 1/2 rather than 1/4, since in the final step 2 would send to 3 with probability 1.

10. Summary

The principal concern of this chapter has been the analysis of time data in an action quantized situation. It was first pointed out that if the individual latency distributions are assumed known, the group latency distribution results from selecting one latency from each individual distribution and taking the largest of these. A form for this function was given. Second, from a plausible assumption as to the nature of the decision-making process, it was shown that the individual latencies could be expected to be of the form

$$\lambda(t) \exp\left[-\int_{t_o}^{t} \lambda(\tau)\, d\tau\right].$$

Indeed, it was suggested that for some purposes one can take $\lambda \exp\left[-\lambda(t - t_o)\right]$, and an attempt was made to fit the data by a χ^2 technique using this assumption. The fit proved to be good, but it was pointed out that the very selection process used to get the group latency makes the result very insensitive to changes in the rising limb of the individual latency distribution. Finally, an analysis of variance indicated that the pair of assumptions we made accounted for the data in a simple fashion.

These results are considered of importance only insofar as they are a step toward an analysis of group communication when decisions and times for decisions are not independent. It is surely inviting in any model building of the complex situation to attempt to use the exponential function which is mathematically tractable. On the other hand, it is almost certain that it is only the falling limb that is approximately exponential, and that the rising limb is far from vertical. Without the selection process of action quantization this will be important, and an error of some magnitude will be introduced. Some care will have to be taken in this matter.

The final section of the chapter was concerned with the nonindependent case. The mathematical difficulties are very great, but one computationally feasible suggestion was offered as a partial solution. Essentially, it is based on the notion that in some situations there are a comparatively few orderings of message flow that account for almost all of the cases. It is possible, for any reasonably simple set of assumptions about the nature of the interaction, to estimate the probability of a specific ordering. Whether this will in fact be satisfactory remains to be seen.

CHAPTER V – NOISE

1. Introduction

In previous chapters, the concepts of communication, information, and noise have been defined and discussed (see sec. I.2.3). The treatment of group learning and the time basis for group action and group learning in Chapters III and IV, however, has been entirely for the noise-free case. In this chapter we will discuss some of the effects of noise on the task-oriented group, and present data from one series of experiments (No. 3 and No. 6) in which the communication was noisy. It should be emphasized at the outset that very little experimental work has been done in this area; most experiments which have been performed in the laboratory on task-oriented groups have either been noise-free or have been considered noise-free in the sense that the presence of noise was ignored in the treatment of the results. Some experiments using very small groups and oral communication, with measured amounts of acoustic noise, have been performed by Heise and Miller (29), but little work has been done with other forms of noise. This lack of experimental work can be attributed to the difficulties encountered in any attempt to measure or control noise in an experimental situation, particularly in the most interesting areas of semantic and coding noise. The experiments reported in this chapter are subject to these difficulties, since they were originally conceived with an entirely different purpose in mind, and the discussion in this chapter is based only on the data recorded, which does not include any attempt to measure the noise involved.

In spite of the difficulties inherent in experimental work in this area, it remains an extremely important field for theoretical and experimental development, since, in general, all real situations are more or less noisy. The "noise-free" case is a fiction, which real situations may approximate more or less closely. No application of theoretical or experimental results to practical situations can succeed unless the effects of noise are taken into account in some way, and we feel that definitive experiments in this field are badly needed.

The experimental results reported in this chapter involve the intuitively clear concept of "errors" on the part of the group. The question of errors has not been previously discussed, and the experimental problems described in previous chapters have generally been too simple for errors to arise. When the conditions under which the group operates includes some provision for ending the experiment, i.e. if a certain feedback from the environment is specified as a "stop" signal, and the experiment ends whenever this signal is received, this provision can be regarded as part of the boundary conditions under which the group operates. Such conditions allow the possibility that the experiment will end before the group has succeeded in fulfilling all the requirements of the task. In such a case, the ending of the experiment without full completion of the task is called an "error." Obviously, certain real situations include this possibility; very often the task includes some sort of time limit, and the group may be unable to meet all the conditions of the task before the expiration of this time.

This chapter presents some consideration of various forms of noise and their effects on the performance of task-oriented groups – in particular, the relation of the noise present and the network used to the relative frequency of errors.

2. Noisy Communication in the Task-Oriented Group

2.1. Definitions

In section I.2.3 the concepts of communication, symbol contents, symbol designs, and message were discussed, and information was treated in terms of mathematical information theory. Certain deficiencies in the theory in regard to the study of task-oriented groups were pointed out, and the application of these concepts to symbol contents and the semantic communication problem as well as some of the problems concerning psychological information theory were discussed. Keeping in mind the limitations on these extensions of mathematical information theory, we will discuss here the treatment of semantic or symbol-content noise in a restricted case. In general, we will use the terms "symbol" or "channel" noise to refer to noise which affects the symbol designs during the transmission process, and the terms "semantic" or "coding" noise to refer to uncertainty arising in the assignment of symbol content to a particular symbol design. We have assumed that the reader is familiar with conventional information theory and will make free use of its definitions and theorems.

2.2. Channel Noise

We will first consider the effects of channel noise in a task-oriented group. Following Shannon, we shall describe channel noise by a set of conditional probabilities. In the case of a discrete noisy channel with a finite number of states, we have a set of probabilities $p_{\alpha, i}(\beta, j)$, which are the conditional probabilities that if the channel is in state α and symbol i is transmitted, symbol j will be received and the channel will be left in state β. Thus, α and β range over all possible states for the noisy channel, i over all possible transmitted symbols, and j over all possible received symbols. This most general case results in great complexity when any attempt is made to handle a task-oriented group with noise present; simplifying assumptions will usually be needed in any practical example. For example, we may assume that the set of symbols S from which the transmitted messages are drawn, and the set S' which the receiver may get, are equal; thus, in the expression above, i and j range over the same set of symbols. We may also assume that the channel has a single state in regard to noise. This second assumption is the more limiting one, but in many cases of interest this assumption is not unrealistic. One further assumption is appropriate in most cases: In general, the characteristics of the noise depend on the channel so that the probabilities mentioned above will be different for each link of the communication network. Consequently, we actually have P_{ij}^{kl} as the probability of receiving symbol j when symbol i is sent over the link from node k to node l; this probability will be a function of both k and l. While many examples of interest exist, particularly in applied situations in which an important characteristic of the noise is its variation from link to link, considerable simplicity in this discussion is gained by assuming P_{ij}^{kl} invariant over all links in the network, so that we may write the same probability P_{ij} for any link. Even with this simplifying assumption the introduction of noise in a network greatly complicates the problem of theoretical consideration.

Under these assumptions suppose there is a series of nodes forming a chain of length n; that is, nodes 1,2,3,. . ., n, connected so that node 1 has a symmetric link to 2, and n-1 has one to n, and node k is linked symmetrically to nodes k-1 and k+1. Assume that the P_{ij} are given. If state E_i refers to the receipt of symbol i by a node, then the transmission of a message from node 1 to node n forms a Markov chain. The P_{ij} are then the transition probabilities from E_i to E_j. We know from existing work on Markov chains that if we define $P_{ij}^{(k)}$ as the probability of symbol i being received as symbol j after passing over k links, we have

$$P_{ij}^{(k)} = \sum_{\mu} P_{i\mu} P_{\mu j}^{(k-1)}$$

and

$$P_{ij}^{(1)} = P_{ij}$$

where μ ranges over S.

For finite Markov chains algebraic methods may be used to determine values for $P_{ij}^{(n)}$ (81). In all situations of interest to us at present, the networks are sufficiently small so that the algebraic manipulations needed to calculate the $P_{ij}^{(n)}$ are simple enough to be useful. By combining this calculation with the methods previously mentioned for calculating decision probabilities for nodes which are branch points, we may be able to extend this Markov analysis to arbitrary networks. This requires that the entire process be Markov, in which case it is essentially the problem of the random walk in n dimensions.

It should be mentioned, however, that it is possible for the transmission of symbols and the decisions to be Markov processes, but for the flow of information through the net to be a non-Markov process and hence not susceptible to ordinary methods of analysis.

The essential feature of this discussion is that channel noise in itself, as applied to noisy networks under certain simplifying assumptions, does no more than complicate an already complex problem. The features of the noisy channel which are of interest to us in the network can be treated by known methods.

One further question of interest in the noisy case is that of channel capacity. In the noise-free case, it was assumed that the time needed for the group to circulate a given amount of information, or for a node to transmit a given amount, was dependent only on the transfer functions of the individuals. In the noisy case we have the additional limitation of channel capacity. If we know the entropies of the sources for a given link, and the characteristics of the noise in terms of the conditional entropies of the corresponding receivers, then the channel capacity is given by

$$C = \text{Max } H(y) - Hx(y) .$$

Also, we know that for any source with entropy $H \leq C$, there exists a coding system which will transmit information at an average rate of H bits per sec with an arbitrarily small frequency of errors. This is not true for any source with entropy $H > C$, so channel capacity poses an additional limitation on a network. Also, in the noisy case, if redundancy is used to attain this maximum possible rate, a delay must generally be introduced at the receiver since a large sample of the transmitted signal must be received before judgment is made as to which signals have been distorted. In many cases, both of these factors may be neglected, but examples will exist where one or both play an important part in the behavior of the network.

Very little experimental or theoretical work has been done to investigate the effects of different networks in a problem involving channel noise. The most relevant contribution has been by Heise and Miller (29). Until more has been done in this line, the general considerations given in this section will point out the complexities introduced by channel noise, which is the simplest form of noise to treat from the experimental or theoretical point of view.

2.3. Coding Noise

Semantic or coding noise, as defined above, may in certain cases be treated in exactly the same manner as channel noise. If we can define a finite set R of physical objects, to which all semantic symbol contents refer (possibly two sets, R and R', for the source and receiver) and also define sets S and S' of semantic symbols (words, phrases, and the like) used by the source and receiver to code these referents, we may consider the two coding processes and the transmission channel as one channel and calculate noise either on the basis of conditional entropies or sets of probabilities that a given referent, chosen by the source, will be received correctly by the receiver. In this case the considerations of the previous section apply in toto and will not be repeated here.

In the general case of coding noise, however, other factors are present. As in the case of channel noise, coding noise may be combated by the use of redundant transmissions, but this redundancy will be in the form of alternate coding schemes for a given referent. In any practical use of these considerations, important differences between the channel and coding noise case become evident. In the case of channel noise, a given use of redundancy in transmissions has the same effect regardless of circumstances, but with semantic noise the peculiarities of the individual subject become important, i.e. differences in vocabulary, and so forth. The relative position in the network also plays a part. In subsequent sections we present experimental evidence for differences between networks in the presence of semantic noise. Many of the differences pointed out do not exist in the noise-free case, and it is doubtful whether they would all be present in the channel noise case. For instance, the effective use of redundancy seems to depend, among other things, on the presence of a two-way link. This effect has no parallel in the channel noise case and serves as an example of the factors which are important in the semantic noise problem and which are as yet imperfectly understood.

Difficulties in experimentation and measurement in the semantic noise case have already been pointed out in section I.2.3. These difficulties stem largely from the absence of any absolute measuring methods, and the experiments reported here are subject to these difficulties.

2.4. General Considerations

In section I.2.3 certain properties of what was termed "psychological information theory" were pointed out: questions of the unexpectedness of a message, its estimated truth value, and the like. These considerations form a further extension of the concepts of information theory as it is currently known, and very little work has been done in this area as yet. In general we may say that as the concepts of information theory are extended from the problems of channel noise to coding noise and ultimately to "psychological information theory," the theoretical treatment comes closer and closer to being a theory of "information," as we intuitively understand the term. However, at the same time, progress in this direction introduces increasing theoretical complexity, increases greatly the number of relevant factors, and places increasing reliance upon psychological measures of the individual. Experimentally, we can deal with channel noise, and we have begun to deal with coding noise, although much is lacking in the theoretical side of this treatment. Any further progress toward a theory of communication and information which takes into account all the factors awaits considerable theoretical and experimental work. In particular, since communication and the exchange of information is essentially a group phenomenon, a theoretically full treatment of these problems must be an integral part of a theory of behavior of groups and must be developed as such. The work reported here is only a small beginning in this direction.

3. Experimental Results

3.1. Experimental Conditions

In Experiments 3 and 6, each subject was given at the start of each trial a box of colored marbles. They were informed that there was one color common to all boxes and that their task was to determine this color. Communication was by written messages, with no restrictions on content. The subjects were also informed that after they had signalled their answer to the experimenter (by dropping the desired marble down a tube), they were free to change this answer by dropping another marble at any time until the trial ended, and that the trial would end as soon as all five subjects had dropped a marble, whether it was the correct one or not.

During the first 15 trials, the marbles used were of plain, solid colors, easy to distinguish and to describe. On the 16th trial and thereafter, the marbles used were of cloudy, mottled, indistinct colors. They were still easy to distinguish if they could be directly compared, but it was very difficult to describe each one clearly and unambiguously.

The experiment consisted of two phases separated by about 18 months. The second phase (Experiment 6) was run in an attempt to answer some of the questions raised in the analysis of the first (Experiment 3). Four groups of five subjects were run on each of three networks in Experiment 3, and four groups were run on three networks by this laboratory in Experiment 6, two of the networks coinciding with two of those originally used. The networks used are listed in Table V.1.

Table V.1

Experiment	Code	No. of Groups	Network	Run By	Remarks
3	C	4	Circle (0)	S. L. Smith	M.I.T. subjects
	Ch	4	Chain (0)	S. L. Smith	M.I.T. subjects
	S	4	Star	S. L. Smith	M.I.T. subjects
6	C'	4	Circle (0)	G.N.L.	Military subjects
	SF	4	Star	G.N.L.	Military subjects Additional feedback
	P	4	Pinwheel	G.N.L.	Military subjects

The second series of three networks were run under conditions approximating as closely as possible those used in the original experiments, but various differences in procedure may have occurred in spite of these precautions. In addition, the second series were run with military subjects from Fort Devens and the First Naval District Receiving Station, rather than with volunteer M.I.T. undergraduates, and this difference undoubtedly influenced the results. A specific change occurred in the star groups SF in that at the end of each trial the experimenter gave the group feedback as to the errors made. This variation will be discussed in detail in section V.6.

It should be emphasized that considering the fairly large variation between different groups on the same network, the four groups run on each network are not nearly enough to establish any sort of reliability in the data. Consequently, the experimental results presented are extremely tentative in nature; the internal inconsistencies in the experiments and the small size of the samples prevent any high degree of confidence in the results. It is the hope of this laboratory to run a similar experiment in the very near future, using revised techniques and far larger samples. At present, the data presented here, although extremely tentative, serve as an interesting illustration of the effects involved and give a good indication of the problems and complexities in this area.

3.2. Noise in the Experiment

The noise occurring in this experiment was present only in the last 15 trials. Two different subjects would often use different terms to describe a given marble, and the same word would often be applied by different subjects to two or three different marbles. Such confusion or ambiguity is an example of the noise discussed in previous sections as encoding-decoding noise.

This method of producing noise experimentally has several drawbacks. In the first place, it is almost impossible to control, since the amount of confusion varies greatly from subject to subject, group to group, and even from marble to marble. In the second place, it is impossible to measure the amount of noise present at any time, even approximately, since we cannot estimate the entropy of any subject, or the "average" subject, as a semantic-information source, and we are even less able to determine the semantic entropy of the source as seen by the receiver.

It is evident from observation of the experimental groups that the amount of noise actually present was dependent on the I.Q., vocabulary, and previous experience of the subject, and particularly on his color vision. All subjects were tested for color vision using the standard Ishihara plates, and those who were color-blind were eliminated, but this procedure by no means guarantees uniform color vision among those who passed the test. In addition, subsequent examination of the messages sent gives the impression that several of the subjects used, although they passed the test and hence were not in that sense color-blind, had definitely peculiar color reactions. As a result of these difficulties, it is almost impossible to give any exact treatment of the noise in this experiment from a quantitative point of view, but the most casual study of the data leaves no doubt that semantic noise was present.

3.3. Evaluation of Errors

All of the errors occurring in this experiment consisted of one or more of the subjects having registered the wrong marble at the end of the trial. If a subject had dropped the wrong marble and had corrected it before the trial ended, the correction was accepted as a correct answer.

In Experiment 3, the errors were originally counted on all-or-none basis; that is, if everyone in the group had dropped the correct marble, the trial was scored "correct," but if one or more subjects had dropped an uncorrected wrong marble, the trial was scored "error." For the purposes of this chapter we considered that a more sensitive tabulation would be more suitable, and for each trial we scored the group 0, 1, 2, 3, 4, or 5 depending on the number of subjects who had dropped an uncorrected wrong marble at the end of the trial. Among other considerations, some networks tended to produce either all correct or all wrong answers on any one trial, whereas others had a higher proportion of trials with one or two subjects having incorrect answers. The original "all-or-none" scoring method tended to equate these two effects, whereas the second method brought out the differences between the networks.

In tabulating errors, the data for each group were lumped in blocks of three trials: 16, 17, 18; 19, 20, 21; and so forth. If we assumed that the characteristics of the group with regard to errors changed relatively slowly as the trials progressed, this method tended to eliminate some of the fluctuations which were caused by the small sample size.

The errors made by the groups were also tabulated as a function of the marble which was held in common, with the figures lumped for all groups run. This plot showed a considerable variation in the number of errors, depending on the marble which was the correct answer. Apparently, some marbles were more difficult to describe than others. The tabulation of errors has been corrected for this effect in order to bring all blocks of trials to the same basis. Figure V.1 presents the uncorrected tabulations, and Figs. V.2 and V.3 give the corrected relations. Figure V.4 presents the error data scored on the original "all correct or all wrong" basis.

Since there remained a large amount of fluctuation in the error figures from one block of trials to the next, it was considered advisable to test the significance of any differences which occurred. Using Students "t" distribution, we may establish 95-percent confidence limits. These are plotted in Fig. V.5. In this figure the rectangles indicate the areas within which the true population mean is expected to lie in 95 percent of the cases, so that those groups whose rectangles do not overlap for a given trial block are significantly different at better than the 5-percent significance level with regard to the error count.

When the corrected error plot is examined with these confidence levels in mind, several points emerge as significant: (a) None of the networks show any appreciable differences during the first fifteen relatively noise-free trials, in which the solid color marbles were used (see Fig. V.1). The differences among these networks are evident only in the

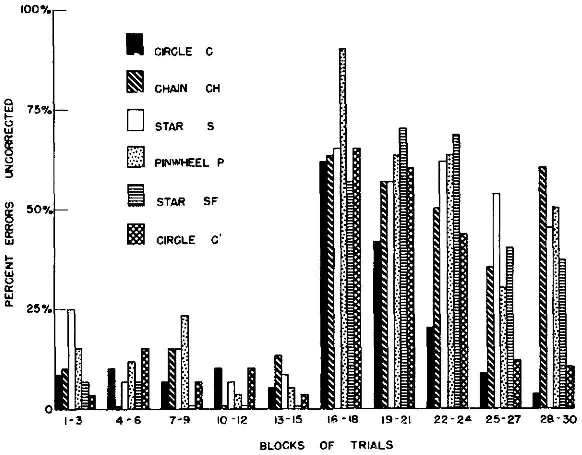


Fig. V.1

Uncorrected error count, Experiments 3 and 6.

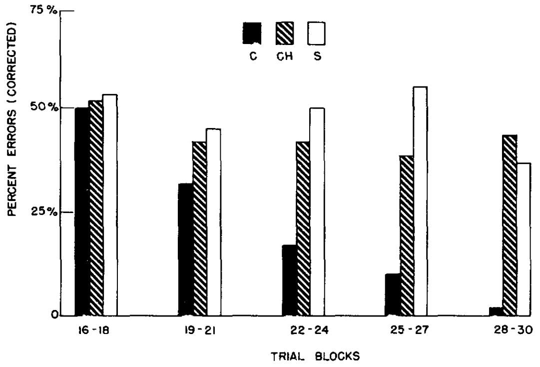


Fig. V.2

Corrected error count, Experiment 3.

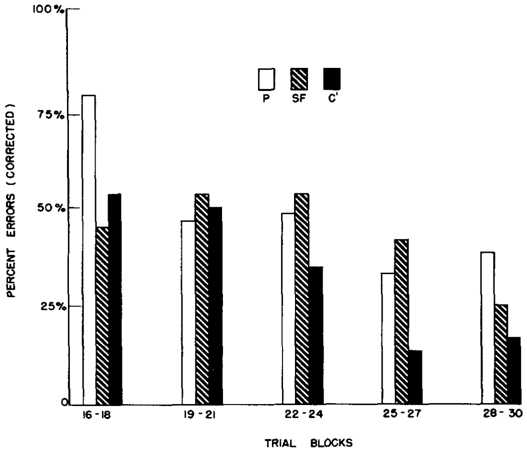


Fig. V.3

Corrected error count, Experiment 6.

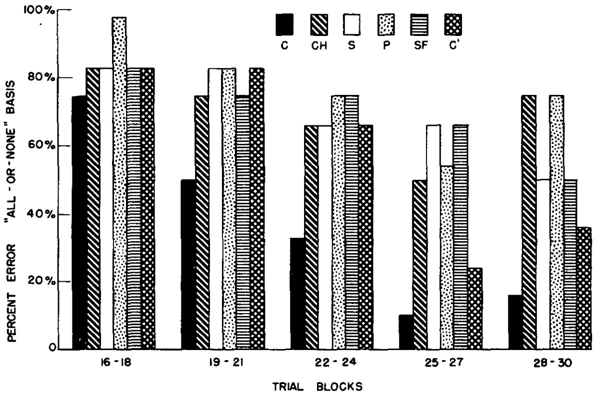


Fig. V.4

Group error count, Experiments 3 and 6.

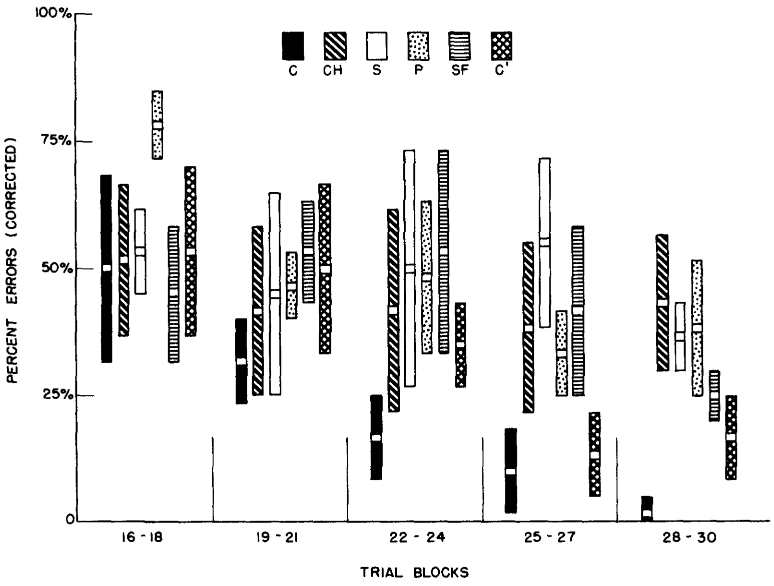


Fig. V.5

Corrected error count and significance limits, Experiments 3 and 6.

presence of noise. (b) Both circle groups, C and C', are significantly more accurate than all other groups in the last six trials. (This is true for the original circles, in which M.I.T. subjects were used, for the last nine trials.) (c) The pinwheel is more subject to error in the first few trials than all other groups. (d) The star with additional feedback is more accurate than the original star or chain groups in the last few trials.

3.4. Conclusions and Hypotheses

From the results mentioned above and from the general trend of the data (even at points where this difference is not significant), we may develop a feeling for the differences among networks. The circle groups seem to learn to reduce their errors more quickly, and after ten to twelve trials they are much more accurate than any other group.

We have assumed that the difference between the first and second groups of circles (C and C') arises from differences between the subjects used. Therefore, it should be noted that the P and SF groups should properly be compared with the C' groups rather than the C groups, because of this subject difference.

The original star and chain groups show relatively little learning and remain at a high level of errors throughout. The pinwheel groups start with an extremely high error count, and while they display some learning, they are only able to bring their accuracy up to approximately that of the star and chain groups. The star with additional feedback has about the same error count as the star and chain in the first ten trials, but in the last few trials it learns to be considerably more accurate than the original star, chain, and pinwheel.

In the next few sections of this chapter we will examine the mechanisms of these effects and will discuss some of the questions raised by the differences noted above. Among these questions are the following: Why do the circle groups display learning and reduce their errors more effectively than any of the others? Why do the star, the chain, and the pinwheel continue at a high error level, and why does the star with feedback show learning toward the end of the experiment? Why is the pinwheel unable to reduce its errors further, in spite of the initial period of learning displayed? While we cannot provide complete answers to these questions, we have been able to gain some insight into the effects of noise on these networks and the mechanisms used to reduce it, and we have given some very tentative explanations for the differences observed between networks. More than this is impossible with the existing data, but the available material points to experiments involving noise as one of the most interesting for future research.

4. Measurement of Noise

4.1. Conventional Measurements

It would obviously be convenient to discuss and measure the noise in these experiments in terms of existing information theory concepts. Unfortunately, several difficulties arise when this attempt is made. The noise in this experiment is not of the symbol or channel type, but is noise of the second kind, that is, semantic or coding noise. While information theory measures can still be applied to this type of noise, such application requires knowledge of several facts unavailable in the present case. We are unable to define the semantic symbol space S, and more important, we have no way of observing the transformations from the set of referents R to the messages formed from units of S. Nor can we observe the corresponding transformations from the symbol units of the received messages to the receiver's referent set R'. Consequently, although we may discuss the noise in this experiment in terms of information theory in a qualitative manner, we are unable to arrive at the usual quantitative measures. Nevertheless, such a measure is needed, and it may be achieved in an approximate manner by considering more closely the characteristics of the noise occurring in the experiment.

4.2. "Ambiguity" as a Measure of Noise

Since noise is fundamentally a question of uncertainty, any single-valued measure of the amount of uncertainty in an experiment can be expected to be monotonically related to the noise. In Experiment 3, the uncertainty arose largely from different subjects applying the same name to different marbles, with the result that comparing the names used by each man to describe the marbles he had of that trial led to several possible answers or, in some cases, to a single incorrect answer. Specifically, during the first fifteen trials, in which solid color marbles were used, the groups generally learned to refer to each marble by a single color name, such as "red," "black," and the like. After the sixteenth trial, even though the marbles used were mottled and streaked, often with more than one color or with shades of one color, this behavior persisted. The subjects usually attempted to use such one-word color names as "amber," "aqua" or in some cases such compound words as "light-green" or "blue-green" to describe the marbles, and the noise in the experiment lay in the coding-decoding process of translating a name in terms of the actual marble to which it referred.

Thus, it is reasonable to attempt to measure the uncertainty by counting the number of marbles referred to by a given name in any one trial, and averaging this number over all the names used. Since some names were used far more frequently than others, a weighted average was indicated.

The following procedure was used to calculate this uncertainty, which will be referred to as "ambiguity," or marbles per name, and denoted by the letter A. For each trial, the message cards sent by any subject were examined, and in all cases in which a definite assignment of names to marbles could be made on the basis of the experimenter's knowledge of the marbles in each man's box, this information was tabulated. In most cases, this method resulted in considering only those cards sent which listed the marbles in the subject's box for that trial. From the results for all five subjects for that trial, lists of names which had been used to describe each marble were compiled, with the frequency of occurrence of each. From these lists a master score sheet was prepared for each group, which listed the number of marbles referred to by each name at any given trial, the number of names used to describe each marble, and their relative frequency of occurrence. The weighted average of the number of marbles referred to by a single name was calculated, and this was assigned as the value of A for that group at that trial. This procedure was followed for trials 16 to 30, for all the groups run. From this master sheet a corrected value of A was prepared as follows: If a given name was used to describe two different marbles on trials i-1 and i+1, but specific evidence for this confusion could not be found during trial i, we assumed that it was present on the strength of its occurrence before and after trial i. From these corrected values of A, an average value was computed for each network during each block of three trials. These results are presented in Fig. V.6.

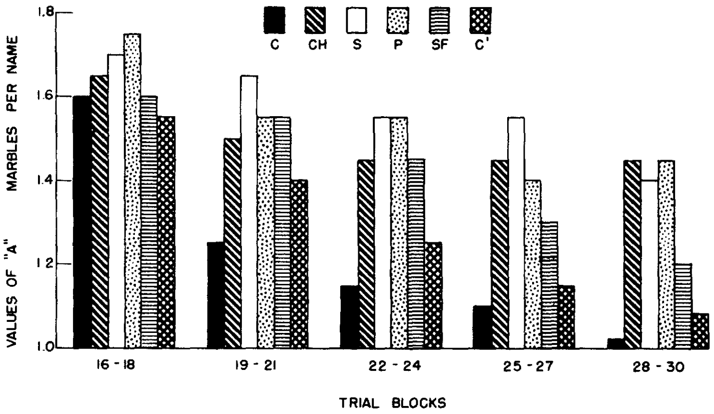


Fig. V.6

Average ambiguity of messages, Experiments 3 and 6.

An examination of this graph shows that the values of A are roughly those which would be expected if A is considered as a measure of noise and if the errors made in the experiments are considered due to the noise. The chain, the star, and the pinwheel groups start at a high level of A and show very little reduction as trials progress. This behavior is in conformity with their error scores. The circle groups (C and C') show a steady reduction of ambiguity, with C somewhat ahead of C'. This agrees again with their error results. The star with additional feedback has a high value of ambiguity during the first nine or ten trials, but shows a reduction in the last five or six trials, again in agreement with its reduction of the error score toward the end of the experiment. From these results we may conclude that the observed values of A serve as a useful measure of noise on the basis that changes in the values of A have roughly the effect that changes in the noise level would be expected to produce.

It is possible to relate A to the semantic noise level calculated on the basis of source and receiver entropies if a few very questionable assumptions are made. Suppose that the source in all cases uses a code of six symbols to refer to the six marbles, and that each of these six symbols occur equiprobably; then we may calculate the entropy of the source

$$H(x) = - \sum_{i} p(i) \log_2 p(i)$$

if p(i) = 1/6 for all i, H(x) = 2.59 bits per symbol. Note that this assumption, while perhaps fairly close to the actual case during the first fifteen trials, is definitely not accurate during the last fifteen trials; the subjects use many more than six symbols to refer to the six marbles, and they do not occur with equal probability. Nevertheless, we may gain some insight into the relationship of A to the semantic noise as defined by entropy if we proceed on this basis.

Since the observed average value for A lies between 1.0 and 2.0, let us assume that all the ambiguity for each received semantic symbol lies in a choice between two possible referents. If a symbol which encodes α is received and A is given, the probability that the referent α will be chosen is

$$P_\alpha(\alpha) = \frac{1}{2}(3-A)$$

which is a linear function. Hence, when A = 1.0, $P_\alpha(\alpha) = 1.0$ and A = 2.0, $P_\alpha(\alpha) = 1/2$. If β is the other possible referent, then

$$P_\alpha(\beta) = 1 - P_\alpha(\alpha).$$

This is definitely an inaccurate assumption, since the marble chosen, when the symbol α is received, is often picked from a set of three or four marbles. However, since these secondary choices are not distributed in an even fashion over the six marbles, it does not seem possible to improve these assumptions on the basis of the existing data.

The average conditional entropy of the receiver, when the source message is known, is defined as

$$H_x(y) = - \sum_{i,j} P(i,j) \log_2 P_i(j)$$

$$= - \sum_{i} P(i) \sum_{j} P_i(j) \log_2 P_i(j)$$

where $P_i(j)$ is the probability of picking marble j when symbol i is sent. We have assumed

$$P(i) = \frac{1}{6}$$

$$P_i(i) = \frac{1}{2}(3-A)$$

$$P_i(j_1) = 1 - P_i(i) = \frac{1}{2}(A-1)$$

$$P_i(j) = 0 \text{ for } j \neq i,\ j \neq j_1.$$

Using these values in the expression above, we get

$$H_x(y) = \sum_{i=1}^{6} \frac{1}{6} \left[(\frac{3}{2} - \frac{1}{2}A) \log_2 (\frac{3}{2} - \frac{1}{2}A) + \frac{1}{2}(A-1) \log_2 \frac{1}{2}(A-1) \right]$$

$$= 1 - \frac{1}{2} \left[(3-A) \log_2 (3-A) + (A-1) \log_2 (A-1) \right]$$

In Fig. V.7, $H_x(y)$ is plotted as a function of A from 1.0 to 2.0. The relationship between these two measures is not linear and would probably be even less linear if calculated on the basis of a knowledge of the actual number of referents among which each choice was made.

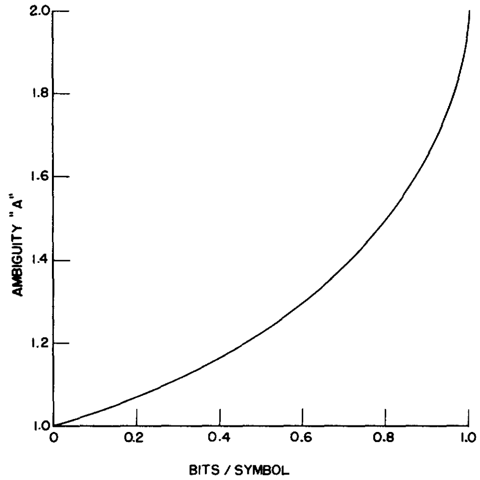


Fig. V.7

Ambiguity vs semantic noise, $H_x(y)$.

The calculation given above, of course, does not constitute a measured value of $H_x(y)$; the values of A are measured, but $H_x(y)$ was calculated from A only after making several assumptions about the coding processes which cannot be verified from the data. Future experiments will be designed to make direct measurement of H(x), $H_x(y)$ and H(y) possible.

4.3 Errors as a Function of Noise

It is possible by a few calculations based on some simplifying assumptions, to confirm the hypothesis that the errors in this experiment are directly related to the noise level as measured by A. Consider a node which has received sufficient messages to reach a decision on an answer. Uncertainty in the coding process renders this decision ambiguous. We will assume that the final choice is between two marbles and that the probability of picking the correct one is $P = 1/2(3-A)$, as above. The answer must be sent to the other members of the group, and we assume they each have probability P of picking the right marble when sent a description of it. So the first man to get the answer has probability P of dropping the right marble, and the others have probability P^2. The expected percentage error will then be

$$P.E. = \frac{100}{5}\left[4(1 - P^2) + (1 - P)\right]$$

This predicted error is plotted against ambiguity as the solid line in Fig. V.8. If we assume, on the other hand, that each member of the group drops a marble independently of the others, with probability P of getting the right one, then the expected value of the percentage error would be $100(1 - P)$. This is plotted as the dashed line in Fig. V.8. The plotted points in this figure are the observed percentage error for various ambiguity values. These points fall along the solid curve (the first assumption) quite closely, considering that it was not a fitted curve. This supports the assumption that errors are a function of ambiguity.

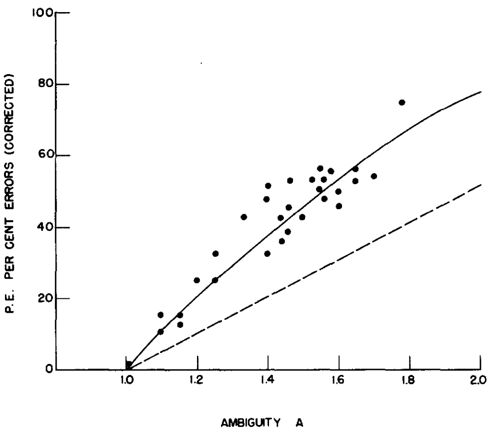


Fig. V.8

Theoretical and observed percent error vs ambiguity, Experiments 3 and 6.

Using the same values of P, we may predict the error curve for all-or-none scoring. Under the first assumption, the expected number of all-or-none errors is

$$P.E.' = 100\left[1 - P(P^2)^4\right]$$

which is plotted against the ambiguity A as the dashed line in Fig. V.9. On the other hand, if we assume all members act independently with probability P of a right answer, the expected number of all-or-none errors is

$$P.E.' = 100\,(1 - P^5)$$

which is plotted as the solid curve in Fig. V.9. In this case the observed values, plotted as points, lie along the solid curve representing the independence assumption. These results show that neither of these two simple assumptions give a completely accurate picture of the group process. Considering the crudeness of the assumptions, this is not surprising. The occurrence of errors as a function of the ambiguity is clear, however.

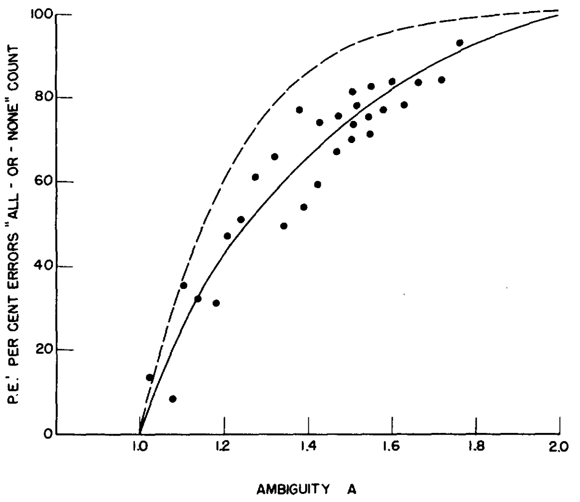


Fig. V.9

Theoretical and observed group errors vs ambiguity, Experiments 3 and 6.

We may then conclude, on the basis of the analysis of these data at this point, that we have established the nature of the noise in this experiment, that the errors in the performance of the group are caused by this noise, and that a knowledge of the amount of noise present, in terms of A, enables one to predict in a straightforward manner the average percentage of errors made by the group. The fact that the assumptions under which this prediction was actually made were not very accurate is not too relevant; the fit obtained by the predicted curves shows that the assumptions are about as accurate as the error data itself, which is all we need ask of the prediction at this point. New experiments, with better data, will permit more complex and accurate assumptions about the relationship between the noise and the percentage of errors. For the present, we can only establish in general terms the role played by noise in this experiment.

5. Redundancy and the Reduction of Noise

5.1. The Use of Redundancy in the Coding Process

Since certain networks manage to achieve a reduction in ambiguity, and hence in their error level, one is led to inquire about the mechanism of this effect. This problem may also be approached by an application of the concepts of information theory, extended to fit this case.

In the conventional case of signals transmitted along a channel, accurate transmission in the presence of noise is achieved at the expense of the transmission rate by the introduction of redundancy; that is, repetition of a given message or the use of symbols in the original which are predictable from others. We shall examine the present case for a similar mechanism. Of course, since the noise here is semantic noise, we shall have to look for semantic redundancy, i.e. duplications in the coding scheme. In our case, these duplications, if they exist, will take the form of synonyms or alternate descriptions of a given marble. We shall show that these duplications do exist and that they are used to overcome the noise.

Since the noise present in this experiment is semantic noise and is measured by ambiguity A, the use of redundancy to overcome the noise and insure accurate transmission of the message will also effect a reduction or elimination of the apparent ambiguity or uncertainty present. This will be reflected in a decrease in the measured value of A. In a sense, this case is not an exact parallel to the usual case of channel noise, since with channel noise the introduction of redundancy in the coding does not remove the noise, but merely removes the errors caused by the noise. Hence, in the symbol or channel noise case, the redundancy must be maintained at a high level in order to insure accuracy. This is not the case with semantic or coding noise, for once the uncertainty in the coding operations has been eliminated, the redundancy may then be reduced without impairing the accuracy of the transmissions. However, semantic noise may also be thought of as having the constant character of channel noise by considering the effect of memory. Once the redundant coding has been used, and the errors reduced thereby, we may assume that the receiver remembers the synonyms used for a given symbol in the redundant code, and that in future messages these synonyms or alternate codes are understood even though not physically present. If the effect of this understood or remembered redundancy is assumed, we may describe the system as one with constant noise but with the effect of the noise overcome by the redundant coding, just as in the channel noise case.

5.2. Measurement of Redundancy

To detect semantic redundancy, we use a method analogous to that previously used to calculate ambiguity. In any one group and at any one trial, six names are sufficient to identify the six marbles. By tabulating from the message cards the names used by the group to describe a given marble, we obtained a record of synonyms or alternate codings used in each trial by each group. This tabulation was corrected for the ambiguity of some of the names used, on the basis that a synonym which was also applied to two other marbles should not be counted as a separate synonym for each, so the tabulation was weighted according to the ambiguity of each term. The table was also corrected for missing data, i.e. for a synonym which was used before and after a given trial, but evidence for the use of which could not be found with certainty during the trial. From these tabulations the average number of names used by each group during each trial was calculated, and from these figures the average number of extra names (that is, the number of names used beyond the necessary six) was calculated. These values were then averaged over all the groups run on a given network and over blocks of three trials apiece, as previously. It is worth noting that these calculations involved several guesses and approximations; names could not always be assigned to a particular marble, even in the reconstruction. In addition, a certain number of arbitrary decisions were necessary to determine whether two names were actually different; for example, "yellow" and "yellowish" were not considered different names, but merely different forms of the same name. In spite of these approximations, the resulting figures are sufficiently accurate for our purposes, since errors made in their calculation will affect them only to a minor extent. The average number of extra names used is called the redundancy R and is tabulated in Fig. V.10 for the different networks by blocks of three trials.

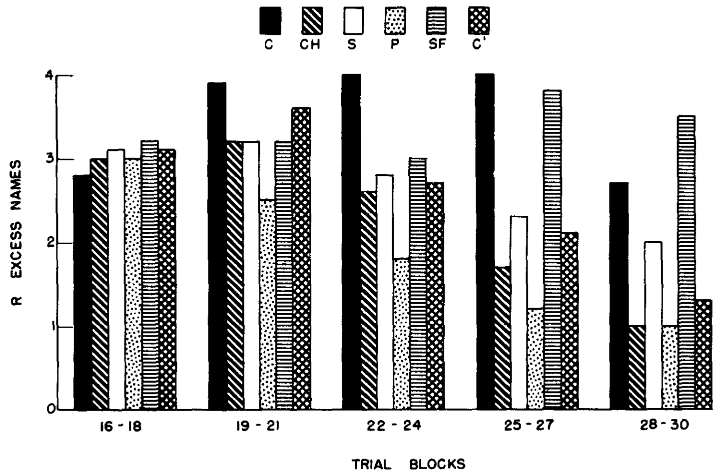


Fig. V.10

Redundancy vs trials, Experiments 3 and 6.

5.3. Redundancy as a Mechanism

Figure V.10 shows several features of interest. The chain, the star and the pinwheel groups start at an R value of about 3 and slowly decrease in R value as trials progress. The circle groups also start at about 3, but show a decided rise in R to about 4 during the middle ten trials, followed by a decrease. These data confirm the hypothesis that the reduction in A is made possible by increasing the redundancy. The circle is the only one of the three original groups to show a decrease in A, and it is the only one of the three original groups to show this rise in R. This use of redundancy as a mechanism for the reduction of ambiguity is supported by the results of the second set of experiments. The pinwheel shows no rise in R and never reduces its value of A or its error count below that for the star and chain groups. The circles (C') show a similar rise in R to that shown by the first group, although the rise is not as high nor of as long a duration. This coincides with the fact that the C' groups did not achieve as good a reduction in A as the C groups, although their performance was qualitatively similar. The SF groups show a sharp rise in R toward the end of the experiment, in agreement with the fact that their value of A was roughly constant until the last five or six trials, when it showed a sharp drop.

We may demonstrate these effects very sharply by plotting a comparison parameter. For each network at each trial block, we define two parameters depending on R and A, as follows: If this or a previous trial block had $R \geq 3.6$, let $\alpha = 1$. If all previous trial blocks had $R < 3.3$, let $\alpha = 0$. If $A \leq 1.4$ for this trial block, let $\beta = 1$. If $A > 1.4$ for this trial block, let $\beta = 0$. Define $\mu = \alpha + \beta$. Then, if a group has reduced its ambiguity by increasing R, μ should be 2. If the group has neither reduced A nor increased R, μ should be 0. An intermediate value ($\mu = 1$) will show contradictory behavior, viz. a drop in A without a previous increase in R, or a rise in R which does not reduce A. The values of μ are plotted in Fig. V.11 for each network and each trial block. The lack of

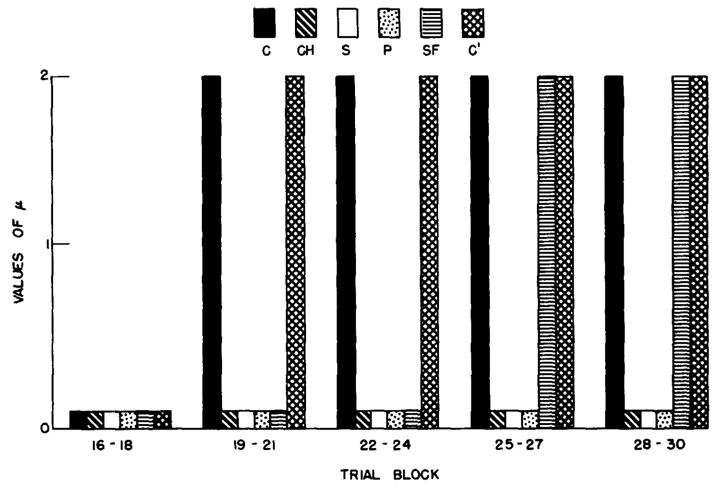


Fig. V.11

Relationship of ambiguity and redundancy vs trials, Experiments 3 and 6.

values of $\mu = 1$ on this graph and the points at which the values of μ jump to 2 indicate clearly that ambiguity is reduced in these groups by increasing the redundancy.

By introducing a measurement of redundancy, appropriately defined, we have been able to demonstrate the mechanism used by these groups to reduce their errors. This mechanism is fully in accord with the concepts of information theory, as extended to the case of semantic noise. We have examined this experiment and answered the question, "what happened?", by measuring the error count and the semantic noise level and demonstrating their relationship. We have answered the question, "how did this happen?", by extending the concept of redundant transmissions to include the semantic noise case. We are now left only with the question, "why did this happen in this way?", i.e. "what produced the observed differences in the behavior of different groups?" We would like to know why the groups run on some networks were able to make use of redundant transmissions to reduce their errors while other networks did not. Unfortunately, we cannot give a clear-cut answer to this last question, but we will discuss the answers and explanations we have produced to date and answer this last question at least in part.

6. The Effect of the Communication Network

6.1. Experimental Effects

In the previous sections, several differences in the behavior of different networks in this experiment have been pointed out. We have shown that the mechanism by which the groups reduced their errors was the introduction of redundant transmissions, but this does not explain the differences in behavior between different networks. There was no reason to suppose, a priori, that the circle would be more successful in reducing its errors than the star or the chain; in fact, certain evidence from noise-free experiments would indicate the star as possibly superior to the circles (34). It is also not immediately obvious why there should be such a difference between the chain and the circle, since their structures are superficially similar.

When the data from this experiment are examined carefully in an attempt to reconstruct the sequence of events and to detect differences in behavior, certain facts emerge which qualitatively give some explanations of the differences between groups. In the star, for instance, it has already been shown that the morale of the peripheral men is very low and that they do not have a strong sense of participation in the group (34). The data show the effects of this. The peripheral men learn during the first fifteen trials that their sole function is to send their information to the center man and then to sit back and wait for the answer to be sent to them. When the noisy marbles are introduced, they continue to act in this manner, leaving all the responsibility of finding the answer to the center man. When he finally sends them his choice for the answer, the descriptions used are usually sufficiently ambiguous so that the peripheral men can find a marble in their boxes which seems to answer this description. The two effects which seem to be of importance here are the lack of a sense of participation on the part of the peripheral men and their ignorance of the confusion which their ambiguous descriptions cause. Consequently, messages from the central node asking for better descriptions are not effective because the peripheral men have no chance to know the deficiencies of their original descriptions, and do not realize the number of errors which are actually occurring, since they see only a small part of the process.

In the circle groups, on the other hand, the work of deciding on the answer is shared more or less equally among the subjects, and messages flow freely around the network; hence, they soon realize both the extent and the source of the confusion and are more easily able to correct it by a joint attempt to clarify their descriptions. The chain groups seem to resemble the stars more closely than the circles in that the structure of the network forces the development of a rigid organization during the first fifteen trials. When the noise is introduced, they are under the same handicap as the stars in attempting to deal with it.

With these considerations in mind, we examined the data from the first series of experiments (Experiment 3) in an effort to formulate more precisely the factors which produced the observed differences between networks.

6.2. Error Feedback

One of the seemingly important differences in method and performance between the circle and the other two groups was the knowledge of the amount and nature of the confusion on the part of all the members of the group. This led to a suggested criterion for predicting the performance of a network on this experiment. We assumed that for any network to be successful, it was necessary to have adequate feedback to each member of the group to acquaint him with the true picture of the group's performance. This meant that the network would have to be such that each member in the natural course of events would handle most of the group's information. Examples of networks in which this is true are the circle and the other networks permitting free flow of information to all members of the group.

In an effort to check this hypothesis, we decided to repeat Smith's experiment with the star network. We decided to give the subjects, in addition, the following information at the end of each trial: the number of errors and the number of different marbles dropped in error. We felt that this information would give them not only a realization of the errors the group was making but would enable them to appreciate the cause of these errors. The groups run under these conditions were denoted SF, and the results of these experiments have already been presented in previous sections. These results formed a partial confirmation of the hypothesis in that this additional feedback significantly improved the accuracy of the stars. However, it was obvious that other factors were also involved, since the SF groups were much slower to reduce their errors than the circles and did not succeed in reducing them as far. Nevertheless, we felt that this experiment demonstrated the importance of the error feedback factor and that this factor is in all probability a necessary but not sufficient condition for good performance in such an experiment.

6.3. Highly Centralized Networks

In attempting to determine other factors in the influence of the network on the performance of these groups, we saw that the notion of some sort of equality between the members of the group played a part. When the circle organizes, it usually does so as a chain, and the central man may be any one of the five with equal a priori probability. In fact, circle groups in which this sort of organization took place were run, and in some of these groups the location of the center man changed from trial to trial.

The evidence suggested that the absence of any centralized structure which would tend to force a rigid, stable organization was a factor in the performance of the circle. The five nodes are all equal, in one sense of the word; this equality was not present in the chain and star. We decided to test this hypothesis on another network, the pinwheel, in which the nodes were equal in this sense. This network is somewhat similar to the circle in that all the nodes are equivalent and any organization can be expected to favor any one of the nodes with equal probability, and indeed the center of the organization may shift from node to node as trials progress. Four groups were run on the pinwheel under conditions duplicating Smith's as closely as possible. The results obtained were presented in the previous sections. The pinwheel does show learning, but only to the extent that it reduces its extremely high error count at the start to about the same value as the S and CH groups at the end.

The existence of learning in the pinwheel is partial confirmation of the effects of equality of position, but these results also make obvious the fact that equality alone is insufficient to assure good performance. The circle obviously has qualities in addition to equality of position, which are absent in the pinwheel.

6.4. Symmetric Linkage

An examination of the properties of the five networks run in this experiment (the C' groups were a complete duplication of Smith's C groups and were run merely as a check on the differences between the subjects used in the two phases) leads to one further fact, in addition to the two mentioned above, which is pertinent to the differences in performance. The circle has symmetric links and the pinwheel does not; that is, in the circle, if A can talk to B, B can also talk to A. This is not true in the pinwheel; therefore, in the pinwheel it is impossible to have a direct exchange of questions and answers. It seems that in the pinwheel position equality and error feedback are present and are sufficient to make the groups realize that ambiguity is present, but that symmetric links are needed in addition to permit efficient use of redundant messages to reduce the errors. The example of the two stars tends to confirm this; the S groups lacked error feedback, and when this was supplied in the SF groups, the presence of the symmetric links made it possible for them to attack the problem with some degree of success. In essence, it seems that questions of the type, "what do you mean by light green? ", prompt the use of synonyms most effectively, and it is just this sort of question which is possible only in the case of symmetric links. Perhaps the lack of equality of position in the SF groups results only in a time delay in the reduction of errors, because the central node must perform almost all the work in reducing errors.

6.5. Conclusions

We are thus led to the belief that no single factor of a geometric or topological nature is responsible for the observed differences in performance, but that a combination of several is responsible. To clarify the effect of these factors, the various networks are tabulated in Table V.2, with the factors possessed by each and the relative performance of the groups using these networks.

Table V.2

Network	Error Feedback	Position Equality	Symmetric Links	Error Performance
Circle	Yes	Yes	All	Learns fast Good error reduction
Star	No	No	All	No learning No error reduction
Chain	Slight	No	All	No learning No error reduction
Pinwheel	Yes	Yes	None	Some learning initially Poor error reduction
Star with Feedback	Yes external	No	All	Slow learning Gradually reduces errors – fair but late reduction

This table summarizes our knowledge to date on the effect of the network in this type of experiment. The three factors we believe important are not stated with the precision we should like, and we are not sure that they are the only important factors, but further work on this will have to await new experiments. We would like to be able to display mathematical relationships between the extent to which these factors are present in a network, as measured by certain topological parameters, and the predicted performance of a group using the network for this general class of problems. Unfortunately, at the present time we are not sure of the factors, and we have not found any topological measures which, either singly or in combination, will describe the presence of any of these factors in a network; hence we are reduced to visual inspection of the network and intuitive judgments of the amount of equality, error feedback, and the like. In addition, we have no relationships which will predict the relative importance of these factors. Eventually, it would be desirable to be able to predict properties of large classes of networks in this respect rather than be forced to examine each network separately as is now necessary.

7. Other Factors and Variables

7.1. Message Content

In an effort to determine the effect of different sorts of messages, a content analysis was performed on the message cards sent in Experiment 3.* Messages were divided into

---

*This content analysis was performed independently by S. L. Smith in 1951 (Unpublished memorandum).

the following categories: (a) Information given concerning task, (b) information given concerning group activity or network, (c) information requested about task, (d) information requested about group or network, (e) orders or suggestions given concerning task, (f) orders or suggestions given concerning group or network, (g) positive valence messages — showing encouragement, and the like, and (h) negative valence messages — showing aggression, criticism, and the like. The messages from the last 15 trials of Experiment 3 are tabulated in these categories in Table V.3. These categories were chosen originally because a high correlation was expected between the relative number of messages not in category (a), or between the relative number of "group-oriented" messages (b + d + f), and the number of errors made by a group. The calculations were made from the data in Table V.3, and the relative numbers in these categories were compared.

These results did not seem to confirm any of the original hypotheses which had prompted the choice of categories; the relative numbers of not–a messages and the relative number of group-oriented messages did not seem to show significant differences between networks.

In an effort to determine any factors determining the reduction of errors, the total number of messages from trials 1 through 30 were tabulated and divided into categories. The following correlations were obtained with the errors made by each group.

Correlation of total messages with errors: -0.75

Correlation of a + b messages with errors: -0.78

Correlation of c + d messages with errors: -0.74

On the basis of these results, the important factors in message content with regard to reducing errors seem to be the number of informational and question messages sent by a group.

On subsequent consideration by this laboratory, however, it was felt that these correlations do not give a true picture of the relative importance of these categories, since they were figured for the entire 30 trials, and since they were straight correlations. It was felt that multiple and partial correlations would give a more accurate picture, and that messages should be counted only for the last 15 trials during which the noise was

Table V.3

	Total Mess.	Totals: Trials 15 - 30								Total	Total	Total	Total
Group	15 - 30	a	b	c	d	e	f	g	h	a + b	c + d	a + b + c + d	e + f + g + h
Circle 1	798	525	18	157	10	34	3	38	13	543	167	710	88
Circle 2	192	155	5	6	4	7	9	4	2	160	10	170	22
Circle 3	289	186	10	37	6	30	7	4	9	196	43	239	50
Circle 4	451	330	3	66	6	10	11	18	7	333	72	405	46
Chain 1	235	169	3	28	2	16	0	13	4	172	30	202	33
Chain 2	543	323	26	82	17	29	16	32	18	349	99	448	95
Chain 3	466	214	18	78	9	45	6	58	38	232	87	319	147
Chain 4	156	131	0	7	0	17	0	1	0	131	7	138	18
Star 1	149	127	0	10	0	4	0	5	3	127	10	137	12
Star 2	182	138	2	14	1	8	0	10	9	140	15	155	27
Star 3	173	142	0	20	0	5	1	3	2	142	20	162	11
Star 4	203	139	1	13	4	22	0	18	6	140	17	157	46

present. With this in mind, tabulated results were divided into classes or types, as listed in Table V.4.

Table V.4

Type	Factors Included
1	Total Error Count; Trials 15 – 30
2	Total No. of Messages, all categories; Trials 15 – 30
3	Total No. of a, b, c, d Messages; Trials 15 – 30
4	Total No. of e, f, g, h Messages; Trials 15 – 30

On this basis, correlations were figured as follows:

$r_{12} = -0.53$ $r_{23} = +0.98$

$r_{13} = -0.58$ $r_{24} = +0.75$

$r_{14} = -0.32$ $r_{34} = +0.61$

These correlations indicate that the number of messages in different categories were strongly related. It is also noticeable that the correlation between total messages and errors drops from - 0.75 to - 0.53 when only the last 15 trials are considered, and the correlation between a, b, c, d categories and errors drops in a similar fashion. To test the strength of these factors, partial correlations were then calculated as follows:

$r_{12.34} = -0.89$ $r_{14.23} = -0.18$

$r_{13.24} = -0.25$ $r_{13.4} = -0.44$

These correlations show that the total number of messages is strongly correlated with errors if the effect of the number of a, b, c, d and e, f, g, h messages are held constant, but that if the effect of the total number of messages is eliminated, there is very little correlation between errors and the a, b, c, d or e, f, g, h messages. This confirmed our belief that this content analysis was not adapted to showing the real differences between groups, since the best correlation is obtained from a count of all messages sent, regardless of categories. An analysis with categories designed in an effort to pick out the use of redundancy as a method would probably show far more significant results, but this has not been done. Meanwhile, the lack of significance in the categories used and the correlation of total messages with errors tends to give additional support to the use of redundancy as a mechanism. Further clarification of the message content problem must await a new content analysis or new experimental data.

7.2. Choice of Marbles

For Experiment 6, the color of marble dropped as the answer was tabulated for each man at each trial. (Unfortunately, this information was not available for Experiment 3.) An examination of these tables produced evidence for a phenomenon which is interesting although not well understood. The six marbles used in the last 15 trials can be divided into two groups: the cool colors (blue, green, aqua) and the warm colors (white, pink, amber). If the number of times a marble in each of these classes is dropped in error is counted for the last 15 trials and totaled for all groups run on a given network as a function of the correct answer, we arrive at the results given in Table V.5.

Table V.5

Warm-Cool Color Choices When Choice was in Error

		Correct Answer					
		Warm	Cool	Warm	Cool	Warm	Cool
Color Chosen	Warm	42	25	30	40	12	22
	Cool	69	39	73	17	65	39
		P		SF		C'	

Table V.5 seems to show puzzling behavior on the part of all three networks; a high proportion of wrong answers were not even in the right class. One would expect the green marble to be confused with other greens and blues, but hardly with browns and whites. In part, this effect may be explained by separating out the effects of guessing as well as possible. If we split these results into two tabulations, one in which we count errors in which no subject in that group had the correct answer for that trial and another in which we count errors for which at least one member of the group dropped the correct marble, we might expect the pure guesses to fall in the first category and the results of ambiguous descriptions to fall in the second. These results are tabulated as before in Tables V.6 and V.7.

Table V.6

Warm-Cool Color Choices for Trials with All Answers Wrong

Correct Answer

		Warm	Cool
Color Chosen	Warm	10	23
	Cool	65	16

P

		Warm	Cool
Color Chosen	Warm	18	40
	Cool	72	5

SF

		Warm	Cool
Color Chosen	Warm	2	5
	Cool	46	10

C'

Table V.7

Warm-Cool Color Choices for Trials with at Least One Right Answer

Correct Answer

		Warm	Cool
Color Chosen	Warm	32	4
	Cool	4	23

P

		Warm	Cool
Color Chosen	Warm	12	0
	Cool	1	12

SF

		Warm	Cool
Color Chosen	Warm	10	17
	Cool	19	29

C'

Here, what is happening seems somewhat clearer. Table V.6, trials with all errors, shows a random distribution of classes with a heavy weighting on the cool class. These are about the results one would expect if the subjects, when they guessed, guessed at random from all six marbles with about twice the probability of picking one from the cool class as from the warm. This behavior is quite consistent from one network to the next and is perfectly acceptable as a phenomenon of human behavior.

When we come to Table V.7 (the class choices of errors when at least one member of the group dropped the right marble) some puzzling effects show up. The entries on the main diagonal tend to be larger in most cases, but there are differences between networks. These differences show up if we examine these tables for contingency and calculate the probability of the observed entries resulting from a random sampling with the same marginals. Here we get:

For the Pinwheel	$p \leq 10^{-3}$
For the Star with Feedback	$p \leq 10^{-6}$
For the Circle	$p \approx 0.84$

Thus we can see that in the star and the pinwheel when one subject in the group has the correct answer and drops the correct marble, the remaining members of the group either drop the right marble or with high probability pick another marble in the same class as the correct one. This is roughly the behavior we would expect; each marble is confused almost entirely with those marbles most similar to it.

In the circle, however, the value of p shows that this could well be a random result, i.e. when one subject drops the correct marble, the other subjects, if they do not drop the correct marble, pick their choices from all the other marbles with equal probability. There seems to be no tendency for marbles to be confused only with other marbles most similar to it. This behavior on the part of the circles shows a decided difference from that of the star and pinwheel, and at present we have no really adequate explanation for this difference. Two factors may be considered, however, which probably play a part in this effect. In the star, information may pass from the center node, which almost invariably arrives at the answer first, by traversing only one link, i.e. after being subjected to only one coding-decoding operation. In the pinwheel, information may pass from any node to any other node after traversing no more than two links, but in the circle, which is organized like a chain in most cases, information often passes along two links on its way to the central point, and over another two links when the answer is transmitted. This is in contrast to the pinwheel, which is almost never organized about a central point. We have listed in Table V.8 the average number of coding-decoding operations undergone by each piece of information from the start of the experiment to its final arrival at its destination as part of a transmitted answer. Table V.8 assumes that the circle is organized like a chain and that the pinwheel is not organized about a central point.

Table V.8

Network	Average Number of Coding Operations
Star	1.8
Circle	2.8
Pinwheel	2.0

We see that on this basis we would expect the distortion to increase and hence the guesses to become wilder as we change from the star to the pinwheel to the circle. This bears out the results mentioned above. Unfortunately we have no good experimental evidence bearing on this point, and so we may only offer this explanation as a hypothesis to explain an otherwise surpising difference between networks. Future experiments of this sort may well shed light on these differences.

7.3. Subject Variations

The influence of variations in the subjects on this experiment has been mentioned and it is now evident in what way these variations are important. Differences in color vision are perhaps the most obvious cause of variations in performance. As we mentioned, the subjects were all screened for color-blindness, but an examination of the data indicates that decided differences in color vision seemed to be present. However, we can only assume that these differences are randomly distributed in our subject populations, since we have no evidence that they are correlated with other factors.

A more important difference is evident if the mechanism for reducing errors is considered, i.e. semantic redundancy. Since this mechanism relies on alternate descriptions and synonyms, the ability of any subject to use it depends in part on imagination, but to a much larger extent on vocabulary. Since it is known that most of the common ways of measuring intelligence, such as those used by the armed forces and for various educational purposes, are strongly dependent on vocabulary, it is evident that subjects drawn from populations with different distributions of I.Q. will differ in their performance in this experiment. This is true in this case to a greater extent than in the experiments previously described. In comparing Experiments 3 and 6, for instance, the differences in the population from which the subjects were drawn must be kept in mind. It seems justifiable to consider that the volunteer M.I.T. undergraduates used in Experiment 3 would have a larger vocabulary, on the average, than the enlisted Army personnel used in Experiment 6. This is borne out by the differences in results of the two circles, C and C'. The circles run in Experiment 6 showed a poorer performance than those from Experiment 3, particularly with regard to the values of R. These differences must be kept in mind in designing future experiments and particularly in attempting to apply such experimental findings to real situations. The difference arising from differences in vocabulary will generally hold true in any situation in which semantic or coding noise is present.

8. Summary

Although the data presented in this chapter were not as complete nor as good as one might hope, the analysis made above has shown several points. It is clear that experimental and theoretical results from the noise-free case will often be invalid in predicting relative behavior in the presence of noise. In this chapter several of the effects of noise have been pointed out, as well as some of the factors which influence the behavior of a network in the presence of semantic noise. Some of the results observed, which are rather unexpected on the basis of noise-free considerations, will serve as an example of the pitfalls of applying such results to task-oriented groups in the presence of noise.

It has been shown that semantic or coding noise may be treated in a manner analogous to that for channel noise and that the errors in the experiments described here are caused by the presence of such coding noise; indeed, the error frequency may be predicted with fair accuracy on the basis of straightforward and somewhat naive assumptions, using measured values of this noise. The use of redundant coding in the semantic sense has been shown to be the mechanism for the reduction of errors due to coding noise, and parallels have been drawn with the well-known case of channel noise.

Finally, certain differences between groups using different communication networks have been shown, and the factors which seem to be important in accounting for these differences have been discussed. The need for further experiments to widen our knowledge of these factors is evident, and the direction of such experiments is plain. Several requirements which such future experiments should meet have also been established. Any future experiments of this sort should make it possible to directly measure the source and receiver entropies; they should be designed to test further the hypotheses suggested to account for the differences between networks, and if possible to develop new and more exact explanations of these differences. Some correlation of the experimental results with the vocabulary or I.Q. of the subjects would also be desirable, since at this point we have no knowledge of how sensitive performance on this type of experiment is to such variations.

CHAPTER VI – ATTITUDES AND LEADERSHIP

1. Introduction

In this chapter we shall examine data of a class which has been of primary interest in most of the published work on group studies. These data concern the motivational and attitudinal aspects of group participation, leadership, group cohesiveness, and other like factors. It should be clear at this stage that our primary interest has been in the objective data (time, errors, and decisions) on group performance, in the objective data on individuals in the group situation, and in the relationship between these two sets of data. For the present, these interests and the resulting restrictions on the experimental groups have caused us to relegate data on attitudinal factors to a secondary role. This procedure does not follow from a belief that these factors are of secondary importance but rather from a belief that the problem can be attacked most effectively in this order.

Historically, the present direction of research is a development of earlier work which had a more conventional emphasis. In Leavitt's paper (34) the primary emphasis is on questionnaire results, and there is very little concern with detailed communicative behavior within the groups. Increased emphasis on the process of communication has led to a sequence of imposed restrictions on the group processes, which have tended to lessen the distinctness of questionnaire results. Furthermore, we cannot claim by an appeal to face validity that our results generalize; our group situations have deliberately been made highly artificial, and our data on attitudes may, in part, reflect this character of the situation.

We do not suggest that it is possible to attain a complete understanding of the dynamics of group processes without simultaneously attaining a correspondingly complete understanding of the dynamics of the behavioral processes of the persons forming the group. Rather, we suppose that by a series of successive approximations, each specifying the sources of individual variability to a greater degree, we may go from a simple but not very generally applicable model to one of greater complexity but of more general applicability. This supposition will cause us in the future to put more and more emphasis on subject attitudes and motivation.

In this chapter we shall present some preliminary data on attitudinal factors. These data are of some interest in themselves but are of questionable generality. Aside from the intrinsic interest of these data, they are important for any future program designed to obtain more precise attitudinal results, for they set up a few signposts in a complex territory.

2. Attitudes and Problem Behavior

Presumably, subject attitude in any of our experiments is a function of trials. We could have broken the experimental session a few times (more often would have been impossible because of time limitations) to administer a questionnaire. Since, in all likelihood, such interruptions would have had an influence on behavior during subsequent trials, and since the ongoing behavior was the main subject of interest, it was decided to use the questionnaire only after the entire experimental run. Therefore, the attitudes tested are those which the subjects held at the end of the experiment or those which they remember having held sometime during the experiment. An example of the latter kind is the question, "How did you like your job?" with a check-list to be filled out for various trial blocks. The results on this type of question support the supposition that subject attitude is not constant. In Appendix 1 reproductions of the questionnaire forms used in the several experiments will be found.

Additional attitudinal data, uncontrolled and unsolicited, exists in the message content for all experiments except Experiment 4. When the message content is unrestricted, it is possible for the subjects to express their feelings toward the experiment, the experimenter, the other subjects, and any of the conditions entailed by the experiment, and the subjects frequently availed themselves of this opportunity. Content analyses of the messages for Experiments 1 and 3 have been done but will not be presented, since they are of little interest in the present context. The results for Experiment 1 can be found in Leavitt (34) and the results for Experiment 3 are essentially similar.

In Experiment 4, where the message content was restricted to input information, there were, by and large, no attitudinal expressions in the messages sent. We did, however, point out in section III.7 at least one possible exception. Recall that in the chain the frequency of sending to the center node from the middle node, when no new information was to be conveyed, was much higher than should be expected, if we assumed that the messages were entirely informational in nature. It seemed plausible to interpret the intent of such messages as that of informing the center node that the middle node needs more information. Receipt of these messages should indicate to the center man that he is in the key position. Since this occurred in the chain groups and no evidence of the same phenomena has been found in the other groups, the chain data are not strictly comparable to the remainder of the data from Experiment 4. For the earlier analyses this discrepancy was not important, but it will be shown to have an effect on the results in this chapter.

The above example clearly demonstrates the influence of problem behavior on attitudes. Equally well, we expect attitudes to have an effect on problem behavior, but we have no simple example from our experiments to demonstrate this. The questionnaire data are only terminal results of the attitudes at particular times during the experiment, and granted that local attitudes will affect local decisions and action times, it does not follow that the terminal attitudes will be well correlated with averaged decisions or action times. This effect is, nevertheless, one of the important interactions of the group process. We believe the solution will have to assume the form of a subtle feedback analysis of this interaction. Though our present data will not support such an analysis, it is our intention to attempt to collect data in the future in a form suitable for such an analysis. The experimental difficulties are formidable.

3. Nodal Properties

Bavelas (78) proposed a model for group structure in which he defined index numbers intended to characterize networks having symmetric links, and these were applied by Leavitt (34) to the data of Experiment 1. We shall describe these index numbers and examine their utility as applied to the data of Experiments 1 and 4.

We may partially describe the relation of a node to the network in which it lies by measuring the (least) distances between pairs of nodes in link units and finding the sum of such distances to each other node from a given node j. We call this index "sum-of-distances" from j. When the sum-of-distances from j is summed over all j in the network and divided by the sum-of-distances from a particular node i, the number obtained is called the "centrality index" for that node. Thus, if d_{ij} is the least distance between nodes i and j, the centrality index for node i is defined by

$$C_i = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}}{\sum_{j=1}^{n} d_{ij}}.$$

The centrality index has the disadvantage of being measured on an arbitrary scale which, among other things, increases with n. Of interest is the divergence of a particular node i from some network norm, i.e. an indication of the variability of node properties which is only minorly dependent on n. Of the many possible measures, Leavitt defined

$$P_i = C_{max} - C_i$$

which he called the "relative peripherality index." The relative peripherality indexes for the various positions in the networks studied in Experiments 1 and 4 are given in Fig. VI.1.

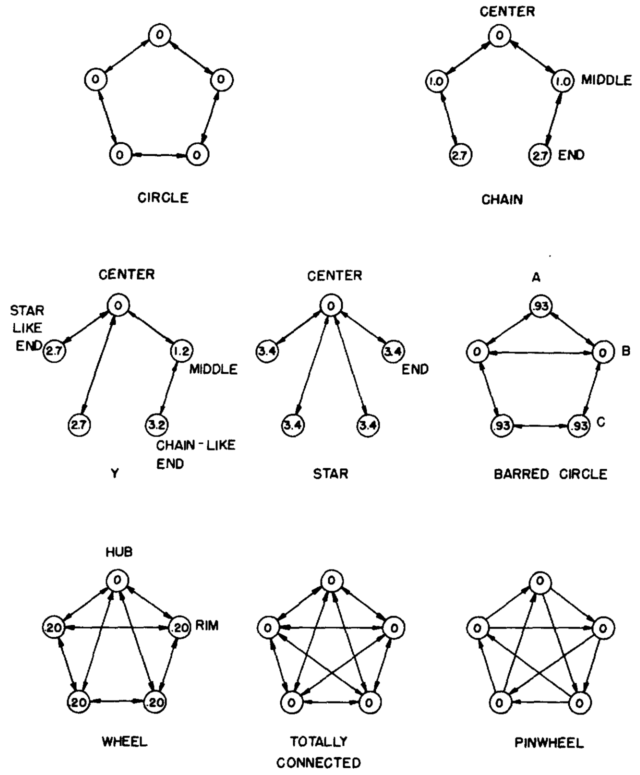


Fig. VI.1

Relative peripherality indexes.

Each of the networks in Fig. VI.1, except pinwheel, has symmetric links, so the indexes are well defined. In addition, for those networks having the sort of symmetry exhibited by pinwheel, but without symmetric links, these indexes are uniquely defined. Such is not the case for the network alpha or for any other network in which the sum-of-distances over incoming links to a node is not the same as the sum-of-distances over outgoing links from that same node. If the word "from" in the verbal definition of sum-of-distances given by Bavelas and used by Leavitt is taken literally, it means distances over outgoing links which should be taken. For our purposes, it will be more appropriate to use incoming links. Figure VI.2 gives the relative peripherality indexes for alpha for distances defined over both incoming and outgoing links.

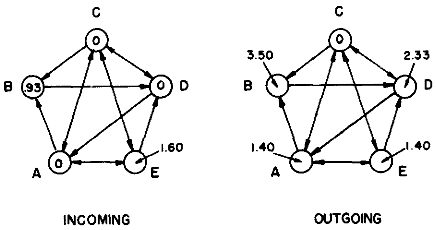


Fig. VI.2
Relative peripherality indexes for alpha.

The relative peripherality index for nodes was devised in order to characterize the sort of job a man would have when occupying a node. The hope was that it would sufficiently well describe the restrictions and demands placed on nodal behavior, and that the attitudes of the subjects would be correlated with the relative peripherality of the nodes the subjects occupied. This hope was realized for the networks studied in Experiment 1. Although the establishment of such correlations is an excellent first step, there remains a problem in explaining the correlations. Relative peripherality is not simply related to a description of the stimulus situations and response possibilities which arise for various nodes. Even more serious, as we have seen in the case of alpha, if nonsymmetric links are allowed, the index is not unambiguously defined.

Let us examine, in some detail, the case of the questionnaire item, "How did you like your job?" For Experiment 1 this item has a high negative correlation with the relative peripherality index, viz. r = - 0.924. If we consider the correlation of peripherality with job enjoyment for the various nodal types in the networks run in Experiment 4, there are two possible results depending on whether we consider incoming or outgoing distances for alpha. Using outgoing links, we find r = -0.496 and using incoming links, we find r = -0.616. This difference indicates that job enjoyment is much more closely related to the input situation at a node than to the output situation. This result is not surprising when we consider that the nodal output is one message at a time, whereas the number of input messages may vary from zero to the number of incoming links.

It will be recalled that it was shown in Chapter III that the near neighborhood of a node had a predominant influence on behavior. This suggests that the simplest structural characteristic on the input side of a node which might correlate with job enjoyment is the number of incoming links. However, it was well demonstrated that behavior was a function of what came over those links, and it is intuitively clear that an unused incoming link might just as well not be there for all it contributes to job enjoyment. Rather than a structural parameter we need a functional one. Let us therefore consider input message density at a node as a candidate for the variable which determines job enjoyment. Our goal will be to show that for all practical purposes this functional property can be reduced to a structural one, admittedly more complex than simply the number of input links.

It would seem likely that there should be some optimal value of input density for which job enjoyment would be highest. For input densities greater than this value, the node would be overloaded much of the time, and for densities less than this value, the node would be idle much of the time. In Experiments 4 and 5, however, the pacing of message flow is determined by the slowest node in the group, and the results of Chapter IV indicate that the node with the largest number of inputs is the slowest node. Thus, no overload situation can arise. The node with the greatest input density pretty much paces the group to yield a satisfactory working rate for himself. Therefore, we should expect job enjoyment to be monotonically related to input density.

Does the structure of the network determine input density? If it does, we can reduce input density to a structural property of the network, thereby accomplishing the purpose for which the peripherality index was invented, with a better understanding of why the relation holds. We have pointed out in Chapter III that there is a strong alternation tendency on the part of our subjects and also a tendency toward locally rational behavior. Both of these tendencies and any random element in the determination of sending probabilities result in nearly equal usage of all outgoing links. The sequence of usage is often not random, but the frequencies are near their equiprobable values. In this situation the average input density at a node or mean number of inputs per unit time can be simply related to the number of output links available to each of the nodes which can send to the given node. For Experiment 4 these values can be calculated readily as "mean inputs per act," and for Experiment 1 the same figures will approximately represent mean inputs per unit time for an appropriate time interval. Table VI.1 gives the inputs-per-act distributions and means on an equiprobable sending basis for each type of node studied.

Leavitt's results on the job enjoyment question were presented as percentages of a unit interval, and the corresponding results for Experiment 4 are also easily expressed as distances within a unit interval. It is therefore appropriate to reduce the figures for mean input, which have a possible range from zero to infinity (for a net with an infinite number of nodes), to the unit interval. Using the transformation,

$$f(I) = \frac{e^{2I} - 1}{e^{2I} + 1}$$

where I = mean input, a preliminary plot of the data of Experiment 1 was made. It was found that the relationship was strong, but that it was curvilinear to such a degree that it could not be adequately represented by a linear correlation. Therefore an arc sine transformation was also employed. The final formula for the scale change in I is

$$g(I) = \frac{1}{\pi} \sin^{-1} \left\{ 2 \frac{e^{2I} - 1}{e^{2I} + 1} - 1 \right\} + \frac{1}{2}$$

Let us call g(I) the "input potential." The correlation between the input potential and job enjoyment for Experiment 1 is r = 0.948, an improvement over the value previously found by using relative peripherality. Thus 90 percent of the variance in the job enjoyment is accounted for by its relation to input potential. A scatter diagram for this correlation is given in Fig. VI.3.

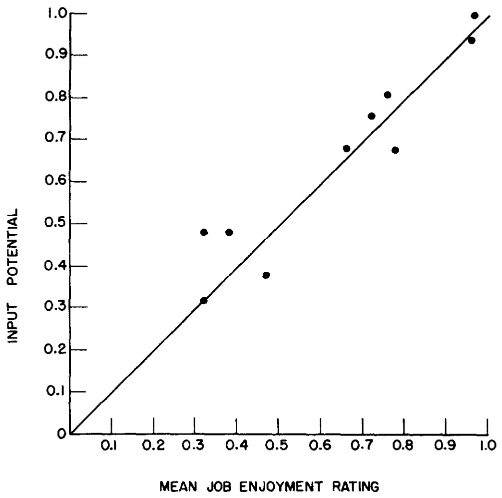


Fig. VI.3

Input potential vs job enjoyment, Experiment 1.

No theoretical significance is attached to the analytic form of the equation for g(I). It is merely a transformation of scale on a stimulus attribute to render its relation to the arbitrary scale of a response attribute linear for the purpose of measuring the strength of the relation by a correlation coefficient. Note, however, that g(I) is not a curve with fitted constants. On the other hand, the skew signoid form of the relation between job satisfaction and input potential (or input density) is a feature of the data which probably has psychological significance.

When the transformation g(I) is applied to the input potentials of the node types in Experiment 4 and the results are correlated with the job enjoyment scores, we find r = 0.683. This is again an improvement over the correlation found between relative peripherality and job enjoyment. A scatter diagram is presented in Fig. VI.4. The number of individual responses averaged to give the job enjoyment scores for various node types is larger in Experiment 4 than in Experiment 1, and for this reason, the Experiment 4 scores would be expected to be more stable. On the other hand, the subjects used in

Table VI.1

Input Distributions and Means by Node Types for Equiprobable Sending

Node	Inputs					
	0	1	2	3	4	Mean
Circle	0.250	0.500	0.250	----	----	1.00
Pinwheel	0.250	0.500	0.250	----	----	1.00
Totally Connected	0.316	0.422	0.211	0.047	0.004	1.00
Chain End	0.500	0.500	----	----	----	0.50
Chain Middle	----	0.500	0.500	----	----	1.50
Chain Center	0.250	0.500	0.250	----	----	1.00
Chain Star End	0.750	0.250	----	----	----	0.25
Chain Center	----	----	----	----	1.000	4.00
Wheel Rim	0.333	0.444	0.194	0.028	----	0.92
Wheel Hub	0.198	0.395	0.296	0.099	0.012	1.33
Barred Circle A	0.444	0.444	0.111	----	----	0.67
Barred Circle B	0.167	0.417	0.333	0.083	----	1.33
Barred Circle C	0.333	0.500	0.167	----	----	0.83
Alpha A	0.250	0.458	0.250	0.042	----	1.08
Alpha B	0.500	0.417	0.083	----	----	0.58
Alpha C	0.222	0.444	0.278	0.056	----	1.17
Alpha D	----	0.500	0.500	----	----	1.58
Alpha E	0.500	0.417	0.083	----	----	0.58
Y-Chain-like End	0.500	0.500	----	----	----	0.50
Y Middle	----	0.667	0.333	----	----	1.67
Y-Center	----	----	0.500	0.500	----	2.50
Y-Star-like End	0.667	0.333	----	----	----	0.33

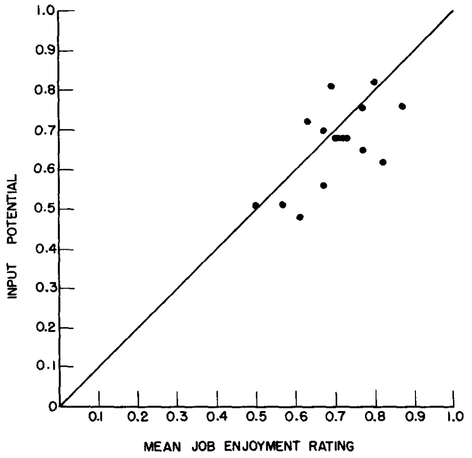


Fig. VI.4

Input potential vs job enjoyment, Experiment 4.

Experiment 1 were a much more homogeneous group than those used in Experiment 4, which should mean greater stability of the scores for Experiment 1. These two factors, acting in opposition, seem to result in scores of about the same stability in the two cases. The explanation of the difference in the correlations found must, therefore, be sought elsewhere.

The sample of node types over which the correlations were taken were not the same in the two experiments. In Experiment 1 both star and Y contain nodes with more extreme input potentials than do any of the networks run in Experiment 4; in fact, the star center node has the highest and the star end node has the lowest input potential possible in any network on five nodes. Thus, the range of node types in Experiment 4 was less than in Experiment 1; so, for the same column variance, a lower correlation is to be expected in 4 than in 1. The fact that the correlation in Experiment 4 accounts for only 47 percent of the variance of the job enjoyment scores indicates that although input potential may explain a considerable part of job enjoyment, there are probably other factors involved. The difference between input potential and input density is not the source of the imperfect prediction, since the correlation between job satisfaction and input density for Experiment 4 is r = 0.641. An examination of the points which deviate most from a perfect correlation suggests that those inputs which contain only redundant information, and which, of course, contribute to input potential as we have defined it, contribute little or nothing to job enjoyment. As an example, consider the inputs to the middle man in the chain. Every message he receives from the end man except the first is redundant. If these messages were not counted as part of his input potential and a new measure, nonredundant input potential, were defined, the relation to job enjoyment would be strengthened. It would be possible to calculate nonredundant input potentials for each type of node we have considered and thus test this hypothesis. We have not done so because the calculation, although straightforward, is excessively tedious.

In order to round out the present argument, it is necessary to show that the structural property of the network, input potential, is closely related to the behavioral feature of the stimulus situation, input density. Data are available to examine this question only in the case of Experiment 4. The tabulations of experimentally obtained input densities are compared with input potential in Table VI.2. The correlation between the two is r = 0.993.

Table VI.2

Comparison of Inputs Observed and Inputs Expected with Equiprobable Sending

Node Type	Mean Equiprobable Input	Mean Input Density
Circle	1.00	1.00
Pinwheel	1.00	1.00
Totally Connected	1.00	1.00
Chain End	0.50	0.53
Chain Middle	1.50	1.50
Chain Center	1.00	0.95
Wheel Rim	0.92	0.94
Wheel Hub	1.33	1.24
Barred Circle A	0.67	0.63
Barred Circle B	1.33	1.27
Barred Circle C	0.83	0.89
Alpha A	1.08	1.09
Alpha B	0.58	0.65
Alpha C	1.17	1.13
Alpha D	1.58	1.54
Alpha E	0.58	0.58

By definition, the input potentials averaged over the network must come to unity in the case of every network in which the topological relation of each node to the rest of the network is the same. It is also true, since the instructions to the subjects in Experiment 4 required them to send a message on every act, that the network average of input density for the same class of networks is 1; although group interaction may not be such that unity is the appropriate value for each node. If the group adopts some distinctive manner of performing its task and adheres to it for many trials, we may expect deviations. We know that such consistent modes of performance did develop in the case of circle (see Chap. III), and we know in detail the sending sequences which yield minimum solutions in the circle. If we consider a set of groups in which each adopts one of these patterns, we can calculate the resulting input potentials for each node. If we assume that each possible pattern is as likely as any other, an unweighted average will give us the mean value of the input potential. However, we know that those patterns which consist of alternations on the part of each node are strongly favored by the subjects. Averaging over these patterns will give a third possible mean value for input potential. The values are: equiprobable sending, 0.677; average of all minimum patterns, 0.667; average of all alternation minimum patterns, 0.670. Thus, the value we have used in our correlations, 0.68, is a close approximation in any case.

Relative peripherality was found by Leavitt to correlate with several attitudinal items other than job enjoyment and with certain objective features of group performance. We have examined in detail only the case of the highest of these correlations and have found a more clear-cut explanation of the relationship in terms of the stimulus situations which arise at the various node types. It may be possible to find a formal relation between relative peripherality and input potential, but such a formal relation would be extremely unwieldy because relative peripherality involves the entire network, while input potential depends only on immediate neighbors.

In addition, in networks of large n, it is clear that the relative peripherality index can depend upon the centrality of a node very far removed from the node in question. This contradicts the very reasonable finding that the near neighborhood of a node is dominant in influencing its behavior, and consequently, relative peripherality is an inherently unsuitable choice as a determiner of this behavior.

The sort of analysis we have presented above for job enjoyment exemplifies what we propose to attempt in the way of explanation of the attitudes displayed in task-oriented groups. We do not feel that the questionnaires we have used so far are sufficiently reliable instruments with which to make more than a bare beginning on this program. Therefore, an improvement in our techniques for the assessment of attitudes is felt to be necessary. A first step in this direction will be described in section 5 of this chapter.

4. Leadership

In the present state of knowledge we cannot expect to give a precise definition of leadership except in a way which is so arbitrary as to be of doubtful general utility. The best we can do is to give a specification of the term, leadership, which seems appropriate for our present purposes. This section deals with our subjects' views on leadership and does not purport to be an analysis as to who was, in fact, the leader. We asked the subjects, "Was there a leader in your group? If so, who?" We shall attempt to account for their answers in terms of the information a node receives about the amount of influence each node has on the group performance. However, it will be impossible to analyze the choices of a leader by our subjects in a manner similar to the way in which job satisfaction was analyzed because the data are too meager. To apply a correlational treatment, as is well known, one must have a relatively large number of pairs of observations to insure stability. Our data on leadership do not permit this, so we must work with empirical choice frequencies, testing the fit of any theory by χ^2.

In Experiment 4, the only basis a subject had for identifying the leader of the group was the source and content of his incoming messages. Also, from the content of a message, a subject could sometimes deduce which nodes could communicate to the sender of the message. It seems reasonable to assume that the amount of new problem information contained in messages received from a node would be used as an indication of that node's importance to the group. We know that under the conditions of Experiment 4, the nodes closest to a given node exert the greatest influence on that node. Therefore, we should expect a node to select a leader among the nodes having input links to him. Let us call nodes so related to a given node "input-adjacent nodes," and all other nodes "input-distant" nodes. With this definition our expectation is that among all leader choices made, other than choices of self, input-adjacent nodes would be favored over input-distant nodes. We can test this hypothesis against a random nonself choice expectation for each network run, except totally connected, for in that network each node is input-adjacent to every other node. The χ^2 tests are given in Table VI.3. In the case of each network, input-adjacent nodes are chosen significantly more frequently than they would have been had the choice been random.

The foregoing discussion takes no account of choices of self as leader. It is possible to examine such choices for any divergence from random expectation. Table VI.4 gives the results of this procedure. The expected frequencies in the self-choice category are too low, in most cases, to apply a χ^2 test, but the consistency of the trend bears out the notion that the subjects tend to avoid self choices. The avoidance is most prominent in the case of the chain middle nodes, and it is reversed in the cases of the chain center node and barred circle. In the latter case, it is the B positions (see Fig. VI.1) which produce the reversal since they chose self over nonself in the ratio of four to three. The divergence of the chain from the common trend is probably related to the previously mentioned instrumental use of messages by the middle node.

The tendency to choose self on the part of chain center and barred circle B can be explained readily. Let us first consider the chain since it is the more clear-cut case. The way in which message sending choices are made (see Chap. III) ensures that the center man is far more likely to get the problem solution first than any other man in the chain network. When this happens, his subsequent messages contain all the problem information; in other words, they constitute "answer" messages. The center man must be aware of this fact, since one of the requirements imposed on his actions in the experiment is that he press the answer button as soon as he has all the problem information in his possession. Therefore, he is very likely to realize his important central position and thus nominate himself as leader. This description can easily be seen to apply, although not so strongly, to the B-position men in barred circle, and seen not to apply to any other position in these two networks or in any other network among those run in Experiment 4.

Table VI.3

Leadership Choice and Input Adjacency
Experiment 4

(ν=1 in all cases)

Circle

	A	D	
f_o	21	2	$\chi^2=15.896$
f_e	11.5	11.5	$p<<0.001$

Chain

	A	D	
f_o	34	11	$\chi^2=23.016$
f_e	18.25	26.75	$p<<0.001$

Pinwheel

	A	D	
f_o	14	5	$\chi^2=4.263$
f_e	9.5	9.5	$0.05>p>0.02$

Wheel

	A	D	
f_o	19	0	$\chi^2=6.333$
f_e	14.25	4.75	$0.02>p>0.01$

Barred Circle

	A	D	
f_o	8	1	$\chi^2=5.639$
f_e	5.25	4.75	$0.02>p>0.01$

Alpha

	A	D	
f_o	15	3	$\chi^2=3.740$
f_e	11	7	$p\cong 0.05$

Table VI.4

Self vs Nonself Choices of Leader
Experiment 4

(ν=1 in all cases)

Circle

	S	$\bar{S}$	
f_o	3	23	$\chi^2=3.257$
f_e	5.2	20.8	$0.5<p<0.10$

Chain End

	S	$\bar{S}$	
f_o	3	17	$\chi^2=0.313$
f_e	4	16	$0.50<p<0.70$

Chain Middle

	S	$\bar{S}$	
f_o	1	23	$\chi^2=3.760$
f_e	4.8	19.2	$0.50<p<0.10$

Chain Center

	S	$\bar{S}$	
f_o	6	5	$\chi^2=8.205$
f_e	2.2	8.8	$0.001<p<0.01$

Pinwheel

	S	$\bar{S}$	
f_o	1	18	$\chi^2=2.579$
f_e	3.8	15.2	$0.10<p<0.20$

Wheel

	S	$\bar{S}$	
f_o	2	21	$\chi^2=1.836$
f_e	4.6	18.4	$0.10<p<0.20$

Barred Circle

	S	$\bar{S}$	
f_o	6	9	$\chi^2=3.75$
f_e	3	12	$p\approx0.05$

Alpha

	S	$\bar{S}$	
f_o	4	18	$\chi^2=0.045$
f_e	4.4	17.6	$0.98<p<0.99$

Totally Connected

	S	$\bar{S}$	
f_o	4	13	$\chi^2=0.132$
f_e	3.4	13.6	$0.70<p<0.80$

From Experiment 1, the self choices of star center are 100 percent and of Y center, 80 percent; these high values would be expected by the same argument.

So far, we have shown that when a choice of leader is made, the choice will be self only under special (and explicable) circumstances, and if it is not self, it will very probably be a choice of an input-adjacent node. If there is more than one input-adjacent node, but each is related to the network in the same way as the others, we have no basis for expecting one to be selected as leader more often than any other. This is the case for most nodes in the networks we have studied. There are, however, a few cases where there are two or more dissimilar nodes input-adjacent to a given node. This situation obtains for wheel rim nodes, barred circle nodes B and C, and for all nodes of alpha. Nothing can be done to analyze barred circle B and C because the choice frequencies are very low. The remaining cases provide data for a further test of our hypothesis that the value of a node as a source of information determines the frequency with which it is chosen as leader.

The information that one node gains from another depends upon the frequency of messages over that channel and upon the richness of content of those messages. We may, as we saw in the preceding section, closely estimate the frequency of use of a channel by assuming equiprobable sending. The richness of content of messages from a given node is very closely estimated by the input potential of that node. Without an exhaustive combinatorial analysis, it is impossible to know how to combine these factors or how to allow for redundancy. Therefore, let us simply take the product of the relative frequency of use of a channel (on an equiprobable sending basis) and the mean input of the sending end of the channel as an estimate of the sending end's value to the receiving end. When the resulting index is normalized over all nodes input-adjacent to a given node, we have an estimate of the relative frequency of leadership choice for the various input-adjacent nodes. In the case of wheel rim, these estimates can be applied directly to the choice frequencies since the total number of choices by wheel rim nodes is large enough to yield expected frequencies which can be tested by χ^2. For alpha, in order to get large enough expected frequencies, it is necessary to lump together the two nodes, A and C, which have the largest expected frequencies, and to lump together the three nodes, B, D, and E, which have the smallest expected frequencies. Table VI.5 shows the results. The χ^2 test of goodness-of-fit is excellent in the first case and fairly good in the second. An incomplete, but more detailed analysis of alpha indicates that the fit can be improved by refining the estimates of choice frequency according to the discussion; however, the paucity of data does not warrant further effort along this line.

This analysis is heavily dependent on the severely restricted nature of our experimental situation. In the cases of the circle and chain networks, it is possible to make a comparison of the Experiment 4 data with the data from the much less restricted situation of Experiment 1. The first difference noticeable is that refusal to make a choice is much more common in Experiment 4 than in Experiment 1. The percentages of subjects making leadership choices is shown in Table VI.6. The frequency of choice seems to lie around 30 percent when there is little basis for a choice. For example, in the case of totally connected, for which there is no network differentiation to determine leadership choice and which shows no divergence from random assignment of the choices, 34 percent of the subjects make a choice. With increasing differentiation, the percentage of choices made increases. For circle and chain the difference between the percentage

Table VI.5

Leadership Choice Among Adjacent Nodes
Experiment 4

Wheel Rim

	Node Chosen		
	Adjacent Rim	Hub	$\chi^2=0.214$
f_o	12	7	$\nu=1$
f_e	11.3	6.7	$0.50<p<0.70$

Alpha

	Node Chosen		
	A or C	B, D or C	$\chi^2=1.191$
f_o	10	5	$\nu=1$
f_e	7.88	7.10	$0.20<p<0.30$

Table VI.6

Leader Choice vs No Choice

	Experiment 4	Experiment 1
Network	Percent Choice	Percent Choice
Circle	35	48
Chain	55	72
Pinwheel	40	
Totally Connected	34	
Wheel	46	
Barred Circle	30	
Alpha	44	
Star		92
Y		80

choice in Experiment 1 and in Experiment 4 cannot depend on network parameters but must be explained in terms of the different conditions under which the two were run. Experiment 1 allowed for interaction much more nearly like that to be found in natural groups. The message content was unrestricted, and as a result, the functions of leading and following could be, and were, explicitly expressed in the content of the messages. The selection of the leader is obviously much easier in such a case, provided some node acts as the leader.

The unrestricted message content of Experiment 1 also means that a man can know more about the function of input-distant nodes; hence, we should not expect input-adjacent nodes to dominate the choices in Experiment 1, whereas we found that they did in Experiment 4. Table VI.7 shows a four-fold table for this comparison for the circle network (the only case in which the comparison can be made uniformly over the nodes). The exact probability, p=0.029, of obtaining this table if there were no contingency between adjacency and experiment type shows that our contention is well founded. We may also consider the choices by chain end men of the adjacent middle man vs the (nonadjacent) center man and compare Experiments 1 and 4. Table VI.8 presents this comparison, yielding the exact probability, p=0.021, which shows very nicely that the true group center (leader) will be chosen in the case of unrestricted message content, but that the adjacent man in the direction of that center will be chosen in the other case. This relationship is actually even stronger than the table indicates, because in one chain group in Experiment 1, a middle man was the leader in the sense that all information was funneled to him, and he sent out the answer. In this case (which never occurred in Experiment 4) he was picked as the leader by the two end men.

We have taken the convention that the leader of a group is the member whose part in the group process most influences the behavior of the other members. Insofar as our subjects' responses to the questionnaire were made in terms of this same convention, the problem of predicting their reports of leaders becomes that of assessing the indications of influence in the messages they receive. With the highly restricted situation we have used in Experiment 4, the extent of influence is obviously a matter of the amount of problem information sent. Using this analysis, we have been able to show a good fit to the data.

The theory of leadership given in Appendix 3 is based on the same convention as that used in this chapter. However, it does not take account of how much information one node can have about the importance of each other node, and it covers only situations in which

Table VI.7

Adjacency vs Message Content Restriction in Circle Groups

	Restricted	Not Restricted
Adjacent	21	6
Nonadjacent	2	5
	p=0.029	

Table VI.8

Choices by Chain End Men of Adjacent Middle or Center vs Message Content Restriction

	Restricted	Not Restricted
Middle	11	1
Center	2	4
	p=0.021	

there is only one (repeated) piece of information. Therefore, it could not be applied here. It is a more sophisticated theory and a modification of it which permits taking account of these two factors would be very useful.

5. Factorial Analysis

The questionnaire was intended to yield information about attitudes, leadership, and incidental knowledge of the network structure. The attitudes it was designed to assess were selected on an intuitive basis, and items included were selected on the basis of their face validity as indicators of the intuited attitudes. For example, it was felt that the subjects would develop an attitude toward the excellence of their group's performance; so they were asked, "Do you think your group could do better?" The questionnaire items were not selected with the hope that they would be independent but rather in such a way that more than one attitude would enter into the determination of responses to many of the items. In this situation the description of the subjects given by the raw questionnaire results is not parsimonious since there are more items than there are presumed attitudes. In this section we shall present the results of a factorial analysis of the questionnaire data. This analysis will do two things: (1) It will check whether the data supports the presumed attitudes. (2) It will disclose any additional attitudinal factors which may exist in the data, thereby providing a basis for the construction of more effective questionnaires for future experiments.

A reproduction of the questionnaire form used for Experiment 4 is included in Appendix 1. Table VI.9 gives the questionnaire items and the method of scoring for each item. (The items have been renumbered for the purposes of this section.) Since the frequency split on the dichotomous items was never extreme, it was

Table VI.9

Numbering and Scoring of Questionnaire Items

Item	Question	Scored
1A	Show who can communicate to whom. (A drawing was to be made.)	For direct links*, percentage reported.
1B	Show who can communicate to whom. (A drawing was to be made.)	For first degree indirect links*, percentage reported.
2	Did your group have a leader?	Yes = 1, No = 0, and Blank = 0.
3	Was there anything at any time that kept your group from performing at its best?	Yes = 1, No = 0, and Blank = 1/2.
4	Do you think your group could do better?	Yes = 1, No = 0, and Blank = 1/2.
5	How many more problems do you think it would take before you would get "fed up"? Circle the number. 0 10 20 40 60 80 100 More	As indicated, with "More" = 120 and Blank = 0.
6	Circle the words which describe how well your group did on this experiment. Extremely-Mediocre-Average-Better-Excellent Poor Than Average	From 0 for "Extremely Poor" to 4 for "Excellent," and Blank = 2.
9A	Check in each column how you felt about your job as the experiment proceeded. Trial – 5 10 15 20 25 liked it very much ☐ ☐ ☐ ☐ ☐ liked it ☐ ☐ ☐ ☐ ☐ disliked it ☐ ☐ ☐ ☐ ☐ disliked it very much ☐ ☐ ☐ ☐ ☐	For 5th trial from 0 to 4; missing data supplied by interpolation or extrapolation.
9B	Same as 9A	For 15th trial.
9C	Same as 9A	For 25th trial.

---

*Direct links are those from or to the node in question. First degree indirect links are links over which inputs arrive at those nodes which are input-adjacent to the node in question.

felt reasonable to treat all scores as continuous variables and compute product moment correlation coefficients. Table VI.10 presents the matrix of intercorrelations. Each of these correlations is based upon 450 pairs of observations so that correlations which differ from zero by 0.05 or more are significant at the 5 percent level or better.

Table VI.10

Intercorrelations of Questionnaire Items

	1B	2	3	4	5	6	9A	9B	9C
1A	+0.135	-0.085	+0.199	-0.129	-0.199	-0.195	-0.205	+0.005	-0.078
1B		+0.007	-0.005	+0.031	-0.046	-0.040	+0.196	-0.004	-0.015
2			+0.239	+0.056	+0.127	+0.086	+0.015	+0.023	+0.078
3				+0.372	+0.072	-0.217	+0.049	-0.076	-0.052
4					+0.103	-0.246	+0.051	+0.036	-0.231
5						+0.030	+0.116	+0.388	+0.465
6							+0.025	+0.084	+0.102
9A								+0.530	+0.087
9B									+0.688

The matrix of item correlations was factored by the centroid method, and four factors were found before the residuals became negligible. The factoring was done three times, using as communalities in each case the values found in the preceding analysis in order to improve the stability of the estimates of the communalities. It was considered important to do this since the communalities constitute estimates of the percentage of variance of the test items accounted for by the factors. Table VI.11 contains the final values of factor loadings for the centroid solution and the communalities. It is to be noted that some of the communalities h^2 are very low, particularly items 1B and 2. Item 2 is the one questionnaire item which has to do with leadership, and it is therefore encouraging that it shows this degree of independence. Item 1B is an indicator of incidental learning which should also be expected to have a small communality in this battery since the only other item testing memory of incidental features of the experiment, item 1A, presumably involves little learning. From a partial analysis of the probability of messages containing information from which the existence of indirect links can be deduced, it can

Table VI.11

Centroid Factors

	Factors				
	I	II	III	IV	h^2
1A	-0.381	-0.139	-0.456	0.289	0.46
1B	-0.027	-0.149	-0.194	-0.110	0.07
2	0.268	0.213	0.080	0.213	0.17
3	0.215	0.444	-0.420	0.273	0.49
4	0.185	0.484	-0.350	-0.284	0.47
5	0.570	-0.068	-0.025	0.170	0.36
6	0.091	-0.183	0.439	0.031	0.24
9A	0.353	-0.196	-0.143	-0.471	0.36
9B	0.775	-0.555	-0.307	-0.133	1.00
9C	0.546	-0.477	-0.020	0.350	0.65

be shown that the item 1B reports are highly correlated with frequency of receipt of the information upon which the report must be based.

Aside from incidental learning and leadership, the questionnaire was intended to assess the subjects' liking for their jobs and their attitude toward how well their group had done. It was therefore desired to rotate the factor matrix to show these factors clearly, if that could be done. Accordingly, trial factors were defined as follows: factor I, job enjoyment, the average of items 5 and 9C; factor II, attitude toward quality of performance, the average of items 3 and 4; factor III, a temporarily unidentified factor, item 1A; factor IV, the centroid factor IV. The result of this procedure is an oblique factor matrix. Factor IV is unchanged and orthogonal to the other three factors and the angles θ between pairs of factors I, II, and III are:

$$\theta(I, II) = 95^{o}45'$$

$$\theta(I, III) = 64^{o}12'$$

$$\theta(II, III) = 91^{o}49'$$

When plots of the test vectors against each of the 6 pairs of reference factors were made, it appeared the tests could be accounted for by orthogonal factors as well as, or better than, by oblique factors. Therefore, further rotations were done, one plane at a time, to adjust to orthogonality and to place the orthogonal set of factors in the way which, by eye, gave the best simple fit to the data. The final factor matrix is given as Table VI.12.

Table VI.12

Rotated Orthogonal Factor Matrix

	Factors			
	I	II	III	IV
1A	-0.104	0.115	0.557	0.353
1B	0.108	0.027	0.232	-0.085
2	0.094	0.151	-0.343	0.180
3	0.096	0.631	-0.114	0.274
4	0.031	0.616	-0.103	-0.285
5	0.493	0.106	-0.294	0.141
6	0.025	-0.400	-0.246	-0.003
9A	0.424	0.060	0.042	-0.468
9B	0.987	0.025	0.065	-0.127
9C	0.694	-0.152	0.072	0.357

Clearly, factor I is a job enjoyment factor. It was initially found by taking the average of items 5 and 9C, and item 9B winds up solidly in the middle of the factor. It was item 9B that was used in section 3 of this chapter as the measure of job enjoyment and the present results motivated that choice. It is notable that the three items, 9A, 9B, 9C, which were worded in exactly the same way, except for the trial to which they referred, have intercorrelations well below unity. In particular, the correlation of 9A with 9C is only 0.087. Factor II is clearly an estimate-of-performance factor. It was found using the average of items 3 and 4 and winds up with item 6 loaded more heavily on it than on any other factor. Although items 3 and 4 are loaded most heavily on factor II, each has an appreciable loading on factor IV, but of opposite signs. Factor III is the most difficult to interpret. Its heaviest loading is on item 1A, a simple factual question, and its second heaviest loading (negative) is on item 2, the question as to whether there was a leader, which cannot clearly be answered in the affirmative except on the basis of a broad understanding of what constitutes leadership. Its pattern of loadings suggests, tentatively, a factor of literalness in response to the task and to the questionnaire items. Factor IV is fairly clear and very interesting since it was wholly unexpected. The loadings upon which its interpretation is based are: positive on item 3, negative on item 4, and positive on item 9C, negative on item 9A. It is, like factor II, an estimate-of-performance factor, but, whereas factor II represents an absolute judgment, factor IV represents a relative judgment or level of aspiration which changes as the number of trials done increases. It is a bipolar factor which can be called optimism vs pessimism in level of aspiration. It is indeed interesting that it is orthogonal to factor II.

6. Summary

This final chapter has dealt with one of the most interesting features of group studies: the emotional interaction of the subjects within the group process. As we pointed out, the work in this area is just beginning, and the tools so far available are crude. We have dealt with the subjects' emotional responses as obtained on a questionnaire. The analysis assumed two forms. Some of the questions were shown to correlate highly with some average measures of the communication process (e.g. job satisfaction and input potential). The entire battery of questions for Experiment 4 was subjected to a factor analysis which revealed the existence of four factors to account for the intercorrelations of the responses. This is of only minor interest in and of itself, but it serves as an important tool in the design of improved questionnaires.

Thus, in a crude way, we have dealt with the responses of the subjects to the communication process and, in the factor analysis, with the interrelations of responses. No work has been completed on the difficult problem of the influences of emotional states on the communicative process itself. We can argue that in the experimental situations discussed in this report the effect of emotional factors on the group process was minimal and could, to a first approximation, be neglected. This cannot remain true in future work; however, the methods to cope with the problem are by no means apparent. What is most desirable is a probe or measuring instrument which will ascertain the required features of the emotional state of the subjects but will not seriously influence the communication process. Whether such a technique can be developed, or whether (as seems more probable) it will be necessary to develop simultaneously both the theory of the process and of grosser probes, remains to be seen.

APPENDIX 1 – EXPERIMENT DESCRIPTIONS

The apparatus used in running each of the six experiments to be described included a round table of the type described in section II.3.3. Two versions of this table were used. Style A had slots through the partitions as well as through the center post (see Fig. A1.1), and these slots were sufficiently large so that some visual communication may have been possible. The partitions and work spaces of table A were each painted a different color (red, white, blue, yellow, and brown), and the subject seated at the table was referred to by the color of his compartment. Style B had all the slots through the center post (see Fig. A1.2), and these slots were only large enough to permit a message card to pass conveniently. All the partitions and work spaces of this table were uniform in color.

Each experiment was designed for five subjects.

The communication between group members was, in all cases, by written messages.

1. Experiments 1 and 2: Common-Symbol

Style A table.

Message cards: 8 inches by 1.5 inches, similar in color to the compartment in which they were placed.

Communication within the group: free of constraints other than that the messages be written.

Input information: at each trial each subject was given a card bearing five of the six symbols shown in Fig. A1.3. These sets were so chosen that there was exactly one symbol common to the five sets of five symbols presented. The cards for all the trials in which a given subject was to participate were arranged in loose-leaf fashion in his compartment with the blank side facing him. The cards were numbered consecutively from 1 to 15 to correspond to trial numbering.

Communication to the environment: in each compartment there was a set of six switches, each associated with one of the six symbols. Throwing such a switch activated one of 30 lights before the experimenter.

Task: each subject was to determine in each trial the symbol which was common to the five sets of symbols and throw the corresponding switch. Subjects were allowed to change their choices prior to the end of the trial. When each subject had thrown a switch (not necessarily the correct one), the experimenter verbally ended the trial.

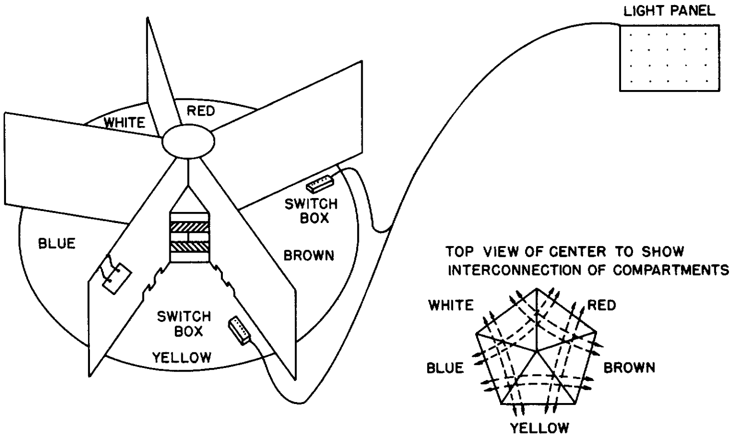


Fig. A1.1

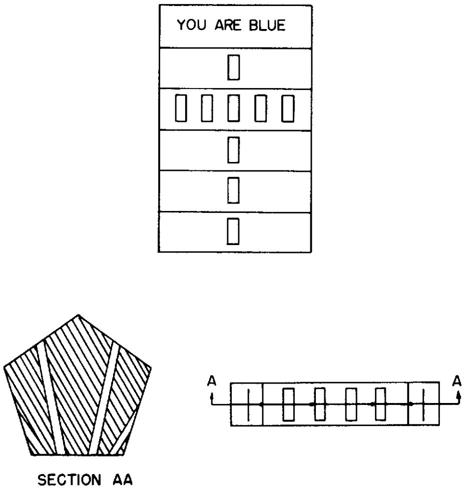


Fig. A1.2

Fig. A1.3

Networks studied:	Experiment 1:	circle (0)	5 groups
		chain (0)	5 groups
		Y	5 groups
		star	5 groups
	Experiment 2:	circle (0)	4 groups
		chain (0)	4 groups

Network positions were rotated through the color positions of the compartments in order to minimize color influence.

Experimental run: each group was given 15 consecutive trials on one network and the subjects were not used again. Each run took approximately 50 minutes. Following the trials, the subjects filled out a questionnaire.

Subjects: Experiment 1, volunteer male M.I.T. third- and fourth-year students drawn from the "humanity" courses. Experiment 2, volunteer Radcliffe College (female) students from one lecture course in Personnel Administration.

Data record: message cards, time to complete a trial as recorded by the experimenter using a stop-watch, the answers signalled by the subjects and recorded by the experimenter, and answers to a questionnaire given at the end of the experimental run.

Instructions

We've asked for your help today in an experiment on the ability of groups to solve abstract problems. This question is a basic one in any research team or other groups organized for solving problems.

Now, before we get started, let me lay down one general rule. Once we get under way, please don't talk to any other member of the group. Any conversation can throw the results off considerably. That's the only general rule.

Before starting the final experiment, we want to familiarize you with the problem you'll be doing. So, we're going to have each of you do, alone at first, what you'll later be doing together.

Each of you will get five large cards, on each of which will be five symbols like these. There is one symbol and only one which appears on all five cards. Your job is to find out what the common symbol is. When you find it, raise your hand.

Questions?

O.K. When I say go, turn over the cards, find the common symbol, and when you've got it, raise your hand.

(Time interval for sample run.)

Now we come to the main problem. The puzzle is the same, but this time, instead of having five cards apiece, each of you will have only one card. Your job, then, is to find out with the help of the others on your team what the common symbol is.

You still can't talk to one another, but you can communicate by writing messages on these little cards and passing them to your neighbors through the appropriate holes in this apparatus. But again, as you see, you can't send messages to everybody, only to those to whom you have open channels. Look in your booth now and see which channels are open. For every open channel to someone, there is an open channel from him. That is, you can get messages from anyone you can send to.

You will find large cards with symbols posted on the wall and plenty of small cards for messages. At the "go" sign, turn over the first large card and then send any messages you want to the men to whom you have a channel. Each of you, of course, will have a different symbol card since there is only one common symbol. Your job as a team is to find the common symbol.

You must not pass the same message card along. You can copy any messages you get and pass the copy along, but you can't sent on the same card you have received from someone else. And you can write any message you want to. Each man's message cards are in his own color.

Your job is not done until everyone on your team has the answer. Then, and only then, is the puzzle solved. When you have the right answer, you can pass it along. So when anyone thinks he has the right answer, he can push the proper switch in his booth and then can go on working. When I see all five lights on my panel, I'll know the job is done.

You can push only one button at a time, so if you change your mind about the answer, switch off the first guess and switch on the second.

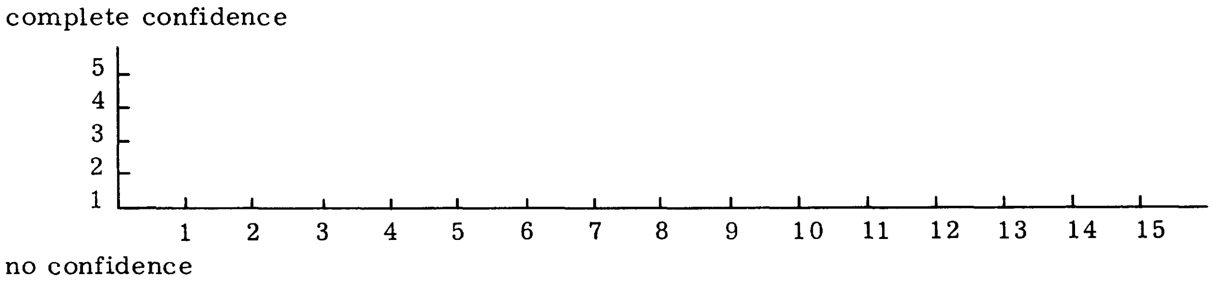
Your team will be competing with other five-man teams to see how long it takes you to get the answer. The important thing is to get the answer in as short a time as possible. The shorter the time, the better your team's score.

Start when I give the buzzer signal and stop when I give the buzzer signal. Ready?

After first trial: (1) Put a rubber band around all the messages you have received. (2) Mark the top one "Trial No. 1," and (3) Drop them in the basket. (4) Turn all your switches back to the "off" position.

Questionnaire

1. How confident are you (check on the line below) that your group got all the answers right?



2. Describe briefly the organization of your group.

3. Did your group have a leader? If so, who?

4. Was there anything at any time that kept your group from performing at its best? If so, what?

5. Do you think your group could improve its efficiency? If so, how?

6. How many more problems do you think it would take before you would get "fed up"?

0 25 50 70 100 or more

7. Rate your group on the scale below.

extremely poor mediocre average better than average excellent

8. How did you like your job in the group?

disliked it liked it

9. a. Who had the best job?

   b. Who had the worst job?

10. a. Check below the proportion of necessary messages you personally sent on the last five trials.

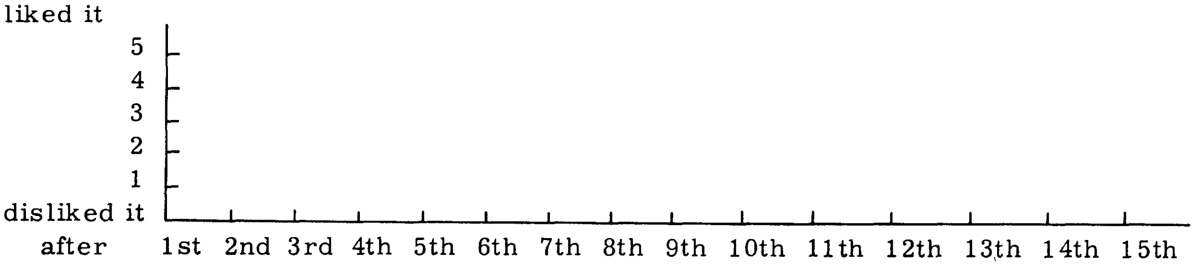
    0 all necessary

    b. Check below the proportion of necessary messages you personally received on the last five trials.

    0 all necessary

11. Do you think you solved the problem in the fewest messages possible?

12. See if you can recall how you felt about your job as you went along. Draw the curve below.



2. Experiments 3 and 6: Noisy-Marble

Style A table.

Message cards: 8 inches by 1.5 inches, similar in color to the compartment in which they were placed.

Communication within the group: free of constraints other than that the messages be written.

Input information: at each trial each subject was given five of six colored marbles. These were so chosen that there was but one color common to the five sets of five colors presented. The colors used during the first 15 trials may be described as red, blue, black, yellow, green, and white. During the last 15 trials the marbles were not solid colors but were mottled, streaked, and so forth. These have been described as green, blue, aqua, brown, white, and amber. In each compartment there were 30 small closed boxes, numbered consecutively, each containing the five marbles of the corresponding trial.

Communication to the environment: from each compartment there was a rubber tube (sufficiently large to take a marble) running to a container before the experimenter.

Task: each subject was to determine in each trial the color of the marble common to the five sets and send the corresponding marble down the tube to the experimenter. The subjects were allowed to change their choices by sending a second marble down the tube; the last one was considered the final choice. When each subject had dropped at least one marble down his tube (not necessarily the correct one) the experimenter verbally ended the trial.

Networks studied:	Experiment 3:	circle (0)	4 groups
		chain (0)	4 groups
		star	4 groups
	Experiment 6:	circle (0)	4 groups
		pinwheel	4 groups
		star	4 groups

In contrast to the other cases, the star groups of Experiment 6 were told at the end of each trial the number of wrong answers and how many different incorrect answers had been given.

Experimental run: each group was given 30 consecutive trials on one network, and the subjects were not used again. Each run took approximately 2 hours in Experiment 3 and 2 1/2 hours in Experiment 6.

Subjects: Experiment 3, volunteer M.I.T. undergraduates, most of whom were in their third year. Experiment 6, enlisted army personnel from Fort Devens, selected from those whose scores in the General Classification Test were in the upper three categories, and enlisted naval personnel from the First Naval District Receiving Station. All subjects were given the Ishihara color-vision test, and those who failed were rejected.

Data record: message cards, time to complete a trial as recorded by the experimenter using a stop-watch, the written record of the marbles sent down the tube, and answers to a questionnaire given the subjects after the experiment.

Instructions

We've asked for your help today in an experiment on the ability of groups to solve abstract problems. This question is a basic one in any research team or other groups organized for problem solving.

Now, before we get started, let me lay down one general rule. Please don't talk to any other member of the group. Any conversation can throw the results off considerably.

Before starting the actual experiment, we want to familiarize you with the problem you'll be doing.

Here you see five boxes, in each of which are five marbles. There is one marble and only one which appears in all five boxes. Your job is to find out what that common marble is. When you see it, raise your hand. Now let's go into the other room and begin.

(Time interval for move)

Take any seat around this table.

The problem is the same, but this time, instead of seeing all five boxes, each of you will see only one of these boxes. Your job, then, is to find out, with the help of the others on your team, which marble appears in all of the boxes. Just one kind of marble will be common to all your boxes in each trial, so there will be just one correct answer each time. To find this answer, you still can't talk to one another, but you can communicate by writing messages on these little cards and passing them to your neighbors through the appropriate holes in this apparatus. But, as you see, you can't send messages to everybody, only to those to whom you have open channels. Some channels have been blocked deliberately, so don't worry if they're not all open. Look in your booth now and see which channels are open. You'll notice that anyone you can send to can also send messages to you; that is, the channels all work two ways.

To keep the messages straight, the messages you send out should all be on cards of your own color (show them sample cards), so if you receive a message that you want to pass on, you must copy it over on one of your own cards and then send it. Whatever message cards you receive, you must keep, because they're of someone else's color and you're not supposed to send them.

You will find the boxes of marbles numbered by trials and the small cards for messages on your desks. At the "go" sign, open the first box. You may then send any messages you want to the people to whom you have channels. Each of you, of course, will have a different group of marbles, and there is only one common marble. Your job as a team is to find the common marble.

When you have found which marble is the common one, take it out of the box and drop it in the tube in front of you. It will roll into a box that I have here, so I can see how many have gotten the answer.

Your job is not done until everyone on your team has the answer. Then, and only then, is the problem solved. When you have the right answer, you can pass it along if you wish. So when anyone thinks he has the answer, he can drop the proper marble in his tube and then go on working. When I see a marble from each of you, I'll know that the job is done and tell you to stop.

If you should change your mind about the answer after dropping a wrong marble through the tube, you may send the correct marble along after it, but any errors, of course, must be counted against the group performance.

Don't open the box for the next trial until I give you the signal to start. Your team will be competing with other five-man teams to see how long it takes you to get the answer. The important thing is for all five of you to get the answer in as short a time as possible. The shorter the time, the better you are doing. Any questions? O.K. Open the boxes labelled "one" and start to work.

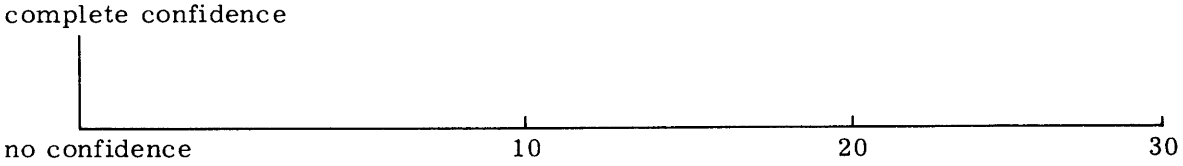
After first trial: (1) Put a rubber band around all the messages you have received. (2) Mark the bunch "Trial No. 1." (3) Drop them in the bag to your left.

Ready for trial two? Remember, you are trying for speed.

Questionnaire

Date______________ Color______________ Position______________

1. How confident are you (draw the curve below) that your group got all the answers right?



2. Describe briefly the organization of your group.

3. Did your group have a leader? If so, who?

4. Was there anything at any time that kept your group from performing at its best? If so, what?

5. Do you think your group could improve its efficiency? If so, how?

6. How many more problems do you think it would take before you would get "fed up"?

0 10 20 30 40 or more

7. Rate your group on the scale below.

extremely poor | mediocre | average | better than average | excellent

8. How did you like your job in the group?

disliked it | liked it

9. a. Who had the best job?________ b. Who had the worst job?__________

10. See if you can recall how you felt about your job as you went along. Draw the curve below.

liked it

disliked it 10 20 30

3. Experiment 4: Action-Quantized Number

Style B table, compartments unpainted.

Message cards: 8 inches by 1.5 inches, in red, blue, yellow, brown, and white, one color to each subject, who were designated by these colors in the experiment. These cards were printed as shown in Fig. A1.4, and they were numbered serially within each trial.

Communication within the group: the subjects were allowed to write only numbers in the indicated positions on the cards, thus allowing messages of the form, "From red that blue has the number x," where the colors refer to the designating colors of the people. Furthermore, the message sending was action-quantized (as described in sec. II.3.4), and each subject was required to send one and only one message card on each act. In addition, he was required to send all the information he had at the time, even if he knew it was redundant.

Input information: on a set of loose-leaf cards, similar to that used in Experiments 1 and 2, each subject was given one number between 0 and 99 for each trial. The backs of the cards were serially numbered to correspond with trial numbers.

Communication to the environment: in each compartment there were two buttons, one labelled "ready" and the other "answer." The ready button was pressed when the subject was ready to send a message, and it activated his part of the action-quantization mechanism. The answer button was pressed when each subject had fulfilled his portion of the task.

Task: each subject was to learn what the other four had as input information. When he knew this, he was to press the answer button. When each subject had pressed the answer button, a series of 5 relays was closed. A buzzer was then activated, signalling the end of the trial.

Communication from the environment: each of the ready buttons activated one of five relays in series, so a circuit was closed when all the subjects had pressed their ready buttons. This activated a 2-sec time delay relay which in turn activated a bell which was the signal for the subjects to send their messages. The delay was introduced because we discovered in pilot runs that without it the subjects rushed at a very hectic speed. Apparently, the last subject to decide did not like to be the last one.

It was found necessary to put all the relays involved in these two systems in a different room from the subjects because the clicking served as an uncontrolled and undesirable communication.

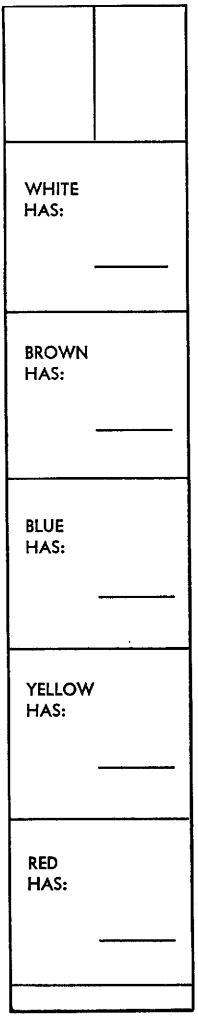


Fig. A1.4

Networks studied:

circle (0)	5 groups
circle (x)	10 groups
chain (0)	5 groups
chain (x)	20 groups
pinwheel	10 groups
barred circle	10 groups
wheel	10 groups
totally connected	10 groups
alpha	10 groups

In each case, except 10 of the chain(x) groups, the subjects were told the minimum number of acts in which a trial could be completed. The exceptional chain groups were told that the minimum number of acts was 3, when 5 was the actual number.

Experimental run: each group was given 25 consecutive trials on one network, and the subjects were not used again. Each run took approximately two and one-half hours.

Subjects: Same categories as Experiment 6.

Data record: message cards, time each act occurred as recorded on an Esterline Angus pen recorder (see Fig. A1.5), time each subject signalled the answer as recorded on a pen recorder, and answers to a questionnaire given after the experimental run.

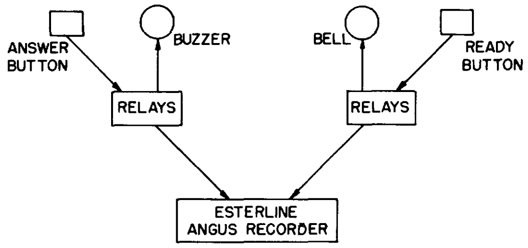


Fig. A1.5

Instructions

We're asking your help today in an experiment in which we want to discover the methods which people will use to communicate with one another to solve problems. These problems involve the sending and receiving of information by each person in the group. We're not interested in your individual abilities but in the performance of the group as a whole.

Now, before we get started, let me lay down one general rule. Please don't talk to any other member of the group. Any conversation can throw the results off considerably.

Before starting the experiment, we want to familiarize you with the job you'll be doing.

When we go into the next room, you will be seated around a table divided into five sections. You will each see a card with one number written on it which is circled in red. The job is for each of you to find out what all the numbers are. Obviously, you must communicate with each other in some way. There will be no talking. The only way for you to give or receive any information about these numbers is by passing messages to each other. There will be many ways in which messages can be passed. Some of these ways are better than others. Your problem is to work out the most effective way.

The job is for each of you to figure out the best way to send your messages so that every one of you will get all of the numbers.

You will have in front of you a board through which you will send and receive messages. Enlarged, it looks like this (show board). These are the channels through which you pass messages. In some sections all the channels may be open. This means that you can pass messages to the people at the other end of the channel. In other sections, channels may be blocked (block blue and yellow with chalk). Now, red cannot send to blue or yellow.

There are colored message cards on which you send these messages. For each experiment you will use one packet of message cards (show packet), the color depending on the color of your section. Each card in the packet is numbered from one through eight. This is the order in which you must use the cards.

To keep the messages straight, the information you send out must be on cards of your own color. So if you receive information that you want to pass on, copy it over on one of your own cards, and then you can send it. Whatever cards you receive, you keep, because they are of someone else's color and you're not supposed to send them.

Now here's how it is done: when I say "start," write down your number on the designated place on the message card of your own color. Then press the red button, which you will find on the black box in your section, until the light goes on. When you have pressed the button, wait. A bell will ring as soon as every single one of you has pressed his button. The bell is the signal for you to send your message card. Then the next time you want to send a card, go at it this same way. First, fill in all the numbers you know on one of your own cards. Then press the red button and wait until the bell rings. Then you can send.

Every time you send a card, you must have on it all the numbers you know at the time. No matter if you have sent part of the information to the same person before, you must again include this part in your next message card. You must also send a card every time the bell rings, even if you have nothing new to send. Do not worry if you do not receive a message every time the bell rings. It is possible that you may not receive a message, or you may receive only one, or you may receive more than one.

Let me repeat: you must send a card every time the bell rings. Every card you send must always contain all the information you have at the time. There are no exceptions to these rules.

Just as soon as you know what all five numbers are, press the answer button. Then after pressing the answer button, fill out the next card, wait for the bell, and pass the answer on to someone else. The problem is not over until all of you have pressed the answer button and a buzzer has sounded.

You must keep sending message cards until the trial is finished. This may mean that you keep sending the same stuff to the same people several times. Don't let this worry you.

The idea in this experiment is for all of you to get the right answer by passing the least number of message cards. The number will depend on the method you use in sending your cards. Some methods are more efficient than others. You should try to find the best method, thereby reducing the number of message cards to a minimum. It is possible for all of you to have all the numbers by sending only three message cards. The only way you can do this is by taking your time and thinking things out. Your group is competing against other groups who have also worked these same problems. Your goal as a group is to do better at this than any other group.

Again, here are the rules which must be followed: (1) Do not talk. (2) Send only cards of your own color. (3) Include all the information you have on each card. (4) Press red button once as soon as you decide where to send your message. (5) Send card only when the bell rings. (6) Press answer button just as soon as you know all five numbers.

Now, is everything clear? Are there any questions? When I say start, turn down your first number card, look at your number, and begin.

(Time interval for trial)

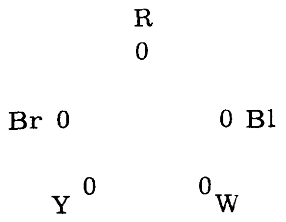
Stop! This trial is over. Pick up the messages you have received and all of the remaining message cards you have not used and put an elastic band around them. Mark the number "one" in the upper left-hand square of the top card in your pack. Now drop the bundle of cards in the brown paper bag at your left.

Ready for trial number 2? Start a new bundle of message cards. Take your trial 2 number card, turn it down, look at the number, and start.

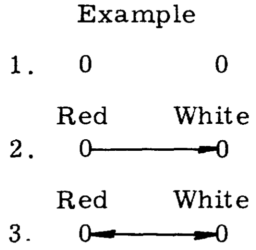
Questionnaire

Date________________ Color________________

1. Show who can communicate to whom.



Example



2. Did your group have a leader?______ If so, who?________________

3. Was there anything at any time that kept your group from performing at its best?

______________________________ If so, what?__________________________

4. Do you think your group could do better? If so, how?

5. How many more problems do you think it would take before you would get "fed up"? Circle the number.

0 10 20 40 60 80 100 More

6. Circle the words which describe how well your group did on this experiment.

<u>Extremely poor</u> <u>Mediocre</u> <u>Average</u> <u>Better than average</u> <u>Excellent</u>

7. Who had the most interesting job?________ Who had the dullest job?________

8. How did you like your job?

9. Check in each column how you felt about your job as the experiment proceeded.

	5th trial	10th	15th	20th	25th
liked it very much					
liked it					
disliked it					
disliked it very much					

4. Experiment 5: Action-Quantized Marble

Style A table.

Message cards, 8 inches by 1.5 inches in colors red, blue, white, yellow, and brown, one to each subject.

Communication within the group: action quantization as in Experiment 4, except that the message content was unconstrained.

Input information: same as trials 1 to 15 of Experiments 3 and 6.

Communication to the environment: same as in Experiment 4 but without the answer circuit.

Task: same as in Experiments 3 and 6.

Networks studied:	circle (x)	6 groups
	chain (x)	6 groups
	star	7 groups

Experimental run: each group was given 15 consecutive trials on one network, and the subjects were not used again. Each run took between two and three hours.

Subjects: paid undergraduate M.I.T. students employed through the M.I.T. employment office at standard student rates.

Data record: message cards, record of the answers given by each subject on each trial, the time for each act, and answers to a questionnaire given at the end of the experimental run.

Instructions

We're asking your help today in an experiment in which we want to discover the methods which people will use to communicate with one another to solve problems. These problems involve the sending and receiving of information by each person in the group. Your group will be competing with other previous groups to see how fast you can solve these problems.

Now before we get started, let me lay down one general rule. Please don't talk to any other member of the group. Any conversation will throw the results off considerably.

Before starting the actual experiment, we want to familiarize you with the problem you'll be doing. Here you see five boxes, in each of which there are five marbles. There is one marble, and only one, which is in all five boxes. Your job is to find out as quickly as possible what the common marble is. Now, will you look and, when you see it, raise your hand.

When we go into the next room you will be seated around a table divided into five sections. Each section of the table will be designated by a color (show diagram). You will each see a stack of little white boxes numbered from one to thirty. Number one box will contain the marbles for trial number one, number two box, the marbles for trial two, and so on. Each of these boxes contains five colored marbles. The job is for each of you to discover what color marble is common to all the five boxes in each trial. Obviously, you must communicate with each other in some way. There will be no talking. The only way for you to give or to receive information about each other's marbles, or about anything else, is by passing messages to each other. For this purpose each of you will be provided with message cards like this (show message cards). The color of the message cards you have corresponds to the color of the section you are sitting in. You are permitted to write anything at all you wish on these cards. These cards come in packets of ten. Each packet is numbered from one to ten. You must always use these cards in the order in which they are numbered. Packets for the following trials are also numbered in this fashion.

When you are sitting at your section of the table, you will see directly in front of you a board through which you can send and receive messages. Enlarged it looks like this (show board). These are the channels through which you pass messages. But, as you see, you can't send messages to everybody – only to those to whom you have open channels. Some channels have been blocked deliberately, so don't worry if they're all not open. For instance, if the channels to yellow and white are blocked, you will not be able to send to them directly. The only way to get your message to them is to send it through one of the open channels and hope those persons in turn can send it on.

To keep the messages straight, the messages you send out must all be on cards of your own color. So if you receive a message that you want to pass on, you must copy it over on one of your own cards and then send it. Whatever cards you receive, you must keep, because they're of someone else's color.

Now here's how it works: When I say, "Open box number one and start," open box number one, see what marbles you have, and then write out on card number one the first message you want to send. Then decide which open channel you wish to send your card through. We want all of you to send your messages at the same time. So when you are ready to send, press the red button which you see mounted on that little black box on your right. That little light will go on to show if you've pressed the button firmly enough. When you have pressed the button and have seen the light go on, wait. As soon as every single one of you has pressed his button, a bell will ring. The bell is the signal for you to send your message card. For the next message you want to send, go at it this same way: Write out your message on card number two, decide where you will send it, press the button, and wait until the bell rings. Then you can send. You may send out only one message card every time the bell rings.

You must send a card every time the bell rings, even if you have nothing new to send. If you want, you can simply send a blank card, or whatever you want to send as a message. But always send a card.

Do not worry if you do not receive a message every time the bell rings. It is possible that you will not receive a message, or you may receive only one, or you may receive more than one.

Let me repeat: you must send a card every time the bell rings.

Just as soon as you know what color marble is common to all five boxes, press the answer button which is shown here. (Experimenter indicates position of answer button in table diagram.) However, your job is not done until everyone on the team has the answer. Then, and only then, has your group solved the problem. When you get the right answer, you may pass it on if you wish. So when you get the answer, keep on working and send messages until the end of the problem when I tell you to stop.

When you press the answer button, you are required to write down on a piece of paper the answer you got. If, later on in the problem, you decide you were wrong, simply draw a line through the answer you wrote down and write the new answer beside it. Be sure always to press the answer button as soon as you get the answer, because we want to know which person gets the answer first on each trial.

Your team will be competing with other five-man teams who have also done these problems. The objective in these problems is to get the answer in the shortest time possible. The shorter the time, the better you are doing. Remember you are working as a group and competing with other groups.

Any questions? O.K. Let's go into the room where we will work.

. . . .

Take any seat around this table. Look at the board in front of you and see who you can send to. If you have any questions to ask, do not say them aloud, but put up your hand and I will come over to you. Look at the black box marked "ready." That is the red button you push every time so you can send the messages cards. Notice the stack of little white boxes which contain the marbles for each trial. Notice the message cards They are in the color of your section. Notice the black box on your left which is marked "answer." Remember that you must push the answer button just as soon as you get the answer. Also, don't forget to write down the answer when you get it. Remember you must send one card every time the bell rings. Never forget to push the ready button or you will hold up your whole group while they wait for you. Notice that your message cards are numbered in order. This is the order in which you will use them.

To start each trial: Is everybody ready? O.K. Open box number (use trial number) and start. Remember that you are trying for speed and that you are competing with other groups who have also worked on these problems.

After each trial: O.K. Stop. Put an elastic band around all the messages you have, mark the bunch "Trial number, (use number of trial completed)" and put them in the bag.

The questionnaire used in this experiment is the same as the one used in Experiment 4.

APPENDIX 2 – OCTOPUS*

1. Design Criteria

Past experiments in this laboratory have suggested that for certain purposes, modifications in experimental technique would be desirable. These modifications were desired, in particular, to make possible two features: short actual times for the running of any experimental trial to enable the collection of large amounts of data, and the elimination of the tedious and time-consuming recording and analysis of the data by hand which has been necessary in the past. In an effort to design an apparatus to fulfill these criteria, we made the following list of desirable features which any such apparatus should possess.

a. A single trial should be brief, on the order of one minute, to permit the running of many trials within the usual two- or three-hour experimental period.

b. A large number of messages should be used for each trial to permit easy statistical treatment of message flow.

c. The number of possible categories of messages should be reduced to a minimum to facilitate subsequent analysis.

d. The mechanical operation of any such device should be simple to learn, so that the individual's learning time to operate the apparatus skillfully will not be a major factor.

e. The communication network should be easily variable, at any time, with or without the subjects' knowledge. Also, any desired communication network should be possible.

f. Repeated trials under a given set of boundary conditions should not become trivial, so that a single group of subjects may be used for a relatively long time.

g. A complete record of all occurrences within the group should be made automatically and in such a way that it may be reconstructed in its major time sequence at a later date. This reconstruction should be as automatic as possible, and provision should be made for the major part of the analysis to be done mechanically or electrically and in a relatively automatic manner.

h. As far as possible the subjects must be allowed free choice of action and, in general, must be allowed to behave as human beings in a relatively normal manner.

---

*The design criteria mentioned below are based on unpublished work of Dr. O. H. Straus; the preliminary design of the apparatus with the exception of the recording unit was executed by Dr. Straus.

i. The equipment should be relatively simple to use; it should be sturdily constructed and easily moved from place to place without dismantling or excessive precautions.

We feel that in many ways the apparatus described below, which has been nicknamed "Octopus," comes very close to fulfilling these requirements. In some ways it is perhaps extremely restricted in nature; it seems impossible to avoid this and at the same time to get all the other advantages from an experimental point of view which are inherent in a simplified and highly automatic apparatus of this nature.

2. Description of the Apparatus

2.1. General

Octopus consists of a central control unit, a recording unit attached to this, and five stations. It is designed to be used with five experimental subjects or less. When it is in use, each subject is seated at one of the stations, which are all interconnected through the central control unit. The transmissions made from one subject to another with this device are made by binary switching operations, which permit free choice of message (within the given categories), destination, and time of transmission. Messages successfully received are acknowledged to the sender. The subjects need not be given any information about the presence or absence of links for a particular experiment, in which case they must deduce this information from the presence or absence of acknowledgments to their transmissions.

Messages which may be sent may be written in the general case as "JKL(0,1) (t,r)," read as "node J sends to node K the information that node L has a 0 (or a 1). This message is a transmission (or a reception)." This last symbol is necessary, since all transmissions do not necessarily result in receptions. For the particular case of a five-man group with these restrictions, the matrix referred to above may be simplified to a three-dimensional matrix, 5 by 5 by 5, with entries identifying the message as 0 or 1, t or r, or "no message." This specialized matrix is referred to as the "master" matrix, and the apparatus is laid out with this matrix in mind.

2.2. Subjects' Stations

Each subject has in front of him a panel, called his station, carrying two matrices of switches. On his right is a matrix representing a section of the master matrix in the KL plane at the value of J for that station. The main diagonal and the row of entries

KK are omitted, since a man will not send to himself, nor will he send to K that K has a certain piece of information. On the subject's left is a similar matrix representing a section of the master matrix in the JL plane at the value of K assigned that station and with similar omissions. The right-hand matrix consists of 16 momentary-contact, spring-return, neutral-center switches, which are used to transmit various messages to any chosen destination. The choice of switch determines the JKL part of the message, and the direction in which the switch is thrown determines the 0 or 1. The left-hand matrix consists of 16 double-pole, neutral-center, permanent contact switches, used for the acknowledgment of the subject's receipts. Each switch has two possible positions, corresponding to the sending or receiving of the information 0 or 1, and each switch has two signal lights which signal the transmission or reception of information. The method used in sending is as follows: J sends to man K that man L has a 0 or a 1 by pushing the corresponding switch. This action, if such a link exists, lights the corresponding light on man K's receive panel, and man K acknowledges receipt by throwing his receive switch. This permanently lights the signal light on J's transmit matrix, allowing a check on both the receipt of the message and the correctness of the reception. Figure A2.1 is a photograph of a typical station, showing the switch matrices and labeling.

In addition to the send and receive matrices, each man has a pair of lights giving him his input (0 or 1) for the particular trial, a switch and signal lights with which to transmit his answer or output, and two lettered panels which light up to show starting and stopping instructions for each trial. A trial is run by giving to each subject an input of 0 or 1. The task of the group is to find out whether the sum of the inputs is even or odd. This answer may be required to be known to any or all of the subjects and is signaled with the station's output switch. The trial ends automatically when the required number of stations have submitted the answer.

2.3. Central Control Unit

The control unit is designed to be operated by the experimenter. All transmissions from one station to another go through this unit, and the communication network may be varied at will by throwing different switch combinations in the control matrix. This may be done without the subjects' knowledge, if so desired.

Figure A2.2 is a photograph of the control unit. Also shown here are the other controls: a bank of switches to determine the inputs to each station and a similar bank to control the output required from each subject for automatic termination of the trial, the

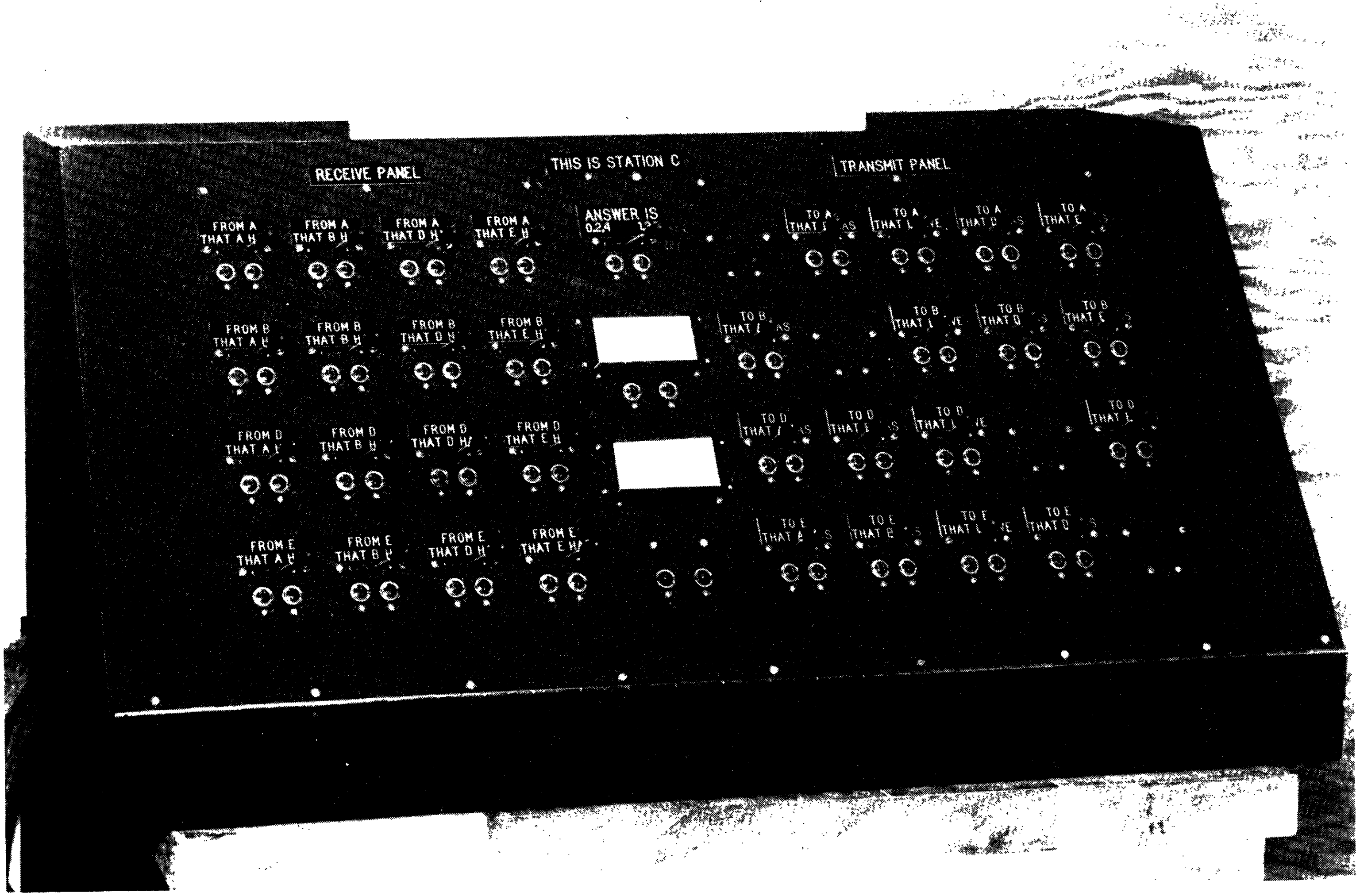

RECEIVE PANEL
THIS IS STATION C
TRANSMIT PANEL
FROM A THAT A H
FROM A THAT B H
FROM A THAT D H
FROM A THAT E H
ANSWER IS 0.2A
FROM B THAT A H
FROM B THAT B H
FROM B THAT D H
FROM B THAT E H
FROM D THAT A H
FROM D THAT B H
FROM D THAT D H
FROM D THAT E H
FROM E THAT A H
FROM E THAT B H
FROM E THAT D H
FROM E THAT E H
TO A
TO A
TO A
TO A
TO B
TO B
TO B
TO B
TO D
TO D
TO D
TO D
TO E
TO E
TO E
TO E

Fig. A2.1

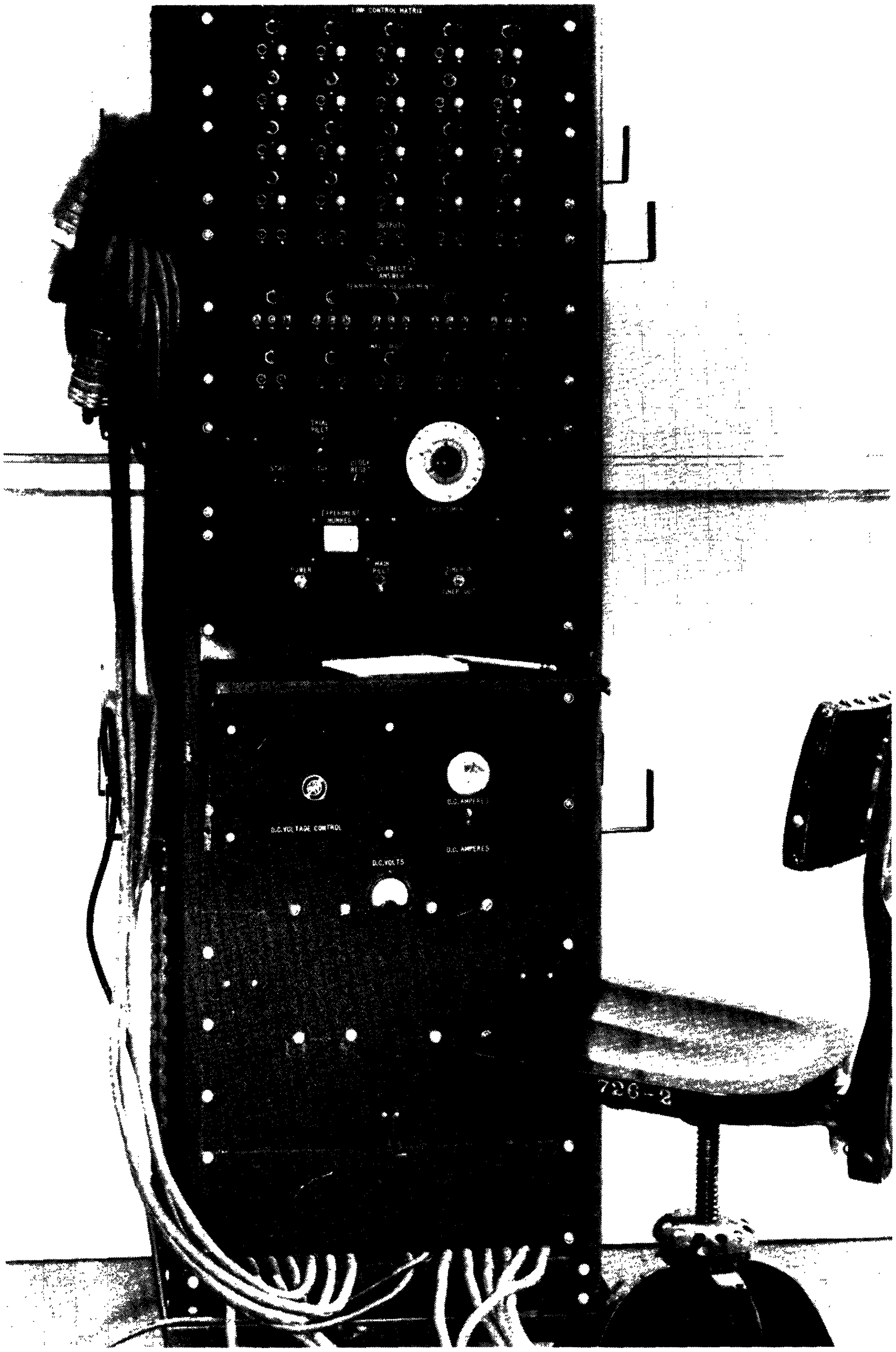

LINK CONTROL MATRIX
OUTPUTS
CORRECT ANSWER
TRIAL PULSE
START
EXPERIMENT NUMBER
POWER
MAIN PULSE
D.C. VOLTAGE CONTROL
D.C. AMPERES
D.C. AMPERES
D.C. VOLTS
726-2

Fig. A2.2

choice of requirement for each station being either the correct answer, any answer, or no requirement at all. The control unit also contains start and stop buttons, voltage controls, a timer which may be optionally used to limit the length of a run, and indicators of the correct answer and the answer arrived at by each station. Provision is made for incorporation of circuits to generate various types of noise in any or all of the links, and interlock circuits are included to prevent guessing and premature transmission of the answer. Figure A2.3 shows a view of the internal wiring of control. These methods are used throughout Octopus in an effort to build in as much strength and ruggedness as possible. The entire apparatus is designed to be operated from this control unit, and after the initial instructions are given the subjects, the experimenter may remain at the control unit. All of the data produced during an experiment, including times, experiment number, boundary and initial conditions, as well as a record of all transmissions during an experiment, are recorded automatically without requiring the attention of the operator.

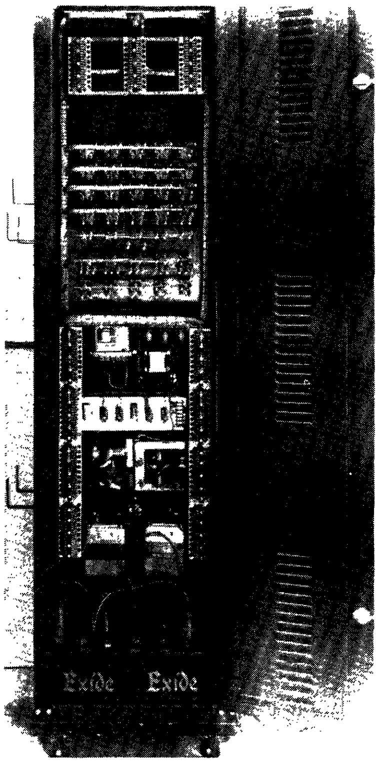


Fig. A2.3

2.4. The Recording Unit

This unit was designed to record and play back all information generated in Octopus, although the magnetic recording section has application to other uses.

The Recording Unit is pictured in Fig. A2.4. The upper portion presents information as arrays of lights, expressed in all cases in a binary code, so that on-or-off presentation is possible. The following information is presented: (a) Experiment timer, in seconds from start of experiment, automatically reset; (b) Experiment number; (c) Setting of output requirement switches; (d) Inputs given each subject; (e) Correct answer; (f) Output from each station, and time of occurrence; (g) Projection of the master matrix on the JK plane, with time of occurrence of each transmission; (h) Projection of the master matrix on the KL plane, with time of occurrence of each transmission; and (i) Pilot light, lit during the running of an experiment.

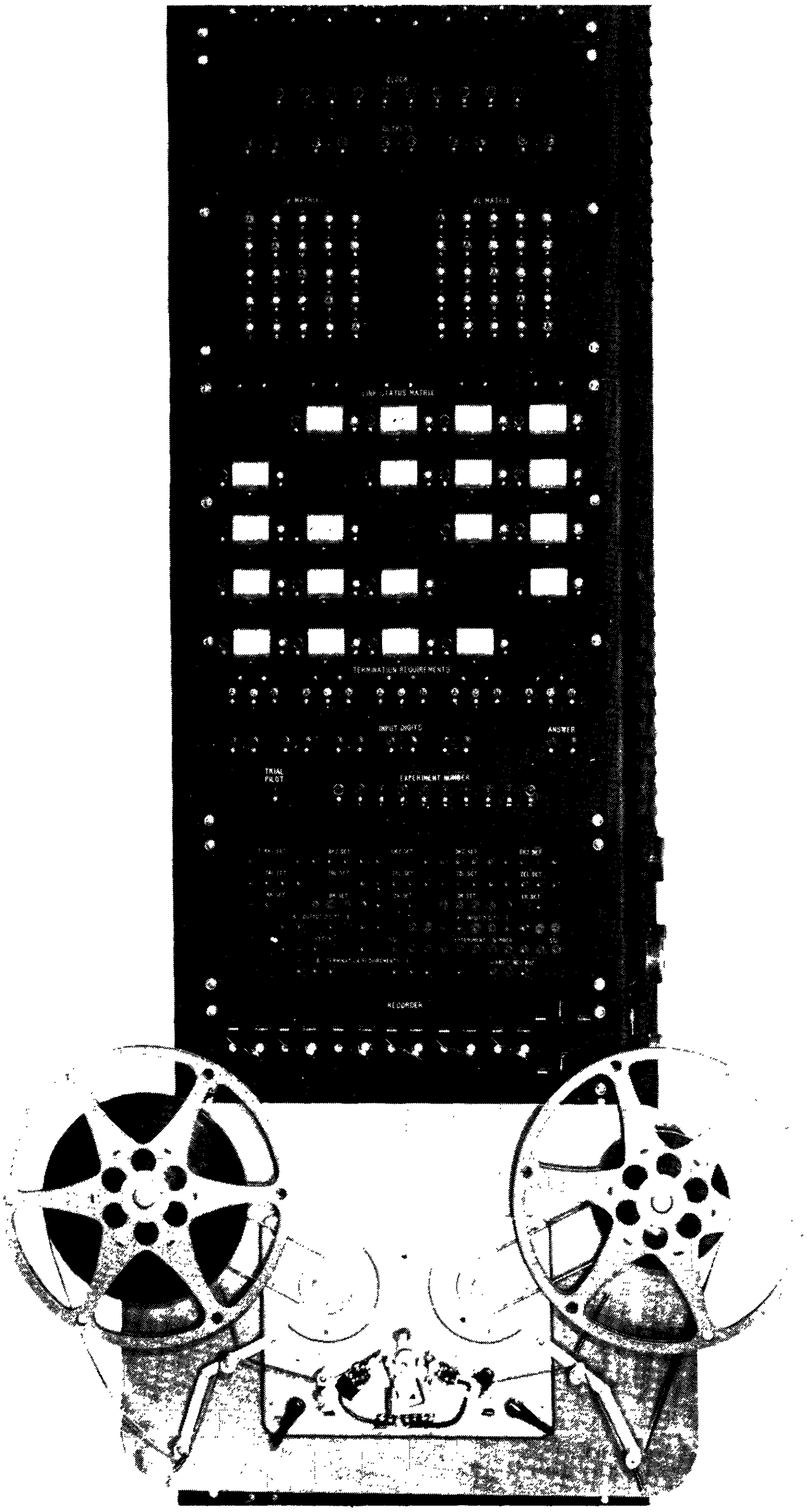

TERMINATION REQUIREMENTS
INPUT DIGITS
ANSWER
TRIAL PILOT
EXPERIMENT NUMBER
RECORDER

Fig. A2.4

Since these data are all expressed in a binary code, i.e. a light off or on, their recording and treatment is simplified. All the above circuits are led to a plug-in board in the center of the recording unit and also to the magnetic tape recording section in the bottom. The plug-in board is arranged and coded to permit the use of various computing circuits, so that sums, averages, conditional probabilities, and the like may be calculated from the data by merely plugging the correct set of relay units into the desired circuits. This should permit a great deal of the manual labor involved in the usual data analysis to be replaced by automatic operations.

Below the calculator plug-in board is a bank of switches for selecting either recording or playback, and these lead from the data presentation circuits to the magnetic tape recorder. This tape recorder was developed to solve this particular problem, but the design has turned out to have wider application; a somewhat more detailed description of this section will consequently be given. The general method used is the following: Each of the data circuits is sampled by a mechanical rotary switch after passing through the record-playback switches. The resulting pulses serve as a trigger for an oscillator which records on magnetic tape, driven by a special transport mechanism to avoid slip. On playback, the pulses are picked up from the tape, amplified, rectified, and returned to the correct circuit by the rotary switch. In each circuit the pulses close a relay which reactivates the circuit (see Fig. A2.5).

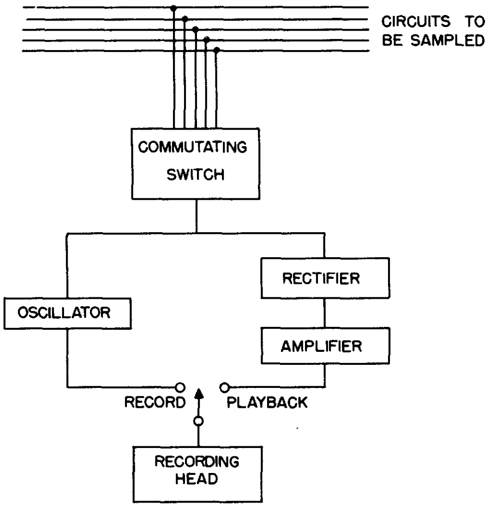


Fig. A2.5

The two limiting factors of this design are the frequency response of the magnetic recording tape and the minimum permissible sampling rate needed for the circuits involved. Some representative figures, giving maximum capacity of the system, are given below. In each case it is assumed that the information contained in each circuit or channel can be expressed as one binary digit, "yes or no."

A. Tape Speed: 3 3/4 inches per sec	3 bands frequency multiplex
Sampling Rate:	Capacity:
2 per sec each channel	150 channels
5 per sec each channel	60 channels
B. Tape Speed: 7 1/2 inches per sec	4 bands frequency multiplex
Sampling Rate:	Capacity:
2 per sec each channel	200 channels
5 per sec each channel	80 channels
10 per sec each channel	40 channels
C. Tape Speed: 15 inches per sec	6 bands frequency multiplex
Sampling Rate:	Capacity:
1 per sec each channel	600 channels
2 per sec each channel	300 channels
5 per sec each channel	120 channels
10 per sec each channel	60 channels

The figures in the table above are for one recording head and one track used on the recording tape. If several recording heads and tracks are used, the figures should be multiplied accordingly. The recording device used in Octopus samples each of the 120 data circuits twice per second and uses two recording heads and two tracks on the magnetic tape, at 3 3/4 inches per sec tape speed, recording at one single frequency.

When using 5,000 foot reels the device has the following uninterrupted playing times

15 inches per sec	1 hr, 15 min
7 1/2 inches per sec	2 hrs, 30 min
3 3/4 inches per sec	5 hrs

Thus, it can be seen that this device is adapted to recording a large number of binary bits of information in such a way that automatic playback and re-creation of the original circuit is possible. The device is quite flexible in operation and can be adapted to many different data-recording jobs, and yet it is simple, fairly rugged, and far less expensive than other electronic devices with a similar range of performance.

The sampling circuit is given in Fig. A2.6. This system is repeated for each channel of the tape recorder (two, in the example constructed) and also for each sampling frequency. Operation of this sampling circuit is as follows.

Presence of a dc potential in the circuit to be sampled causes the commutating switch contact to be energized in the commutating switch as the arm passes the appropriate contact, and this dc potential is picked up as a pulse at every revolution of the switch arm.

On playback, a train of suitable pulses is fed to the commutating switch and individual pulses are delivered by it to the various circuits. With SW_1 (see Fig. A2.6) now open, this pulse closes relay 2, and C_1 holds it closed for the duration of the sampling cycle. If another pulse is then received, it stays shut. If not, it opens. C_1 and resistances are used to adjust this delay time to the appropriate value.

The train of pulses from the commutating switch and a 1-kc pulse delivered to one pole of the commutating switch, for use as a check on synchronization, are separated by two filters (see Fig. A2.7). The dc pulses operate a trigger circuit which, when open, transmits a 3-kc signal from the oscillator to an amplifier; the resulting train of 3-kc pulses is thus amplified to a suitable voltage for the recording head, which impresses them on the tape.

On playback, the train of 3-kc pulses (and the 1-kc synchronization signals) are picked up from the tape by the playback head, amplified, and separated by filter networks (see Fig. A2.8). The 3-kc pulses are then rectified and smoothed to produce a train of dc pulses. These are separated by the commutating switch and sent to the same individual circuits from which they originally came. The 1-kc pulses are fed through the commutating switch unaltered, and if the tape and commutating switch are properly synchronized, they will appear as a 1-kc "beep" in a small monitor speaker.

In operation, the commutating switch and the wheel which drives the tape are locked on a common shaft and driven by the same motor. This insures that the wheel and the commutating switch will always be synchronized.

In the conventional tape recorder, the tape is driven by a rubber-tired wheel. This allows a certain amount of tape slip and flutter. For the applications used in this device, the tape must be driven by a completely positive drive, such that no slip between the tape and the driving wheel is possible. This is accomplished by utilizing a toothed sprocket as the driving mechanism and punching a row of holes in the tape. The tape used in this device is the regular 1/4-inch plastic base, red oxide-coated magnetic recording tape.

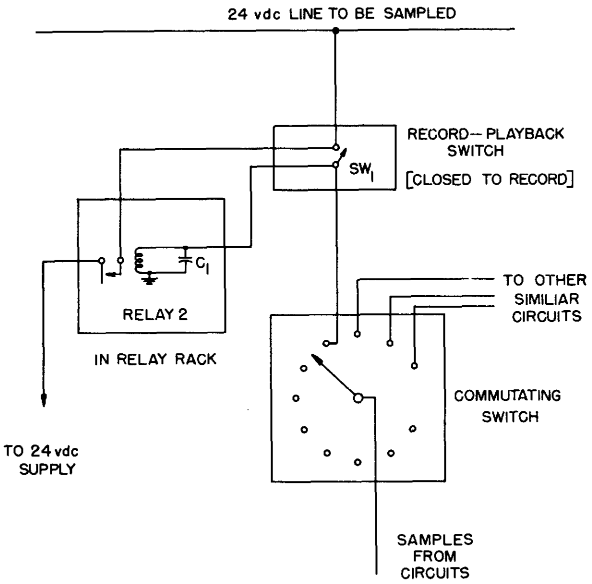


Fig. A2.6

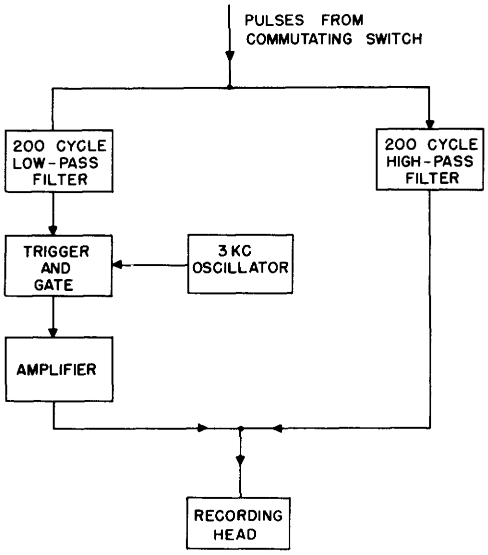


Fig. A2.7

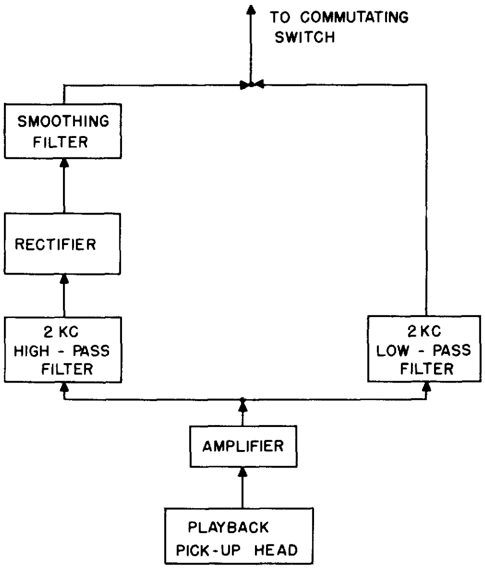


Fig. A2.8

A rotary punch and die is used to punch a row of holes (1/16 inch in diameter and spaced about 3/8 inch between centers) down the center of the tape.

Tests on the punched tape on an ordinary tape recorder (twin channel) indicate that audio fidelity is not impaired to any marked extent by the holes in the tape, provided the tape is erased after punching.

The driving sprocket is toothed to fit the holes in the tape. The teeth have the approximate shape of a paraboloid of revolution, and are equal in size to the diameter of the hole at the base, with a 0.001 clearance. The tape passes over about one third of the circumference of this wheel, so at least three teeth are fully engaged at all times (see Fig. A2.9).

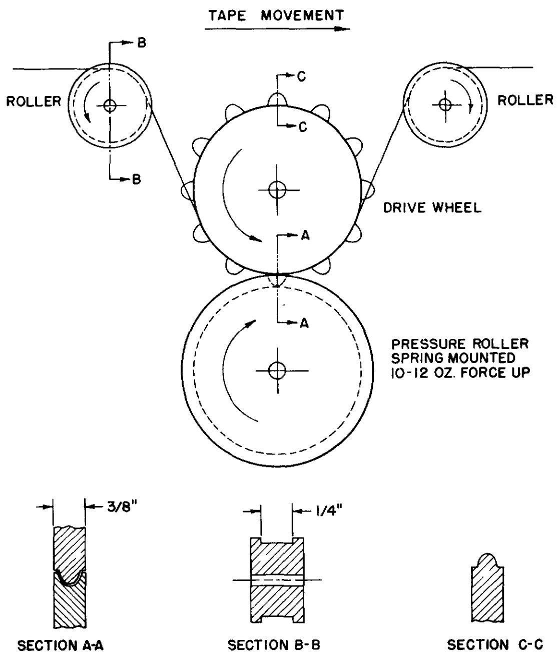


Fig. A2.9

The recording heads, one for each channel (two in this case), are mounted immediately before and after the sprocket; pressure pads and ball bearing guide rollers are mounted in a more or less conventional manner (see Fig. A2.10).

The arm carrying the recording-playback heads and the feed rollers swings down, as indicated in Fig. A2.10, to facilitate threading the tape, and the drive sprocket carries an index mark to insure accurate alignment of the tape for playback. Auxiliary units to the tape transport mechanism are the two knee-action mechanisms to insure constant tension on the tape (see Figs. A2.11 and A2.12).

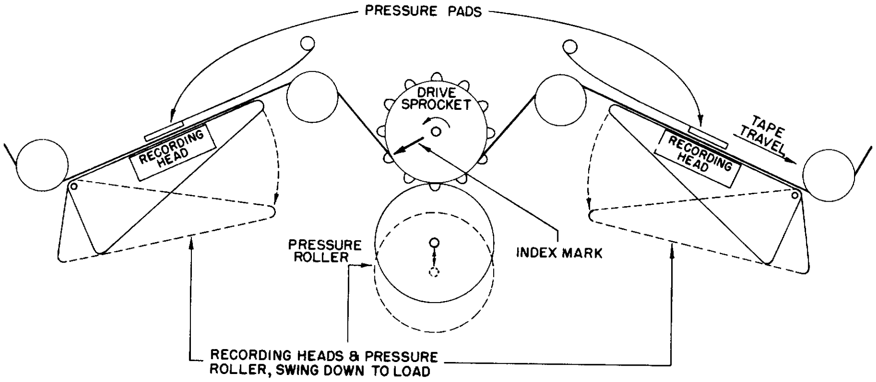


Fig. A2.10

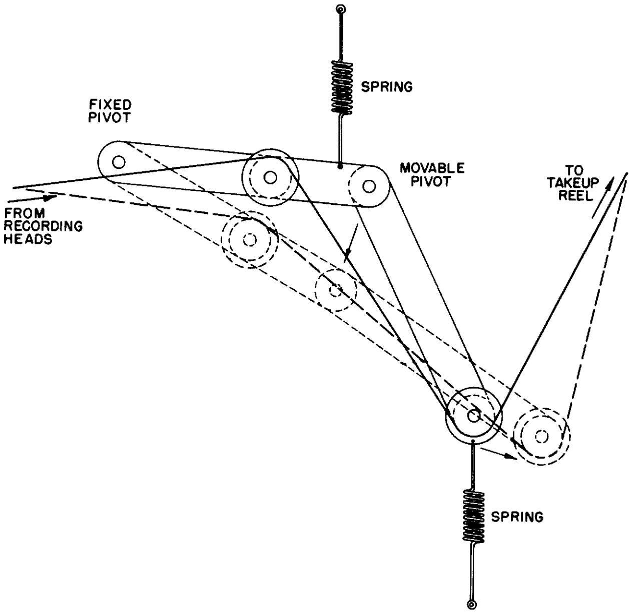


Fig. A2.11

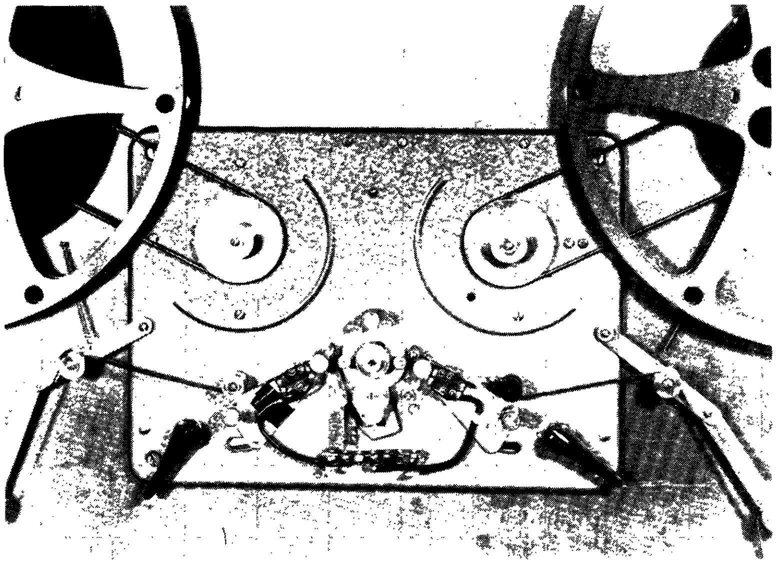


Fig. A2.12

The two knee-action units and the spring drive on the take-up reel insure even tension on the tape as it passes the heads and drive sprocket. A fairly high tension, of the order of 1 to 2 pounds pull, is maintained on the tape at all times, and the knee-action units must be quite flexible to maintain this tension during different conditions of tape loading, especially when the large 14 1/2-inch aluminum reels, holding 5,000 ft of tape, are used.

Two shorting contacts are provided on one of the guide rollers, so that in the event of tape breakage or the supply roll becoming empty, the two contacts short across the bare roller, stop the machine, and activate a signal.

3. Summary

We feel that Octopus has gone as far in the mechanization of the group communication problem as is profitable. In order to achieve the results mentioned in the first paragraph, it has been necessary to create an experimental situation for the subjects which is extremely simplified and rather far removed from any real-life situations. Such experiments are valuable in order to obtain large quantities of basic data, but future developments will be in the direction of increased face-validity and closer approximations to real-life situations. It may be possible to adapt such machines as Octopus to handle part of the recording job in such future experiments, or perhaps to use them as auxiliary units in a more complex and realistic communication problem, although these applications have not been worked out as yet.

APPENDIX 3 – THE GROUP AS A MONOCHROMATIC LINEAR SYSTEM*

1. Introduction and Assumptions

To the reader having some knowledge of electrical network theory, it is obvious that many of the formulations in the previous chapters bear some resemblances to the formal structure of electrical theory. How similar are they? To what extent can the assumption of the electrical theory be applied to what group situations? Can the formal results of electrical theory be reasonably interpreted in terms of groups?

This appendix is devoted to reinterpreting the formal structure of electrical theory, making clear what the assumptions of that theory are and how applicable they are to human groups. It will be shown that the assumptions are so stringent that relatively few human groups can be considered to satisfy them, even to a first approximation. Nevertheless, it is worthwhile developing the theory, because in the process of making the assumptions explicit, we will see what the direction of generalization must be if a theory is to be developed which does apply to human groups; and because so far as the present theory is applicable, a more general theory must include it as a special case. Furthermore, it is interesting that even in a most restricted theoretical structure one obtains phenomena which may be identified with human phenomena. A brief summary of the various sections is in order. The remainder of this section explicitly states the four basic assumptions of the theory, and each is examined as to its meaning and applicability to a human group situation. Having these interpretations in mind, we shall point out the types of groups to which the assumptions seem to apply. It will be shown that the class of groups is very limited and does not include most groups that are of interest for psychological theory. Section 2 examines verbally the type of problem which may be solved within the theory. In the third section enough mathematical terminology is introduced to formulate the properties of a node. In the next section the equations for the group are developed, a human group interpretation is given for the various symbols introduced, and the entire problem is reduced to one in matrix algebra. The mathematical solution to this problem is stated in section 5 without proof, since the techniques are well known.

---

*This appendix is a condensation and rewriting (with permission) of the unpublished manuscript, Linear Systems and Group Dynamics, by W. H. Huggins (118). Credit is due Mr. Huggins for the development of the approach described here, but he cannot be held responsible for interpretations made by the authors of this report.

In section 6, a concept of the "leader" of a group is given in terms of quantities introduced in the previous section. The seventh section presents an application of the theory to a simple example. And the final section discusses the value of the theory.

As in the previous discussion we shall take as the basic undefined notion a node having inputs and corresponding outputs. The relation between the input and the output is called the transfer function. That which is either an input or an output is called, following electrical terminology, "flux." For groups, flux is what the group processes; it may be information or it may be materials.

The following assumptions are made:

Assumption 1: Flux is a variable which may be described by a single real number.

This is a very strong assumption which, along with the second assumption, makes the resulting theory comparatively inapplicable to human groups. If we consider the case of discrete pieces of flux, it implies that these pieces or elements of flux are indistinguishable one from another, that the only information extracted from the flux by the nodes are the number of elements present. One simply counts the number of pieces of information one has received. In the case of electrical theory, these discrete elements are electrons, and we assume that the electrical components cannot distinguish one electron from another. This is, in fact, true for electrons. It is not generally true for information flow; a person can tell if two separate inputs are duplications or if they are different. A mathematical structure which is applicable to general human groups would have to be an "electrical" theory in which there are "colored" electrons, any two of the same color being indistinguishable, but two of different colors being treated differently. A discussion of such a theory was given in section IV.9.

Assumption 2: The node response is linear.

An equivalent formulation is that the superposition law holds at a node. Thus, the nodal response to two inputs together is the sum of the responses to each of these inputs separately (by Assumption 1 it is appropriate to speak of "sum"). This states that if my response to the statement, "Your house is on fire," is to call the fire department, then in a situation in which two people tell me my house is on fire, I will call the fire department twice. This is to say, I do not take into account that I have already processed or taken action on the information once, when I hear it again. More generally, a linear node is unable to learn to change its behavior on the basis of its past output. This, of course, is not generally true of human behavior, but there are cases in which it is approximately true. Some of these will be mentioned below.

These two assumptions are not necessary, as we have seen in section IV.9, provided we accept a very complex mathematical problem.

Assumption 3: Each node has the same response (transfer function) as every other node.

This is clearly not generally true, but it is a convenient first approximation, and there are situations where it seems to be comparatively true (see the discussion in section I.3.3. We may, for example, take this response to be the statistical mean of individuals from some population.

Assumption 4: The response of a node is not a function of time.

Given an input at two different times, the node will react in the same way. Specifically, there are no changes in response except those which are dependent on preceding group behavior. For example, the response of the node does not vary with increasing hunger, fatigue, or other needs. This fails to be true for human beings but it is often approximately true. At the present level of analysis, this restriction does not seem important.

From what we have said, it is apparent that our assumptions do not cover most group situations, and the question must be answered whether they include any actual situation, even as a first approximation. It follows from the first assumption that we must consider only situations in which one type of object or information is being processed by the group. Furthermore, the processing of this flux cannot generate different types of flux, for then we have a polychromatic situation, which violates Assumption 1. This is important, for a situation in which it superficially seems as if there is only one type of flux may, in fact, so far as the members of the group are concerned, generate new types. For example, suppose the group is a purchasing organization for a firm, and a single input, an order for materials, is introduced into the group. As far as this single input is concerned, the group satisfies the first assumption, but each person who has handled the order may remember that he has handled it. This memory information may be treated in one of two ways, each of which violates one of our assumptions: it may be treated as a different kind of flux from the order itself, which violates Assumption 1; or it may be treated as causing learning at the node, which introduces nonlinearity of the response and violates Assumption 2. Thus, a group situation with only one type of flux satisfies the assumption if the individuals take no cognizance of what they have processed. This seems to be possible under two conditions: either they are so uninterested in the information or materials being processed that they do not recognize the same thing when they see it again, or it is impossible for them to recognize the same thing twice. The former may occur in a clerical section of a large bureaucratic organization, for example, in the processing of checks in the Veterans Administration. The latter may occur on a production line dealing with many identical products, say, a production line of vacuum tubes, tooth brushes, automobiles, and the like. In these cases where we have but one type of flux, we may also expect that the assumption of linear behavior is not erroneous. It is not clear that any other sorts of groups satisfy the four assumptions; in particular, the groups that have been studied experimentally and reported on in this report do not.

2. Type of Problem

Before we proceed to an explicit formulation of the mathematical problem which results from these assumptions, it may be well to pause and consider the type of result to be obtained. Recall that the group can process only one type of flux or information, each member deals with the same input of flux in the same way, and no member of the group learns on the basis of past performance. Thus, we can ask no questions about group learning. However, we can ask questions of the following kind: If an impulse of flux is imposed on one node, what is the probability that some other node has received some, or all, of this flux after t time units? That is, we may express the time distribution of arrival of flux at one node when it is imposed at another node. This is a problem of considerable importance in many applied situations: How long does it take, on the average, for an industrial product to be processed through a production line when a certain given percentage are faulty and must be deflected from the normal course of production for repair or discard?

A second type of problem which may be considered and answered is the relative contribution of various nodes to the time response of the group. To increase the average speed through the network, which node would it be most important to consider? For example, in the production-line case, the product may be deflected for one of several reasons. Which of these reasons, on the average, delays the output the most? This question can be answered, and we do so in section 5.

To discuss these two questions it is necessary to use some mathematical symbolism. This involves primarily the notation of matrix algebra. No attempt has been made to carry out the argument in full detail; this may be found, for the electrical case, in such volumes as Guillemin: The Mathematics of Circuit Analysis (117), or Bode: Network Theory and Feedback Amplifier Design (115). Rather, we have pointed out the highlights of the solution and have attempted to give a verbal interpretation of them.

3. The Nodal Characteristic

The most important aspects of Assumptions 1 and 2, (i.e. there is only one type of flux, it is measurable on a continuum, and the nodal response is linear) are contained in the important conclusion that a single scalar function $\mu(t)$, known as the "nodal characteristic," completely characterizes the response of the node to any input pattern. Without this, the theory would be very complex; with it, it is comparatively simple. We may derive this conclusion and an understanding of $\mu(t)$ very easily.

Let us call the rate at which flux is being produced by a node at any instant t the instantaneous "stress" of the node. This will be denoted by S(t). In general, the stress S(t) is some function of all the inputs to the node previous to t; that is, the node has a transfer function which is a weighted memory of inputs in the past but not of its outputs, for this would violate Assumption 2. We may always subdivide the past input, f(t), into a collection of bands of input of infinitesimal width (see Fig. A3.1). Consider a band of input at time τ. Then we have a response at time t to that input which is the same as the response at time t to a unit impulse imposed at time τ, weighted by the value of the input $f(\tau)$ at that time. Now let $\mu(t)$ be the response of the node at time t to a unit impulse at time 0; then, since the response of a node is not a function of time (Assumption 4), the response at time t to a unit impulse at time τ is $\mu(t-\tau)$. So the response at time t to an input $f(\tau)$ at time τ is

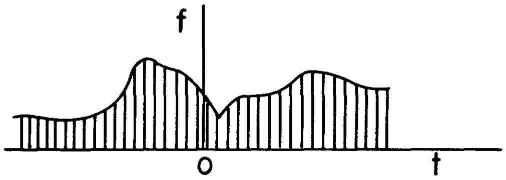


Fig. A3.1

$$f(\tau)\ \mu(t-\tau). \qquad (1)$$

Equation 1 holds only for a single band of input, and we have assumed a continuous distribution of input f(t). But since the system is linear, the response at time t is simply the sum of the response to all the infinitesimal bands of input; in the continuous case this is an integral

$$
\begin{aligned}
S(t) &= \int_{-\infty}^{\infty} f(\tau)\, \mu(t-\tau)\, d\tau \\
&= \int_{-\infty}^{\infty} \mu(\tau)\, f(t-\tau)\, d\tau .
\end{aligned}
\tag{2}
$$

This is a mathematical statement that the stress at a node is the weighted average of its previous inputs taken in a linear sense. Treating the integral as an operator on the input, it is the transfer function of the node.

For the most part, in group situations, it is very unreal to consider the flux (possibly information) to be a continuous variable. Information comes in discrete pieces or quanta, or in electrical terms, electrons. If there are sufficient such quanta, then one may treat the problem as above in terms of a continuous variable. If not, then we must define an analogous nodal characteristic as follows: Perform N experiments in which a single quantum is impressed on the node at time 0. Suppose that in these N experiments the total number of quanta emitted in the time interval from 0 to t is $NQ_N(t)$. If we call $Q_N(t + \Delta t) - Q_N(t) = \Delta Q_N(t)$, then we define

$$
\mu(t) = \lim_{N\to\infty} \frac{\Delta Q_N(t)}{\Delta t} \quad . \tag{3}
$$

Thus, $\mu(t)\ \Delta t$ is the probability that in a large number of trials, a node will emit a quantum of flux during the interval t to $t + \Delta t$ after it has been excited by a single quantum at $t = 0$.

Because of the linearity assumption, we know that if N_i quanta are impressed on the node at times τ_i, where $i = 1, 2, \ldots, n$, then the probability that a quantum will be emitted during the interval t to $t + \Delta t$ is

$$
\left[\sum N_i\, \mu(t - \tau_i)\right] \Delta t. \tag{4}
$$

Expression 4 may be readily transformed into an integral expression formally identical to Eq. 2 by writing the sum as a Stieltjes integral.

Taking Eq. 2 to be typical, we may make a transformation on it which allows an appreciable simplification in the mathematics. It is evident that the integral 2 is in some sense a product, and were it possible to transform the variables in such a fashion that we could write the transfer function of the node as a simple algebraic product, all of the mathematics of the theory would be reduced to algebraic manipulations. It is well known that the use of Laplace transforms will effect this; the result may be written as

$$\widetilde{S(\omega)} = \widetilde{f(\omega)}\ \widetilde{\mu(\omega)} \tag{5}$$

where ω is the frequency. In this notation, the parameter $\widetilde{\mu(\omega)}$ is called the nodal impedance. It represents the stress resulting from a unit flux current of frequency ω. It may also be termed the transfer function of the node in the frequency domain. We solve the remainder of the problem in the frequency domain rather than in the time domain and then finally express the answer in terms of time. In electrical theory it is easy to give a physical explanation for this transformation to the frequency domain; for groups any explanation seems artificial. Thus, it is suggested that this be accepted as a convenient mathematical manipulation and no meaning be attached to it.

4. Network Equations

If we now assume a collection of nodes such that part of the output of one is part of the input to another (a group of people possibly), then we must formulate equations which describe the interaction of these nodes when some flux is imposed on members of the group. The corresponding electrical problem is what happens if we introduce some current into an electrical network. Let there be n nodes.

Since the flux has been assumed to be completely describable by a single number, we can only distinguish what percentage of the instantaneous output of any node i is sent to any other node j. We make the added assumption that this proportion is fixed in time and is independent of the particular value of the stress at the node. Let us denote the percentage of flux sent from i to j by g_{ij}.* Since this is a percentage, it follows that

$\sum_{j=1}^{n} g_{ij} = 1$. This is not necessary for any of the following argument, so if we do not

---

*We shall use the mathematical rather than the electrical engineering notation for matrices; the transpose operation converts one into the other.

wish to make the interpretation of percentage, we need not. From these numbers we form the matrix

$$G = \left[g_{ij}\right] \tag{6}$$

which is known as the "distribution matrix." Let the stress at each node i be denoted by s_i and form the row matrix

$$S = \left[s_i\right] \tag{7}$$

which is called the "stress matrix." Let an amount of flux f_i be introduced from an external source to each node i and form the row matrix

$$F = \left[f_i\right] \tag{8}$$

which is known as the "source matrix." It is, of course, possible that some of the g_{ij} are zero (for example, in the case of an imposed communication network) and that some of the f_i equal zero.

In the frequency domain, the total flux received by each node in the group is equal to the sum of the external sources of flux and the flux distributed from the various nodes in the network, i.e.

$$F + SG. \tag{9}$$

From Eq. 5 and the assumption that each node has the same response, and hence the same nodal impedance $\tilde{\mu}(\omega)$, it follows that

$$S = \tilde{\mu}(F + SG). \tag{10}$$

This equation may be solved for F

$$F = S(\lambda I - G) \tag{11}$$

where

$$\lambda(\omega) = \frac{1}{\tilde{\mu}(\omega)}\ . \tag{12}$$

λ is called the nodal admittance and is seen to be a measure of the passivity of the nodes. If λ is much greater than 1, a large flux may be discharged into a node without producing appreciable stress. If we assume that all the flux must ultimately leave the node, then if

λ is large, it will do so slowly. If λ is much smaller than 1, the nodes react violently to whatever flux may be imposed on them.

Frequently $(\lambda I - G)$ is denoted by the single symbol Y which is known as the group "admittance matrix." Now F and Y are given, and the problem is to determine S as a function of frequency ω or time t. Thus one wants

$$S = FY^{-1} \tag{13}$$

whenever the inverse $Y^{-1} = Z$ exists. $Z(\omega)$ is known as the "impedance matrix."

Before considering this problem, it may be well to examine these matrices in some detail. First, the diagonal terms of Z represent the stress at each of the nodes produced by a unit flux impressed thereon; these terms are called the "driving-point impedances." The nondiagonal terms represent the stress resulting at some other node than the one which is excited, and they are called the "transfer impedances."

It is clear that because of the feedback from other members of the group, the driving-point impedance may greatly exceed that of a single isolated node. That is, the reaction of a single node against an external driving force (input flux) may be greatly increased (decreased) if that node is a member of a connected or feedback group. The difference between these two numbers

$$Z_{ii} - \tilde{\mu}_i \tag{14}$$

is known as the "return impedance." It represents the difference in response between the isolated individual and the same individual in a particular group situation. This may be interpreted, psychologically, as the social pressure or group reinforcement acting on the individual in a group situation. It is easy to see that under some circumstances, it is entirely possible for the return impedance to be very large and of opposite sign to the nodal impedance. In this case the reaction of the individual in the group would be just the opposite to what his response would be if he were isolated. This type of behavior is, of course, common.

Further, one may speculate on the significance of the magnitude of the return impedance. If the return to a given node is very small, that node is in some sense isolated from the rest of the network. This may be true even when that node has a very marked influence on the rest of the network. On the other hand, if the return is very large, the node will be swamped by the flux received from other nodes, and a great stress will result from even a small added flux imposed upon it from an outside source. Such a large stress may lead to fatigue and breakdown in real human situations.

As we mentioned, for very large λ corresponding to deadened nodes, the driving point impedances Z_{ii} approach in value the node impedance $\tilde{\mu}_i$ and all transfer of flux between the nodes of the network ceases. For such values of λ the group can no longer really be considered a group. On the other hand, there will be one or more values of λ for which the driving-point and transfer impedances become infinite. For these particular values of nodal sensitivity, stress patterns may exist throughout the group even though there is no external stimulus. Since these critical values of λ are determined only by the distribution matrix of the network, they will be called "critical nodal admittances" of the network. Associated with each such critical admittance, there is a characteristic distribution of stress which is called a "natural mode" of the network. In general, of course, the nodal characteristic will not assume one of the critical values. Nonetheless, they are important, for it is possible to express the length of time required for the effect of a stimulus applied at one node in the network to be felt at any other node in terms of the decay times of the natural modes. We turn our attention to this problem.

5. Evaluation of Group Response

It is known that if the inverse Z of $(\lambda I - G)$ exists, it may be written as

$$Z_{ij}(\lambda) = \frac{A_{ji}}{\Delta} \tag{15}$$

where

$$\Delta = |\lambda I - G|$$

and A_{ji} is the cofactor of the (j,i) element of the determinant. Obviously, the inverse does not exist at those values of λ such that

$$\Delta = 0 \tag{16}$$

unless the numerator also vanishes. Let the m roots of Eq. 16 be $\lambda_1, \lambda_2, \ldots, \lambda_m$. These are the critical nodal admittances mentioned at the end of section 4.

It is known from matrix theory that there are no repeated roots of this equation, the elements of the inverse Z when it does exist may be written as

$$Z_{ij}(\lambda) = \sum_{\rho} \frac{k_{ij}^{(\rho)}}{\lambda - \lambda_{\rho}} \tag{17}$$

where

$$k_{ij}^{(\rho)} = \left. \frac{A_{ji}}{\frac{d\Delta}{d\lambda}} \right|_{\lambda=\lambda_\rho} . \tag{18}$$

If there are repeated roots, a similar, but more complicated, expression may be obtained, but we shall not go into this here. See Frazer, Duncan, and Collier: Elementary Matrices (81). This presents the answer to the problem in the frequency domain; however, we shall always be interested in the answer in the time domain, so we shall transform the answer to that form

Let $\omega_\beta^{(\rho)}$ be the solutions to the equation

$$\lambda(\omega) = \lambda_\rho \tag{19}$$

for each value of ρ. This is, of course, an equation in the Laplace transform of the given nodal characteristic $\mu(t)$. Define the functions

$$\psi_\rho(t) = \sum_{\beta=1,2,\ldots} \left. \frac{e^{\omega t}}{\frac{d\lambda}{d\omega}} \right|_{\omega=\omega_\beta^{(\rho)}} \tag{20}$$

which are called the characteristic transients corresponding to each value of ρ. Then, the stress at the i-th node resulting from a unit impulse applied to the j-th node at time t = 0 is given by

$$\mu_{ij}(t) = \sum_\rho k_{ij}^{(\rho)} \psi_\rho(t) \tag{21}$$

where the $k_{ij}^{(\rho)}$ are defined in Eq. 18. Equation 21 may be considered the basic solution to the problem in the time domain. For if the flux fed into the j-th node from all external sources is $f_j(t)$, then the total resulting stress at the i-th node will be given by the convolution integral

$$\int_{-\infty}^{\infty} f_j(\tau)\, \mu_{ij}(t-\tau)\, d\tau. \tag{22}$$

If the effects of the excitation impressed on all m nodes is superimposed, as is possible by the assumption of linearity, the total stress at the i-th node produced by all sources is

$$S_i(t) = \sum_j \left[\int_{-\infty}^{\infty} f_j(\tau)\, \mu_{ij}(t-\tau)\, d\tau \right] . \tag{23}$$

In section 7 we shall present an explicit example of how the basic solutions (Eqs. 21 and 23) may be used. Before doing this, we wish to examine more fully the coefficients $k_{ij}^{(\rho)}$ (Eq. 18).

6. A Concept of Leadership

It is clear that one may ask the following type of question: which of the nodes is most important in its contribution to the stress in the group? In a sense, the one who has the most influence is a type of leader. We shall not go into the subtleties of the concept of leadership, but simply use the term here in this special sense.

Recall that we obtained characteristic transients associated to each critical nodal admittance λ_ρ: these are the stress patterns associated with the various natural modes of the system. Now, the coefficients $k_{ij}^{(\rho)}$ express the relative amplitude of the ρ-th natural mode which appears when the system is excited at the j-th node and observed at the i-th node. They are, effectively, coupling coefficients that indicate how strongly any given node is coupled to the various natural modes of the system. Thus, the matrix

$$K_i = \begin{bmatrix} k_{1i}^{(1)} & k_{1i}^{(2)} & \cdots & k_{1i}^{(m)} \\ k_{2i}^{(1)} & k_{2i}^{(2)} & \cdots & k_{2i}^{(m)} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ k_{ni}^{(1)} & k_{ni}^{(2)} & \cdots & k_{ni}^{(m)} \end{bmatrix} \tag{24}$$

represents the amplitudes of the various modes at each of the n nodes when a unit impulse of flux is impressed on the i-th node. However, the corresponding nodal columns in any pair of these matrices, K_i and K_j, will differ only by a constant multiplying factor; these factors are the "driving-point residues" of the nodes. They may be arranged in a matrix

$$K_{jj} = \begin{bmatrix} k_{11}^{(1)} & k_{11}^{(2)} & \cdots & k_{11}^{(m)} \\ k_{22}^{(1)} & k_{22}^{(2)} & \cdots & k_{22}^{(m)} \\ \vdots & \vdots & & \vdots \\ k_{nn}^{(1)} & k_{nn}^{(2)} & \cdots & k_{nn}^{(m)} \end{bmatrix} \tag{25}$$

in which the entries in the i-th row express the amplitudes of each of the various stress modes that would appear at the i-th node if a unit impulse of flux were imposed thereon. K_{jj} and one of the K_i, say K_1, completely characterize all the k coefficients.

Obviously, the first row of K_{jj} is identical to the first row of K_1. The succeeding row will not, in general, be the same, and to make

$$k_{21}^{(1)} = k_{22}^{(1)}$$

the first column of K_1 would have to be multiplied by some factor. This factor expresses how much more strongly an impulse impressed on the second node will excite the dominant (first) mode, compared to the excitation resulting from an impulse applied to the first node. Thus, the effectiveness with which the j-th node is capable of exciting the dominant mode within the group, measured relative to the effectiveness of an arbitrary node, say the first, is given by

$$L_{jl}^{(1)} = \frac{k_{jj}^{(1)}}{k_{jl}^{(1)}} = \left.\frac{A_{jj}}{A_{lj}}\right|_{\lambda=\lambda_1} . \tag{26}$$

The particular node having the largest such ratio is the node that is capable of producing the greatest dominant response within the group of which it is a member. This node may be called the "dominant leader of the group."

For modes of higher order we define

$$L_{jl}^{(\rho)} = \frac{k_{jj}^{(\rho)}}{k_{jl}^{(\rho)}} = \left.\frac{A_{jj}}{A_{lj}}\right|_{\lambda=\lambda_\rho} \tag{27}$$

and then form

$$L_{j1} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ L_{21}^{(1)} & L_{21}^{(2)} & \cdots & L_{21}^{(m)} \\ \vdots & \vdots & & \vdots \\ L_{n1}^{(1)} & L_{n1}^{(2)} & \cdots & L_{n1}^{(m)} \end{bmatrix} \qquad (28)$$

which is called the "leadership matrix" of the group. This expresses the effectiveness of each of the various nodes in bringing about the various stress patterns throughout the group.

In the usual situation, the modes of higher order will be evanescent, so the response throughout the system is composed mostly of the dominant mode. It is, however, possible for situations to arise where these higher order mode components of the response last nearly as long as the dominant mode. The "leader" of the group will then be that node which is capable of eliciting the greatest total response, considering the combined effect of all the modes simultaneously. This is dependent on the area of the stress-time function, which is given by

$$Z_{ij}\left(\frac{1}{m_o}\right) = \sum_{\rho} \frac{k_{ij}^{(\rho)} m_o}{1 - m_o \lambda_\rho} \qquad (29)$$

where

$$m_o = \int_{-\infty}^{\infty} \mu(t)\, dt. \qquad (30)$$

The total effect exerted by any one node upon the group is thus obtained by multiplying the elements of each row of the leadership matrix Eq. 28 by the corresponding factors $(1 - m_o \lambda_1)/(1 - m_o \lambda_\rho)$. From L_{j1} and the column matrix (or vector) formed from these factors, we obtain the "leadership vector"

$$L = L_{j1} \begin{bmatrix} 1 \\ \dfrac{1 - m_o\lambda_1}{1 - m_o\lambda_2} \\ \vdots \\ \dfrac{1 - m_o\lambda_1}{1 - m_o\lambda_m} \end{bmatrix} . \tag{31}$$

We note that when the nodal amplification m_o is such that the dominant mode decays very slowly, the $(1 - m_o\lambda_1)$ will be very small and the leadership values will be determined almost entirely by the dominant mode. If the subdominant leader is different from the dominant leader, it is possible that the combined leadership may shift from one node to another as the nodal amplification m_o is varied.

7. An Example

Consider a chain group of four nodes with a source i at one end and a sink o at the other end (see Fig. A3.2). Let us assume that the output of each node is divided equally among all available nodes. Thus, all of the distribution factors are 1/2, except for the first node which discharges all of its flux into the next node. This may be imagined to be a purchasing group for a large bureaucratic organization in which the input flux is associated with the paper work generated by a purchase request. In the ideal case the order would simply pass down the chain, with each person contributing to the completion of the order. However, if an error is discovered, it is returned for correction to the person from whom it came. We assume that 50 percent of the time such an error is discovered. Furthermore, we assume that the organization is so large and the jobs are so routine that the members simply do not recognize the order if they see it a second time. Furthermore, we shall assume that each of the four nodes have the time response shown in Fig. A3.3. Each node will always delay at least one day before acting upon the flux

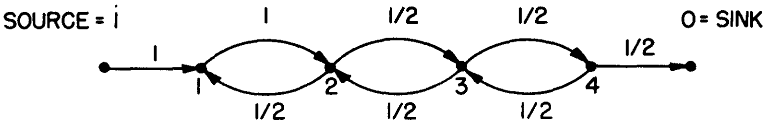


Fig. A3.2

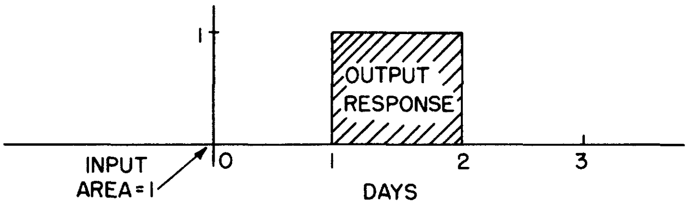


Fig. A3.3

that it receives, but action may occur with equal probability at any time during the second day. We shall assume the area under the output curve is unity ($m_o = 1$), so that, on the average, one action will be taken by the end of the second day. If $m_o < 1$, then there is a probability that the action of the order will never be taken, and if $m_o > 1$, more actions will be taken than are really needed.

By applying the theory of the preceding sections, we shall answer the following questions: (a) What is the probability that the order will be issued within any prescribed number of weeks? In particular, what is the average delay that may be expected? (b) Who is the leader of the group; that is, who should one see to get the quickest possible action? (c) How great an improvement would be effected if the errors were reduced to such a point that the order was passed on two-thirds of the time and was returned to the sender only one-third of the time?

Carrying out the computations indicated in the previous sections we find the average delay is 24 days, and the standard deviation is 20.8 days.

The first node is the "leader" of all the nodes.

If the distribution matrix is altered so that only one-third of the flux is sent back whence it came, it is found, by the same process, that the first node is still the "leader," in fact, more so. In this case the mean delay is 12.37 days with a standard deviation of 9.56 days. This is striking, for one-sixth less feedback at each node reduces the average delay to approximately one-half.

For many applications, the assumption of a rectangular distribution of output response is grossly incorrect. We have seen in section IV.3 that for at least some problems, an exponential $\alpha e^{-\alpha(t-t_o)}$, $t \geq t_o$, is a more appropriate distribution. If μ_1^* is the mean delay of an individual and one-half of the flux is returned, then the mean of the group, μ_1, is given by

$$\mu_1 = 176.8\ \mu_1^*$$

and in case one-third of the flux is returned

$$\mu_1 = 33.1\ \mu_1^*.$$

8. The Value of the Theory

We have included this appendix, in spite of the fact that we have done no experimental work to which it is directly relevant, to illustrate an important type of approach to the development of theories of group behavior. The class of groups to which the theory is applicable is very restricted, but this fact could not surely be known until the attempt to interpret the theory made its implications for groups clear.

The technique, in outline, has been as follows: (a) to formulate clearly the basic assumptions of electrical network theory, (b) to interpret these assumptions for human groups, (c) to give, on the basis of these assumptions, a mathematical definition of the nodal properties and, in the form of network equations, of the group, (d) to present the known mathematical solutions to the network equations and reinterpret this formalism in terms of groups. In the present instance this process of theory construction has involved us in nothing mathematically new, since the mathematics employed is that which is well known from its use in electrical network theory. Beyond the foregoing bare essentials of the method, the theory was further developed (in a way possible for its electrical case but of little interest there) to give a treatment of "leadership" phenomena. Finally, a hypothetical example was presented to illustrate how the theory could be applied to an actual example.

Aside from the apparent fact that the theory is applicable to a certain narrowly defined class of groups, it has merit in that the particular points in which it is deficient can be seen clearly. The task of constructing theories applicable to wider classes of groups can, therefore, be approached more effectively than it could have been without this knowledge (see section IV.9). Insofar as the present theory is found to be true when applied to a properly selected class of groups, a more general theory must reduce, as a special case, to something equivalent to the present one when it is applied to the same class of groups. The linear monochromatic group theory is, therefore, a potentially important reference point for new theoretical developments.

APPENDIX 4 – ABSTRACT NETWORK THEORY

The abstract notion of a "network" is one of several possible abstractions of the concept of a communication network imposed on a task-oriented group of people. In essence, we preserve the routes over which the communication may occur, but we ignore completely the dynamics of the situation by divesting the nodes of their dynamic character, the transfer function. This leaves us the geometrical (or, more precisely, the topological) aspects of a dynamic system. It follows that in such a theory we can expect to prove only topological results. This, as far as a theory of task-oriented groups is concerned, is not desirable, for the problems of the group are dynamic within the boundary conditions of the topology of the given network. Yet, the topology may remain important even in the solution of dynamic problems, for it is possible that certain topological theorems may appreciably reduce the difficulty of the dynamic problem, and it is possible that people sometimes react more to the topology than to the dynamics.

An important contribution in this area would be the development of the mathematics of networks in which the transfer functions of the nodes are retained with sufficient assumptions to yield results but, at the same time, not with such stringent assumptions that the resulting model has no relation to actual situations. This is a goal of a theoretical study of task-oriented groups.

We shall present in this appendix a summary of some results on abstract networks which have been developed primarily from a topological viewpoint. We shall first give a precise definition of the area of study and then introduce several restrictive definitions which we have found useful. The selection of theorems stated here has been based on the following considerations: that they have a fairly clear intuitive meaning and that they are typical of the results that such a study may be expected to yield. For other results in this area the reader is referred to the articles on graph theory in the bibliography under mathematics. No attempt will be made in this presentation to furnish support for the assertions that are made; the proofs are published in full in another place (93).

As we mentioned, an abstract network is a topological entity. On the other hand, one may make the concise definition that a network is a relation over a finite set; that is, a network consists of a set of n elements, called the nodes, and a binary relation R defined as existing between certain ordered pairs of nodes. If R exists between a and b (aRb in the usual notation), we say a link exists from a to b and denote it by [ab]. Since the theory of relations is a part of algebra, we see that a network is both a topological structure and an algebraic structure. If N is a network and N' is another network having only nodes and links that are in N, but not necessarily all those present, then N' is called a subnetwork of N. If the nodes of a subnetwork of N are the same as the nodes of N, then the subnetwork is called "complete."

In a communication network for a group of people, the people are nodes and the directed channels over which the people may communicate are links. In an electrical network, the nodes are components such as resistors and capacitors, and the links are wires over which current may flow. In general, one considers either of these situations to be symmetric in the sense that a link from a to b implies one from b to a; that if a can talk to b then b can talk to a; or in electrical theory, that the current may flow in either direction through the wire. This is not strictly true for either example: In the case of people, man A with a transmitter may communicate to another man B with a receiver, while B cannot communicate to A if B has no transmitter or if A has no receiver. In electrical theory it is considered appropriate to treat the plate of a vacuum tube as a component distinct from the cathode. As the current in normal operation always passes from the plate to the cathode, this current may be represented by a directed link.

In any communication network it is important to consider sequences of links over which a message may flow sequentially. That is, a set of links of the form

$$[ac_1], [c_1c_2], \ldots, [c_{q-1}c_q], [c_qb].$$

Whenever such a set of links exists, we say that a "chain" exists from a to b. A chain from a to a which does not cross itself, i.e. the c_i are all different, is called a "circuit."

Of the class of all possible communicating groups of people, the ones which may be considered to be significant are those in which each member of the group may, in some way, communicate to each other member. This selection is arbitrary, but it is difficult to consider two people to be members of the same group when one is unable to influence the other by communication. Thus, we are led to isolate those abstract networks such that from any node to any other node there exists at least one chain; such networks are called "connected." Other terms have been introduced by various authors for this concept: feedback network, closed network, and re-entrent network. Each of these terms, including "connected," is descriptive; however, here we shall use the mathematical term "connected." If a network is not connected, it is called "disconnected."

These definitions lay out a general area for study. In the following pages additional definitions will be made which will allow the statement of theorems within this realm; however, the choice of definitions is greatly dependent on whether we shall view the network as topological or as algebraic. Some work has been published on the algebraic approach (84), (89), (90), (91), (103). Much of this has had to be new work, for the theory of relations has in large part been built up on the assumption of transitivity, just as our theory of networks is being built up on the assumption of connectedness. A relation R is transitive if aRb and bRc imply aRc. This assumption we cannot make, for in a communication situation the existence of a link from a to b and one from b to c does not often imply the existence of one from a to c. In fact, it is trivial to show that a connected transitive network has all possible links present. Thus, an algebraic approach must be new, and so far, an attempt to apply algebraic techniques to this class of objects has led to problems of great complexity. This has led us to the topological approach and, in particular, to a study of connected networks.

We must now introduce definitions which will allow some analysis of networks. Since connectedness is our primary assumption, it may do well to consider it further. First of all, each of the networks in Fig. A4.1 have the same number of nodes, and each is connected, but certainly no one will deny that the circuit in Fig. A4.1b is "less" connected than the totally connected network, Fig. A4.1a. In part, this difference may be attributed simply to the "density" of links present, but a more important difference is that connectedness may be destroyed more readily in one than in the other. We define a measure of this difference: A network is of degree 0 if it is disconnected; it is of degree $k \geqslant 0$ if there exists at least one set of k links whose removal from the network will result in a disconnected network, whereas the removal of any fewer links will not disconnect it. If a network has m nodes and degree k, $k \leqslant m-1$. The network in Fig. A4.1a is of degree 4, and the one in Fig. A4.1b is of degree 1. A network of higher degree than another is the "more" connected.

This, however, is not the whole story, for the networks in Fig. A4.2 display a difference that we should like to isolate. The network in Fig. A4.2a is of degree 1 and that in Fig. A4.2b is of degree 2; yet the former has a greater number of links than the latter, and in fact the part of the network A4.2a including only the nodes 2, 3, 4, and 5 is of degree 3. The following concept is essentially a condition of "uniformity" or "evenness" of distribution of connectedness: a network is called k-minimal if the removal of any link from it results in a network of degree k-1. In Fig. A4.2a we see that the removal

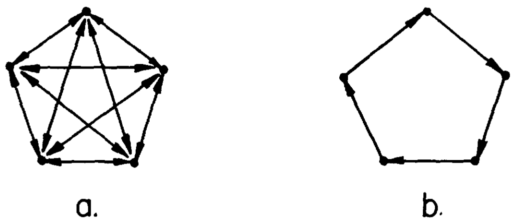
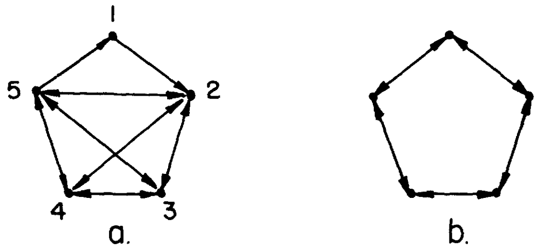

a. b.

Fig. A4.1



a. b.

Fig. A4.2

of the link [23] results in a network of degree 1, and the removal of [12] results in a network of degree 0; hence, that network cannot be k-minimal for any k. The circle, Fig. A4.2b, is easily seen to be 2-minimal.

One may show that if a network is k-minimal and $k \geqslant 2$, then the network is of degree k and any connected subnetwork is of degree $\leqslant k$. This is certainly a property that one would demand of any concept that purports to imply a "uniform distribution of connectedness." Note, however, that we have not shown this for the case $k = 1$; this cannot be done, for any disconnected network, no matter how uneven the distribution of connectedness of subnetworks, is 1-minimal. On the other hand, any connected network which is 1-minimal can be shown to have the property that every connected subnetwork is of degree 1. A member of this class of networks, i.e. connected and 1-minimal, is called minimal. A network is minimal if, and only if, it is connected and the removal of any link results in a disconnected network.

A network N is the sum of t subnetworks N_i, where $i = 1, 2, \ldots, t$, if each link of N is contained in exactly one of the N_i. That is, each link is represented in one and only one of the subnetworks. This is written

$$N = \sum_{i=1}^{t} N_i.$$

We may then state a result which is a "decomposition" theorem for any network: For each network N there exists a unique number k, its degree, and at least one set of k+1 complete 1-minimal subnetworks, N_i, such that

i. $N = \sum_{i=1}^{k+1} N_i$,

ii. N_{k+1} is disconnected,

iii. if $k \geqslant 1$, N_1 is minimal,

and

iv. every connected subnetwork of the N_i, $1 \leqslant i \leqslant k$, is minimal.

To obtain this theorem, find one of the complete subnetworks of degree k of N having the fewest possible links. Let this be N', and define $N_{k+1} = N-N'$. N_{k+1} measures how much more uneven the distribution of links in N is than in the k-minimal network N'. It is then shown that the k-minimal network N' may be decomposed into a sum of k 1-minimal networks. These 1-minimal networks are not necessarily minimal; they may be disconnected. However, there is an element of evenness in the distribution of connectedness in these 1-minimal subnetworks, for any connected subnetwork of N_i is minimal.

This result suggests a particular class of networks which is in some respects experimentally significant. Suppose

$$N = \sum_{i=1}^{k} N_i$$

where the N_i are each a circuit comprising all the nodes. Then N has the property that it is k-minimal and that each node terminates exactly k links and originates exactly k links. Thus, in a sense, each of the nodes is in exactly the same relation to the rest of the network as each of the other nodes. Experimentally, such a property is valuable since it implies that data for the same general category of node will be obtained m times as rapidly as when each node has a different relation to the remainder of the network. It may also be shown that if a network has the property that each node originates and terminates exactly k links, and if the network is of degree k, then it is k-minimal. The condition that the network be of degree k is necessary in the last statement, for Fig. A4.3 presents an example of a network in which each node originates and terminates exactly 2 links. But this network is of degree 1 and is not 1-minimal. Furthermore, it is not true that if a network is k-minimal and each node originates and terminates exactly k links, the network is the sum of k circuits. See Fig. A4.4, for example. In addition to

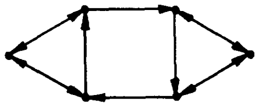


Fig. A4.3

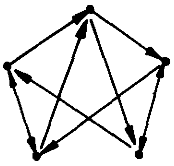


Fig. A4.4

the network of Fig. A4.4, there are other connected networks on five nodes with the property that each node originates and terminates k links. These are shown in Fig. A4.5.

The decomposition theorem indicates that minimal networks are the basic building blocks of connected networks, for the 1-minimal networks N_i, $1 \leqslant i \leqslant k$, consist of isolated nodes, chains which are not a portion of a circuit, and connected pieces which are minimal. Furthermore, a repeated application of the theorem to the connected pieces of N_{k+1} shows that it too may finally be reduced to isolated nodes, isolated chains, and minimal networks. As yet, we really do not know if it will suffice to study minimal networks to obtain information about more general networks, for we do not know enough about this decomposition. Examples may be produced to show that one network N may have two decompositions

$$N = \sum_{i=1}^{k+1} N_i = \sum_{i=1}^{k+1} N_i'$$

which satisfy the above conditions. Furthermore, the decompositions are such that no relabeling of the nodes and the indices of N_i allows one to state

$$N_i = N_i', \quad i = 1, 2, \ldots, k+1.$$

An important unsolved problem of abstract network theory is to state the relation between two such decompositions of a given network. Even though we do not have complete information about this decomposition, it is true, in a sense, that minimal networks are the basic connected networks. Thus, we are led to inquire as to the properties of minimal networks.

In order to discuss minimal networks, it is necessary to be familiar with the concept of a tree in graph theory. Any network having the property that the existence of a link $\left[ab\right]$ implies the existence of the link $[ba]$ is known as a "graph." More strictly,

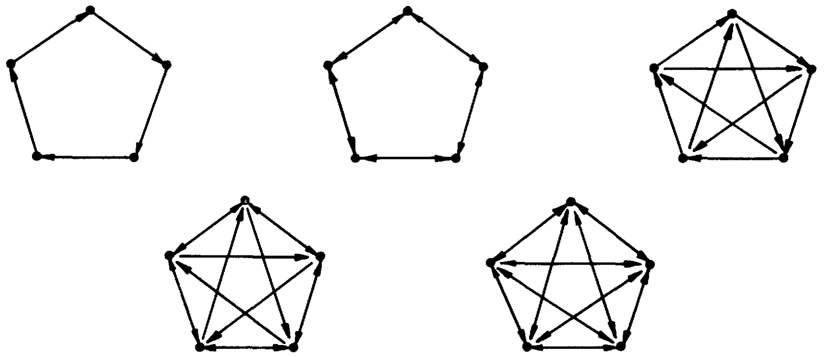


Fig. A4.5

one should say that such a network is isomorphic to a graph; however, we need not make this distinction. In the theory of relations such a network would be called a symmetric relation. The pair of links [ab] and [ba] are known as an arc, ab. A circuit of arcs is a set of arcs of the form

$$ac_1, c_1c_2, \ldots, c_{q-1}c_q, c_qa.$$

A connected network which is a graph and has no circuits of arcs is known as a "tree." This is a concept which is of importance in graph theory, and it has been studied quite thoroughly. Two examples of networks which are trees are given in Fig. A4.6. Without

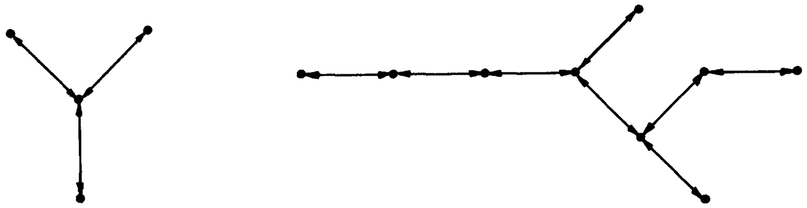


Fig. A4.6

great difficulty one can show that a network which is a tree is minimal; however, the converse, that a minimal network is a tree, is certainly not true since a circuit of links is minimal. Nonetheless, it is suggested that there is a strong relation between the two concepts.

To discuss more fully the nature of minimal networks, we shall need two definitions. A network is a "compound circuit" of order 1 if it is a circuit. Assuming a compound circuit of order s-1 has been defined, one of order s is formed by replacing some node c of a compound circuit C_{s-1} of order s-1 by a circuit C in such a fashion that any link of the form [ac] in C_{s-1} is replaced by one of the form [ac'], where c' is a node of C, and any link of the form [ca] is replaced by one of the form [c"a], where c" is a node of C. For example, the network in Fig. A4.7a is a compound circuit which is formed as shown in Fig. A4.7b. A compound circuit having m nodes and p links and of order s is connected, and s = p-m+1.

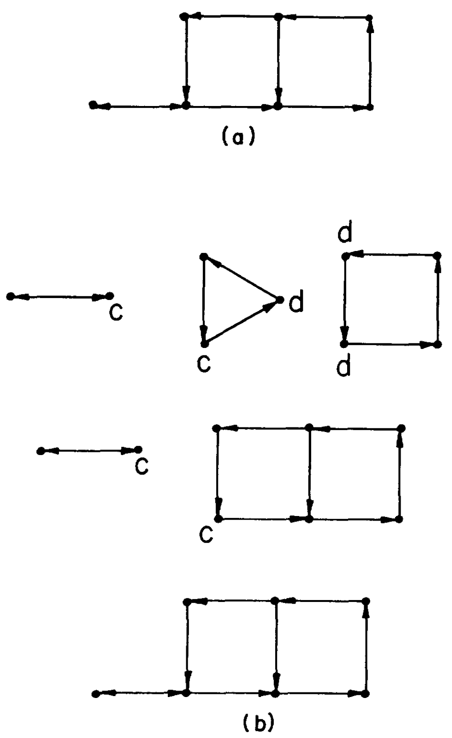


Fig. A4.7

A network is called "reducible" if it consists of two subnetworks N_1 and N_2, having no nodes in common (i.e. disjoint subnetworks), such that there is exactly one link from N_1 to N_2 and exactly one from N_2 to N_1, and the subnetworks are either single nodes or connected. A reducible network is connected. If a network is not reducible, it is called "irreducible." It can be shown that a network which is a graph is reducible if, and only if, it is of degree 1. More important, if N is a minimal network, which is not a tree, then N consists of $t \geqslant 1$ irreducible arc-free disjoint compound circuits and $y \geqslant 0$ nodes which are not contained in these compound circuits. If these compound circuits are treated as nodes but the remainder of the network is unchanged, then the resulting network is a tree. In the case t = 1 and y = 0, this tree is the trivial case of a single node with no links. This result means, roughly, that if you look at a minimal network appropriately, it "looks like" a tree.

From this one can deduce several results. First, let us call a node "simple" if it has only one link entering and only one leaving. Then a minimal network is a compound circuit having at least two simple nodes. (Fig. A4.7 is an example of a compound circuit which is not minimal.) This is a useful result in the proof of other results, not to be stated here, for it allows the development of proofs based on induction. Second, if N is minimal, but not a tree, and if t and y mean what they did above, and if N has p links and m nodes, then

$$p \leqslant \frac{3m+t+y-4}{2} < 2(m-1).$$

If N is a tree, it is well known that

$$p = 2(m-1).$$

A final set of results may be given which, once again, relate very closely the topological to the algebraic view of the situation. It is clear, since a network is a relation, that we may identify a network of m nodes with a matrix N of order m having the entry $N_{ab} = 1$ if $[ab]$ exists and 0, otherwise. There is defined for any matrix its rank r, the maximum number of linearly independent rows. Call the rank of a matrix corresponding to a network the "rank" of that network. Then one may show that if N is a connected network of p links, m nodes, and rank r, then $p+r \geq 2m$. We may term any connected network such that $p+r = 2m$, "rank minimal." It can be shown that a rank minimal network N is minimal, and that it consists of exactly one irreducible rank minimal compound circuit N' and nodes and links not in N', such that when N' is treated as a node, the resulting network is a tree with all arcs meeting at a common node, i. e. a star.

Thus, for example, one may verify that the network in Fig. A4.8a is rank minimal. Its tree-structure is given in Fig. A4.8b.

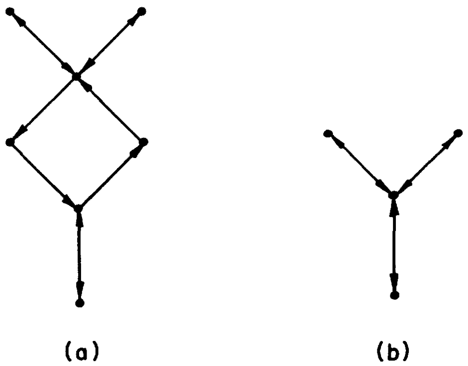

(a) (b)

Fig. A4.8

These results are not exhaustive, but they do include some of the more interesting theorems that have been proved. It is hoped that they will give the reader some feeling for the type of study involved and for the fact that the results are primarily descriptive in nature. The study does not pretend to say anything about any dynamic problem which may involve a network, but it does give some insight into the logical structure of a variable which is clearly important in some experimental studies. It is imperative in experimental work to select the values of such a variable quite judiciously, for the number of networks having $m>3$ nodes is tremendous. Such a study as this isolates classes of networks which have, in some sense, similar characteristics and describes, to some extent, the members of such classes, thus making the task of listing explicitly all members of such a class less tedious than it otherwise would be.

Topological theory may ultimately play a role closer to the study of dynamic problems, for it is quite possible that the topology of the network may impose upper and lower bounds on the possible performance of a group working within the limitations of the network. Thus, for example, it may be possible in certain synthesis problems to say that all networks excepting those in a certain specified class will fail to solve the given problem. If it is then necessary to perform calculations on each network to see which will fulfill the synthesis conditions, then the topological theorems will have reduced the number of networks it is necessary to examine.

BIBLIOGRAPHY

We have divided this bibliography into six sections: Cybernetics; Small Group Psychology and Learning Theory; Information Theory, General Concepts; Information Theory, Applications; Mathematics; and Servomechanisms and Electrical Concepts. Within each section the publications are arranged alphabetically by the author, but the numbering of the bibliography is continuous. Obviously, each of these section headings must be qualified to some extent as to publications relevant to the study of groups. Even so, some arbitrary selection has had to be made. Much of the mathematics developed by the Committee on Mathematical Biophysics for the study of neural networks could have been included as related to the group network problem, even though it was not directly applicable. Where does one draw the line in the vast literature of learning theory?

With the exception of the mathematics section, our policy has been to include publications we have found particularly useful and suggestive. The great bulk of the papers listed under mathematics pertain to graph theory and abstract network theory. These papers are important if the topology of the communication network is important to group problems. Since such a bibliography is next to impossible for the social scientists to find, and requires some time and effort for a mathematician to prepare, it was felt that it was worthwhile to present it, even though it is not in balance with the other sections.

A more complete bibliography of the area of cybernetics has been prepared by F. Dermot Barrett and Professor Herbert A. Shepard of M.I.T. and is available in mimeographed form.

Cybernetics

1. Eric Barrett: Cybernetics as Applied to a Study of Normal and Abnormal Adjustment Mechanisms, J. Psychol. 30, 11-31, 1950
2. E. Barrett, G. Post: Introduction to Some Principles of Applied Cybernetics, J. Psychol. 30, 3-10, 1950
3. Conference on Cybernetics: Transactions of the Sixth Conference, Josiah Macy, Jr. Foundation, New York, 1949
4. Conference on Cybernetics: Transactions of the Seventh Conference, Josiah Macy, Jr. Foundation, New York, 1950
5. Conference on Cybernetics: Transactions of the Eighth Conference, Josiah Macy, Jr. Foundation, New York, 1951
6. W. S. McCulloch: A Recapitulation of the Theory, with a Forecast of Several Extensions, in "Teleological Mechanisms," Ann. New York Acad. Sci. 50, 247-258, 1948

7. A. Rapoport, A. Shimbel: Mathematical Biophysics, Cybernetics and General Semantics, Etc., A Review of General Semantics 6, 145-159, 1949

8. John D. Trimmer: Instrumentation and Cybernetics, Sci. Monthly 69, 328-331, 1949

9. J. D. Trimmer: Response of Physical Systems, John Wiley, New York, 1950

10. N. Wiener: Cybernetics, Technology Press, Cambridge, Mass., 1948

11. N. Wiener: Time, Communication, and the Nervous System, in "Teleological Mechanisms," Ann. New York Acad. Sci. 50, 197-219, 1948

Small Group Psychology and Learning Theory

12. K. W. Back: Influence Through Social Communication, J. Abnorm. Psychol. 46, 9-23, 1951

13. R. F. Bales: Interaction Process Analysis, Addison-Wesley Press, Cambridge, Mass., 1950

14. R. F. Bales, Fred L. Strodtbeck: Phases in Group Problem-Solving, J. Abnorm. Psychol. 46, 485-495, 1951

15. A. Bavelas: Communication Patterns in Task-Oriented Groups, J. Acoust. Soc. Am. 22, 725-730, 1950

16. R. R. Bush, F. Mosteller: A Mathematical Model for Simple Learning, Psychol. Rev. 58, 313-323, 1951

17. R. R. Bush, F. Mosteller: A Model for Stimulus Generalization and Discrimination, Psychol. Rev. 58, 413-423, 1951

18. L. S. Christie: The Measurement of Discriminative Behavior, Psychol. Rev. 59, 443-452, 1952

19. M. Deutsch: An Experimental Study of the Effects of Co-operation and Competition Upon Group Processes, Human Relations 2, 199-232, 1949

20. Edward M. Dooley: Experimental Methodology for Communication Pattern Analysis, Master's Thesis, Dept. of Business and Engineering Administration, M.I.T. 1949

21. W. K. Estes: Toward a Statistical Theory of Learning, Psychol. Rev. 57, 94-107, 1950

22. L. Festinger: Informal Social Communication, Psychol. Rev. 57, 271-282, 1950

23. L. Festinger et al.: Theory and Experiments in Social Communication, Research Center for Group Dynamics, Edwards Bros., Ann Arbor, Michigan, 1950

24. L. Festinger, J. Thibaut: Interpersonal Communication in Small Groups, J. Abnorm. Psychol. 46, 92-99, 1951

25. L. Festinger, S. Schacter, K. Back: Social Pressures in Informal Groups; A Study of Human Factors in Housing, Harper, New York, 1950

26. F. C. Frick, G. A. Miller: A Statistical Description of Operant Conditioning, Am. J. Psychol. 64, 20-36, 1951

27. A. Gibb: The Research Background of an Interactional Theory of Leadership, Austrian J. Psychol. 2, 19-42, 1950

28. H. Guetzkow (Ed.): Groups, Leadership and Men; Research in Human Relations, Carnegie Press, Pittsburg, 1951

29. George A. Heise, George A. Miller: Problem Solving by Small Groups Using Various Communication Nets, J. Abnorm. Psychol. 46, 327-335, 1951

30. J. K. Hemphill, C. M. Westie: The Measurement of Group Dimensions, J. Psychol. 29, 325-342, 1950

31. J. K. Hemphill: Relations Between the Size of the Group and the Behavior of "Superior Leaders," J. Social Psychol. 32, 11-22, 1950

32. D. B. Hertz, R. L. Livingston: Contemporary Organizational Theory: A Review of Current Concepts and Methods, Human Relations 3, 373-394, 1950

33. G. C. Homans: The Human Group, Harcourt Brace, New York, 1950

34. Harold J. Leavitt: Some Effects of Certain Communication Patterns on Group Performance, J. Abnorm. Psychol. 46, 38-50, 1951

35. K. Lewin: Field Theory in Social Science: Selected Theoretical Papers (Ed. Dorwin Cartwright), Harper, New York, 1951

36. R. Lippitt: Socio-psychological Research and Group Work, in C. E. Hendrey: A Decade of Group Work, Association Press, New York, 1948, pp. 166-177

37. G. A. Miller, F. C. Frick: Statistical Behavioristics and Sequences of Responses, Psychol. Rev. 56, 311-324, 1949

38. G. A. Miller:/ Language and Communication, McGraw-Hill, New York, 1951

39. J. G. Miller (Ed.): Experiments in Social Process: A Symposium on Social Psychology, McGraw-Hill, New York, 1950

40. A. Mintz: Non-adaptive Group Behavior, J. Abnorm. Psychol. 46, 150-159, 1951

41. Oskar Morgenstern: Prolegomena to a Theory of Organization, Project RAND, RM-734, Santa Monica, Calif., 1951

42. R. T. Morris, Melvin Seeman: The Problem of Leadership: An Interdisciplinary Approach, Am. J. Sociol. 56, 149-155, 1950

43. C. G. Mueller: Theoretical Relationships Among Some Measures of Conditioning, Proc. Nat. Acad. Sci. 36, 123-130, 1950

44. H. A. Simon: A Study of Decision-Making Processes in Administrative Organization, MacMillan, New York, 1947

45. Sidney L. Smith: Communication Pattern and the Adaptability of Task-Oriented Groups: An Experimental Study, Group Networks Laboratory, Research Laboratory of Electronics, M.I.T. (Mimeographed)

46. Sidney L. Smith: Survey of Some Experiments in Group Communication, Group Networks Laboratory, Research Laboratory of Electronics, M.I.T. (Mimeographed)

47. B. Steinzor: The Spatial Factor in Face-to-Face Discussion Groups, J. Abnorm. Psychol. 45, 552-555, 1950

48. R. M. Stogdill: Leadership, Membership and Organization, Psychol. Bull. 47, 1-14, 1950

49. J. Thibaut: An Experimental Study of the Cohesiveness of Underprivileged Groups, Human Relations 3, 251-278, 1950

Information Theory: General Concepts

50. Y. Bar-Hillel: Semantic Information and Its Measures, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T. April 1952

51. R. M. Fano: The Transmission of Information, Technical Report No. 65, Research Laboratory of Electronics, M.I.T. March 1949

52. R. M. Fano: The Transmission of Information, Part II, Technical Report No. 149, Research Laboratory of Electronics, M.I.T. Feb. 1950

53. D. Gabor: Theory of Communication, J. Inst. Elec. Eng. 93, 439, Nov. 1946

54. R. V. L. Hartley: Transmission of Information, Bell System Tech. J. 7, 535-7, 1928

55. A. E. Laemmel: General Theory of Communication, Polytechnic Institute of Brooklyn, Microwave Research Institute (introductory parts)

56. M. Leifer, W. F. Schreiber: Communication Theory, Advances in Electronics, Vol. III, Academic Press, New York, 1951, pp. 305-343

57. S. O. Rice: Communications in the Presence of Noise: Probability of Error for Two Encoding Schemes, Bell System Tech. J. 29, 60-93, 1950

58. T. Roddam: Communication Theory, Old and New, Science News, No. 12, 105-121, Penquin Books, Harmondsworth, Middlesex, England, 1949

59. C. E. Shannon: A Mathematical Theory of Communication, Bell System Tech. J. 27, 380-423, 1948; 27, 623-656, 1948

60. C. E. Shannon: Communication in the Presence of Noise, Proc. Inst. Radio Engr. 37, 10-21, 1949

61. C. E. Shannon: Communication Theory of Secrecy Systems, Bell System Tech. J. 28, 656-715, 1949

62. C. E. Shannon, W. Weaver: The Mathematical Theory of Communication, University of Illinois Press, Urbana, 1949

63. Symposium on Information Theory: Report of Proceedings, Ministry of Supply, London, Sept. 1950

64. W. G. Tuller: Theoretical Limitations on the Rate of Transmission of Information, Technical Report No. 114, Research Laboratory of Electronics, M.I.T. April 1949

65. N. Wiener: A New Concept of Communication Engineering, Electronics 22, 74-76, 1949

Information Theory: Applications

66. L. Brillouin: Life, Thermodynamics, and Cybernetics, Am. Scientist 37, 554-568, 1949

67. L. Brillouin: Thermodynamics and Information Theory, Am. Scientist 38, 594-599, 1950

68. A. G. Clavier: Evaluation of Transmission Efficiency According to Hartley's Expression of Information Content, Elec. Commun. 25, 414-20, 1948

69. F. Cohen: Some Analytical and Practical Aspects of Wiener's Theory of Prediction, Technical Report No. 69, Research Laboratory of Electronics, M.I.T. June 1948

70. R. M. Fano: The Information Theory Point of View in Speech Communication, J. Acoust. Soc. Am. 22, 691-697, 1950

71. D. Gabor: Communication Theory and Physics, Phil. Mag. 41, 1161-87, 1950

72. G. A. Miller: Language Engineering, J. Acoust. Soc. Am. 22, 720-725, 1950

73. C. E. Shannon: Memory Requirements in a Telephone Exchange, Bell System Tech. J. 29, 343-349, 1950

74. C. E. Shannon: Prediction and Entropy of Printed English, Bell System Tech. J. 30, 50-64, 1951

75. O. H. Straus: The Relation of Phonetics and Linguistics to Communication Theory, J. Acoust. Soc. Am. 22, 709, 1950

76. N. Wiener: Speech, Language, and Learning, J. Acoust. Soc. Am. 22, 696-697, 1950

Mathematics

77. R. P. Abelson: Mathematical Contributions to Group Psychology, Master's Thesis, Dept. of Mathematics, M.I.T. Feb. 1950

78. A. Bavelas: A Mathematical Model For Group Structures, Applied Anthropology, Summer, 1948

79. G. Birkhoff: Lattice Theory, American Mathematical Society, New York, 1948, Chap. 13

80. H. S. M. Coxeter: Self-Dual Configurations and Regular Graphs, Bull. Am. Math. Soc. 56, 413-455, 1950 (This is a survey article with a very large bibliography which extends somewhat beyond our bounds)

81. J. Feller: An Introduction to Probability Theory and Its Applications, John Wiley, New York, 1950

82. R. A. Frasier, W. J. Duncan, A. R. Collar: Elementary Matrices, Macmillan, New York, 1950

83. W. H. Ingram, C. N. Cramlet: On the Foundations of Electrical Network Theory, J. Math. Phys. 23, 134-155, 1944

84. Leo Katz: An Application of Matrix Algebra to the Study of Human Relations Within Organizations, Institute of Statistics, University of North Carolina, Mimeograph Series, 1950

85. M. G. Kendall: The Advanced Theory of Statistics, Charles Griffin, London, 1948

86. M. G. Kendall, B. Babington Smith: Tables of Random Sampling Numbers, Tracts for Computers, No. 24, Cambridge University Press, 1939

87. D. König: Theorie der Endlichen und Unendlichen Graphen, Chelsea Publishing Company, New York, 1949

88. S. Lefschetz: Introduction to Topology, Princeton University Press, Princeton, 1949, Chap. II

89. R. D. Luce, A. D. Perry: A Method of Matrix Analysis of Group Structure, Psychometrika 14, 95-116, 1949

90. R. D. Luce: Connectivity and Generalized Cliques in Sociometric Group Structures Psychometrika 15, 169-190, 1950

91. R. D. Luce: A Note on Boolean Matrix Theory, Proc. Am. Math. Soc. 3, 382-388, 1952

92. R. D. Luce: The Theory of Networks, Group Networks Laboratory, Research Laboratory of Electronics, M.I.T. 1951 (mimeographed)

93. R. D. Luce: Two Decomposition Theorems for a Class of Finite Oriented Graphs, Am. J. Math. 74, 701-722, 1952

94. Saunders MacLane: Some Unique Separation Theorems for Graphs, Am. J. Math. 57, 805, 1935

95. Samuel J. Mason: On The Logic of Feedback, Technical Report No. 153, Research Laboratory of Electronics, M.I.T. 1953

96. W. S. McCulloch, W. Pitts: A Logical Calculus of the Ideas Immanent in Nervous Activity, Bull. Math. Biophys. 5, 115-133, 1943

97. R. Otter: Number of Trees, Ann. Math. 49, 583-599, 1948

98. G. Polya: An Zahl Bestimmungen f. Gruppen, Graphen u. Chemische Verbindungen, Acta Math. 58, 145, 1937

99. John Riordan, C. E. Shannon: The Number of N-terminal Series-Parallel Networks, J. Math. Phys. 21, 83-93, 1942

100. I. C. Ross, Frank Harary: On the Determination of Redundancies in Sociometric Chains, Research Center for Group Dynamics, University of Michigan (mimeographed)

101. S. T. Saunders: A Linear Transformation Whose Variables and Coefficients are Sets of Points, Bull. Am. Math. Soc. 48, 440, 1942

102. James K. Senior: Partitions and Their Representative Graphs, Am. J. Math. 73, 663, 1951

103. Alfonso Shimbel: Applications of Matrix Algebra to Communication Nets, Bull. Math. Biophys. 13, 165, 1951

104. W. T. Tutte: A Class of Self-Dual Maps: Can. J. Math. 2, 184, 1950

105. J. Von Neumann, O. Morgenstern: The Theory of Games and Economic Behavior, Princeton University Press, Princeton, 1947

106. J. H. M. Wedderburn: Boolean Linear Associative Algebra, Ann. Math. 35, 185-194, 1934

107. H. Whitney: Non-Separable and Planar Graphs, Trans. Am. Math. Soc. 34, 339-362, 1932

108. H. Whitney: Congruent Graphs and the Connectivity of Graphs, Am. J. Math. 54, 150-168, 1932

109. H. Whitney: Planar Graphs, Fundamenta Mathematicae 21, 73-84, 1933

110. H. Whitney: The Coloring of Graphs, Ann. Math. 33, 688-718, 1932

111. H. Whitney: On the Classification of Graphs, Am. J. Math. 55, 236-244, 1933

112. H. Whitney: 2-Isomorphic Graphs, Am. J. Math. 55, 245-254, 1933

113. H. Whitney: On the Abstract Properties of Linear Dependence, Am. J. Math. 57, 509, 1935

Servomechanisms and Electrical Concepts

114. W. R. Ahrendt, J. F. Taplin: Automatic Feedback Control, McGraw-Hill, New York, 1951

115. H. W. Bode: Network Theory and Feedback Amplifier Design, Van Nostrand, New York, 1945

116. G. S. Brown, D. P. Campbell: Principles of Servomechanism: Dynamics and Synthesis of Closed Loop Control Systems, John Wiley, New York, 1948

117. E. A. Guillemin: Mathematics of Circuit Analysis, John Wiley, New York, 1949

118. William H. Huggins: Linear Systems and Group Dynamics, Group Networks Laboratory, Research Laboratory of Electronics M.I.T. 1949 (mimeographed)

119. William H. Huggins: Memo on the Experimental Determination of Transfer Functions for Human Operators and Machines, R. F. Components Laboratory, Air Force Cambridge Research Center, October 1949 (mimeographed)

120. H. F. James, N. B. Nichols, R. S. Philips: Theory of Servomechanisms, McGraw-Hill, New York, 1947

121. H. Lamar, R. Nesnick, L. E. Matson: Servomechanism Fundamentals, McGraw-Hill, New York, 1945

122. L. A. MacColl: Servomechanisms, Van Nostrand, New York, 1945

123. H. A. Simon: An Exploration into the Use of Servomechanism Theory in the Study of Production Control, Cowles Commission Discussion Paper: Economics No. 288 (mimeographed copy for private circulation)

Note added in proof:

For a very much more complete bibliography of much the same area as covered here, the reader is referred to F. L. Stumpers: A Bibliography of Information Theory, Research Laboratory of Electronics, M.I.T. Feb. 1953.