
106
058
THS
Ph. D.

MODELS FOR THE CLASSIFICATION OF PROBLEMS AND THE PREDICTION OF GROUP PROBLEM-SOLVING FROM INDIVIDUAL RESULTS

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY

James Henry Davis

1961

This is to certify that the

thesis entitled

Models for the Classification of Problems

and the Prediction of Group Problem-

Solving from Individual Results

presented by

James H. Davis

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Psychology

Frank Restle
Major professor

Date December 17, 1960

O-169

MODELS FOR THE CLASSIFICATION OF PROBLEMS AND THE PREDICTION OF GROUP PROBLEM-SOLVING FROM INDIVIDUAL RESULTS

By

James H. Davis

AN ABSTRACT OF A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

1961

G 18242
11/8/61

ABSTRACT

MODELS FOR THE CLASSIFICATION OF PROBLEMS AND THE PREDICTION OF GROUP PROBLEM-SOLVING FROM INDIVIDUAL RESULTS

by James H. Davis

Eureka problems (word puzzles having unique solutions), were worked on by individuals, and by ad hoc groups in which unrestricted, face-to-face interaction was permitted between members who had no tradition of working together. The first objective of the study was to introduce an ordering of problems and a method measuring that order. Second, the investigation sought to predict group problem solving performance from a knowledge of the problem solving behavior of persons working as individuals. Major emphasis was placed upon the development of a model that dealt concurrently with group product and emergent group structure. Third, the frequently-noted superiority of problem solving groups over individuals working on the same problems was critically examined.

A model for the distribution of solution times of individual solvers was developed from the idea that the solution of a problem is composed of several steps or

stages, and solution occurs upon completion of the k-th stage. Problems can be classified as to the number of stages, which is assessed from the distribution of individual solution times. The Classification Model (statistically a waiting-time model) predicted solution times to be a gamma distribution. The argument proceeded logically to the gamma distribution from such assumptions as that the probability of a stage solution is constant over time until solution occurs, any one time interval is taken so small that one and only one stage may be solved within it, and the stages are independent and equally difficult.

Two models for predicting group from individual performance were proposed, viz., the Hierarchical and Equalitarian Models. These models predicted the distribution of group solution times to be a simple transformation of the gamma distribution that fitted the individual solution times. The Hierarchical Model assumed that group members organized themselves into a hierarchy with the more successful members consuming more than their share of the group's time. The Equalitarian Model assumed each member took his share of the group's working time, whether he contributed to solution or not.

Data were gathered from individuals working alone

and from ad hoc groups of four. These data were: (a) correctness of solution, and (b) the time consumed on each problem. Each group and each individual were tested on a sample problem and three experimental problems, the experimental problems being given in different orders for different subjects. In the groups, not only was solution time recorded, but also a count of the frequency with which each member talked, their choices for future problem-solving partners, observation of member contributions and any unusual events.

The parameter k of the gamma distribution (interpreted as the number of problem stages) was estimated by the method of moments from the sample observations. The theoretical curves thus determined were found to fit the distributions of individual solution times for all three problems. The Equalitarian but not the Hierarchical Model was found to predict the distribution of group solution times in each case. Implicit support for the social psychological assumptions of the Equalitarian Model was found through an analysis of the partner-choice data. The analysis of the communication frequencies of group members, however, was indecisive.

The problem solving behavior of individuals was pooled mathematically and such concocted groups were

found to perform significantly better than the real groups on two of the three problems. This finding was interpreted to indicate that member interaction actually inhibited problem solving, at least under the conditions of this investigation.

Frank Restle

Frank Restle, Major Professor

MODELS FOR THE CLASSIFICATION OF PROBLEMS AND THE PREDICTION OF GROUP PROBLEM-SOLVING FROM INDIVIDUAL RESULTS

By

James H. Davis

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

1961

ACKNOWLEDGEMENT

The author is deeply grateful for the guidance and intellectual stimulation of his teacher, Dr. Frank Restle, who was never too busy to listen and comment.

Important suggestions and valuable criticisms were offered by the members of the committee, Dr. Milton Rokeach, Dr. Charles Wrigley, and Dr. Terence Allen; the author is greatly in their debt.

DEDICATION

To my wife

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

In his review of human problem solving, Duncan (1959) summarized his conclusions thus: "The field of problem solving is poorly integrated. The reasons for this seem to be the use of a great variety of tasks to provide problems, the frequent use of unanalyzed and non-dimensionalized variables,* the lack of an agreed-upon taxonomy of behavioral processes, and to some extent the failure to relate data to other data or to theory. Problem solving particularly needs research to determine the simple laws between dimensionalized independent variables and performance" (1959, p. 426).

Duncan was, of course, referring particularly to investigations of the problem solving process *per se*. But there are two sub-areas of research and theory that are of particular interest to the social psychologist. First, there are studies of group problem solving in which members of a small group collectively address themselves to some problem or task; of major interest are those cases where the group's performance has been contrasted with individual effort, (Lorge, Fox, Davitz, and Brenner 1958).

---

*Emphasis supplied.

Second, small group researchers have often used problems and other tasks to focus the group's effort, while their main interest was in manipulating other situational variables and observing their effects on group structure or various group processes, e.g., creating an out-group threat and observing the increased cohesion of the group under study.

It is the writer's contention that the poor organization and integration of studies, reported by Duncan, creates a difficulty for those who use problems in the course of investigating other phenomena, e.g., communication structures in small, task-oriented groups. At the heart of many difficulties is the lack of knowledge about the task or problem itself: How may problems be ordered or classified to provide some continuity across situations and provide a basis for comparison between studies employing different problems in the study of the same phenomenon?

The lack of a systematic treatment of the problem itself precludes the kind of social-psychological theory that deals simultaneously with the nature of the problem, the group product (task performance), and the social behavior of the individuals in the group. Even regarding the group network studies, which deal with the effects of imposed communication structure on group

performance, Glanzer and Glaser comment, "Perhaps the most surprising thing about the entire area has been the fact that despite the highly formal origins of these studies, the organized body of theory has not yet appeared" (1957b, p. 35).

With this overview in mind, the purpose of this dissertation can be described as an attempt, first, to introduce an ordering of problems and a method for measuring this order, and second, to predict group problem solving performance from a knowledge of the problem solving behavior of persons working as individuals. Major emphasis is placed upon the development of a model that deals concurrently with group product (problem solving performance) and the emergent structure of the group. In addition, the frequently noted superiority of problem solving groups over individuals working on the same problems (Shaw, 1932) is critically examined.

RESTRICTIONS OF SCOPE

Before examining previous research it is well to specify the restrictions of this inquiry. Only verbal problems will be considered, for these generate interpersonal interactions and lead to social organization. Only problems with definite answers, which the subject can solve without special outside knowledge, are useful

in the present experiments. These problems are called "Eureka problems," (Lorge *et al*., 1958, p. 355). Only *ad hoc* groups with unrestricted face-to-face contact will be considered. Groups may restrict their communications or choose leaders, and a structure may emerge, but the present study does not involve imposing a communication network or social organization upon the group.

CLASSIFICATION OF WORD PROBLEMS: STUDIES FOCUSING ON THE PROBLEM AND ITS ELEMENTS

Despite the need for studying the problems themselves, implied above by Duncan, the literature contains little that represents a systematic attack on the behavioral effects of the particular problem--especially word problems. Studies focusing on problems *per se* have been mainly concerned with (a) criteria for selecting problems, (b) presenting problems by different methods, or (c) varying the elements of problems. These will be considered in turn.

a. Criteria for selecting problems. Ray (1955) has developed several criteria for tasks suitable for problem solving studies. For example, one should choose problems which may be scored along a continuum; the problem should restrict subjects to few hypotheses about its solution, etc. Unfortunately, Ray did not

consider verbal problems at all. Marks (1951) suggested that problems should be plausible (stimulate interest), complex and difficult (elicit solving behavior), structured (allow quantification and the recasting of the problem in different contexts), and solvable.

The selection of problem criteria does not really meet the issue of problem classification as set forth in this study. Not only does "meeting the criteria" depend heavily on subjective decisions by the investigator, but it avoids the question by restricting the kinds of problems that may be used. For instance, Marks' criteria would exclude many word puzzles, although these have been among the more important tasks for studying social interaction. However, the problems selected for this study appear to meet the major criteria outlined by Ray and Marks, viz., continuum, restricted hypotheses, plausible, complex, structured, and solvable.

b. Variations in Method of presentation of problems. There have been several studies in which the experimenter varied a problem's "concreteness" (Saugstad, 1957; Lorge, Tuckman, Aikman, Spiegel, and Moss, 1955a, 1955b, Cobb and Brenneise, 1952; and Gibb, 1956). Lorge <u>et al</u>. (1955a, 1955b) presented the "mined road problem" at seven different levels of reality (verbal,

photographic, miniture scale model, real presentation, or various amounts of the model and real situation). Level of reality appeared to have little effect on problem-solving performances. These investigators were not concerned with ordering problems, but rather explored the effects of administering the same problem under different conditions. Style of problem presentation, therefore, does not bear directly upon the question of problem classification.

c. The effects of manipulating the elements of a problem. Katz (in Duncan, 1959) found that increasing the degree of disorder in which digits were presented increased the errors made in computing sums. Benedetti (1956) searched for those aspects of Luchins water jar problems which might be contributing to rigidity; he added one, two, or three jars to the original problems with noticeable effect. The increased freedom of choice seemed to destroy much of the rigidity noted with the series of problems. Judson and Cofer (1956) presented a group of four words to subjects who were to choose the word that was out of place; each contained two ambiguous words and two unambiguous words. The subjects tended to choose on the basis of the first appearing unambiguous word. Dominance of the first occurring ambiguous word was increased by increasing the number of ambiguous words between the

two unambiguous words.

Battig (1957), using a task that met Ray's criteria, had his subjects play a word game in which they were to guess a word, given only the number of letters it contained. A significant difference in variability of performance between words was found, much of which seemed attributable to the specific letters contained in the words. There was limited evidence that word length and frequency of usage were important variables.

Admittedly, the studies dealing with the effects of varying the elements of a problem were not concerned with ordering problems and introducing a method for measuring this order. Various manipulations of the number of pieces, number of alternatives, etc., are not represented in this investigation. But this last group of studies does illustrate the importance of a particular problem and the behavioral effects of slight variations in the problem. Presumably the addition of problem elements would increase a problem's complexity or heterogeneity. The necessity of a solver's choosing from among the added elements would add another stage to the problem-solving process. Duncan has said, "In contrast to the experiments on methods of problem presentation, studies of variation among problem elements consistently reported at least some significant

effects, occasionally powerful effects, on problem solving performance. Thus, performance on a problem may or may not be influenced by contextual variables, such as methods of presentation that do not change relationships among elements of a problem. But changes of a problem's internal structure usually influence performance, even in cases where the problem remains, in some physical sense, the same." (1959,p.410) Duncan's recent review of human problem solving includes an important summary of studies of tangental interest to the issue of problem classification.

PROBLEMS AND GROUP SOLVING

It is no surprise that in studies employing some kind of task the particular task seems to make a great deal of difference in group behavior. In this connection, Lorge *et al*. have said, "...generalizations about problem solving and about group superiority seem to depend upon the nature of the tasks" (1958, p. 356). The same authors generally criticize the research on group versus individual problem solving on the grounds that investigators have paid insufficient attention to the particular task; the usual tasks are petty, unreal, etc. But most of all, the diversity of the tasks used by different investigators has been so great that it is difficult to draw conclusions about the superiority of

groups over individuals.

An ordering of tasks within a problem class would seem to be a first step in studying task-inspired influences on group-individual differences. The present study explores an approach to the ordering of problems, the number of steps or stages to arrive at a solution.

In a recent theoretical study, Roby and Lanzetta (1958) have emphasized the importance of the particular task to group functioning. They suggest that "objective" task properties be given equal status with "modal" task properties. An objective property is one for which a single, definite, value may be specified through either measurement or control. A modal property "...reflects 'typical' behavior of individuals or task variables but is subject to variation due to group characteristics and their interaction effects with other task properties" (Roby and Lanzetta, 1958, p. 89). These authors present a rather complex descriptive paradigm of the interaction of "group-task systems." Although, they do not present a true classification scheme for problems, Roby and Lanzetta point out that a severe limitation on such a design has been the general neglect of task parameters in small group research.

The classification and prediction models, to be presented subsequently, take cognizance of Roby and

Lanzetta's recommendations. The Classification Model represents an attempt to deal with at least one kind of task parameter (stages to solution). The prediction models relate group properties with task properties while focusing on the product of group interaction--problem solution.

GROUPS VERSUS INDIVIDUALS

Formal models of small group behavior appear to have been very slow in developing. Hays and Bush (1954) have expressed surprise that so few mathematical models have been constructed for use with group interaction, inasmuchas the utility of such an approach has been demonstrated in learning theory, etc. These authors speculated that it is the great complexity of social interaction which has hampered the development of mathematical models. They also pointed out that in the early stages of model construction the particular task is crucial. Many suitable tasks could be appropriated from experimental psychology. These should be (a) tasks in which the number of behavior possibilities is limited (important for model building) and (b) tasks whose behavior-producing capacity is well known for individual subjects.

Among the models which have recently been suggested are those of Hays and Bush (1954), and Lorge and

Solomon (1955).

The work of Hays and Bush draws directly upon that of Bush and Mosteller (1953). Two models for group action are proposed and an experimental test of the predictions, using three-man groups and a two-choice, probability-type learning task. The "Group-Actor Model" assumes that the group acts as if its guesses are those of an individual, i.e., individuals and groups were taken to be entirely interchangeable actors in this situation. The "Voting Model" assumes that no basic change occurred between the individual guessing alone and the guessing that takes place in a group, i.e., three men behave as if they were independent voters. Neither model was rejected when compared with group data. In fact, the observed group results fell between the predictions of the two models and the authors interpreted the models as establishing limits for the behavior of such groups.

The authors lamented that they did not collect individual data which would have allowed clearer interpretation. Unfortunate, too, is the fact that the Hays-Bush models appear to be restricted rather definitely to the two-choice learning task they employed. Nevertheless, their work demonstrates the feasibility of such an approach and the possibility of making use of existing models from other fields in the study of small

groups. Their results forced them to shift the focus of attention from the adaptive to the integrative aspects of interaction--something not anticipated until they began to speculate about the requirements of a more suitable model.

Lorge and Solomon (1955) reinterpreted Shaw's 1932 results with a group problem solving model which utilizes results obtained with individual solvers. The Lorge-Solomon model predicts the solution probability of an r-man group to be the probability of one or more solutions by the r individuals working alone. Thus the model relies upon a simple pooling effect, and interaction is assumed to have no effect on the probability that a group will solve a problem. These authors proposed their model as an alternative interpretation of the results obtained by Shaw (1932).

Generally speaking, groups are considered to be superior to individuals in solving problems. The evidence for such a position is much less than one might suppose, considering the extent to which the view is held. The preëminent position of the group in solving problems is reflected by the following quotation from Kelley and Thibaut: "...group problem solving involves individual problem solving, but much more. The 'extra' does not refer to a 'group mind' or a magical 'plus' which arises out of groups. Rather, it refers to the

simple fact that the thinking done by a group member occurs in a different context from that done by the 'isolated' thinker. Because it occurs in communication and interaction with other persons, the products of the group member's thought may be unpredictable from the observation of solutions obtained in isolation" (1954, p. 738).

Watson (1928) conducted one of the early studies to compare groups and individuals working at the same problem. Individuals and groups of three to ten members were to produce small words from long ones. The average number of words produced by the group was considerably higher than that produced by the best individual. Watson then examined his groups more closely; he added together all of the different words of individual subjects and compared the results with their subsequent behavior in a group. The real groups were found to be clearly inferior to the "summated" groups. These results suggested that real groups did not take into account the resources of all members.

The most frequently cited study in the area of group problem solving is that of Shaw (1932). In the first half of the experiment, individuals and groups of four worked on three similar eureka problems. In the second half, subjects were given three rather different problems--unscrambling words to form the

last sentence of a prose passage, unscrambling words to create the last lines of a sonnet, and discovering the shortest route for two school buses, given certain conditions.

Shaw interpreted her results as indicating a clear superiority in favor of groups, and accounted for this on the grounds that groups tended to reject incorrect solutions by members and to guard against errors. However, Shaw grouped all of the problems together when comparing the proportion of correct solutions among groups with that among individuals. Lorge and his co-workers have closely scrutinized the Shaw experiment. "For the first period, on the so-called Eureka problems, three of the 21 individuals and three of the five groups solved the first problem; no individual and three groups solved the second problem; and two individuals and two groups solved the third problem. No individual solved more than one problem, but just three groups made the eight group solutions. Two groups and 16 individuals never solved any of the three puzzles. For the second period problems, three of 17 individuals and four of the five groups solved the first problem completely; a fifth group and seven other individuals made just one error. No individual and no group solved the other two problems. Group superiority rests only on the eight solutions by groups in contrast

with the five by individuals. In general, interpreters of the Shaw experiment have disregarded not only the similarity among the three problems but also the fact that the solutions were based on the sum overall problems rather than on the number of identical solutions by groups. For instance, when only the solutions by individuals and by groups for the first problem are compared, there is no statistically significant difference. Shaw neither discussed the fact that two of the groups never solve any of the three puzzles, nor the relative efficiency of three solutions among 21 individuals versus three solutions for five groups of four members each, i.e., 20 individuals altogether" (1958, p. 355).

Marquart (1955) advanced criticisms similar to those of Lorge. Essentially replicating the Shaw experiment, Marquart obtained comparable results with groups and individuals. She analyzed her data so as to take account of the effect of grouping individuals, and she then discovered the group-individual discrepancy to be negligible. In fact, individuals had a slight advantage over groups. (Note the similarity between Marquart's results and the earlier findings of Watson (1928).

Thorndike (1938) obtained the usual group superiority in comparing individuals with groups on

four verbal problems. The problems were presented so that some subjects could make only a limited number of responses to them; other subjects were allowed an unlimited number of responses. Thorndike hypothesized that as the range of responses increased the superiority of groups over individuals would increase. In general, the results tended to confirm his hypothesis. Unfortunately, Thorndike made no analysis that would throw light on the question of group superiority based only on the pooling of individual solutions.

Husband (1940) found pairs superior to individuals in solving word puzzles and arithmetic problems. However, groups needed to solve the problems in one-half the time required by individuals in order to "justify" themselves. When the data were analyzed in terms of how efficiently groups and individuals used their time, he concluded that the time saved by the pairs was never more than a third.

Taylor and Faust (1952) compared group and individual performance on a game of Twenty Questions which resembled a eureka problem. The proportion of solutions favored groups over individuals. However, in terms of man-hours required for a solution, individuals were superior.

Faust (1959) reported two group-individual experiments, identical except for a slightly different

composition of subjects. In both experiments he found that four-man groups averaged more problems solved, both verbal and spatial, than did individuals. When these observed groups were compared with "nominal" groups, generated from the mathmatical model of Lorge and Solomon (1955), a significant difference was found on only one of the problems used with the second experiment. In this case the real group was superior to the "nominal" group. Thus, the results of Faust's well-executed study are generally in line with previous experimental evidence.

Faust pointed out that the Lorge-Solomon Model A provides a much-needed baseline against which to compare the effects of social interaction in real groups. (The same proposition is advanced later in the present study).

The preceding discussion suggests that more groups are likely to solve a problem than individuals. That the interpretation is not as simple as it first appears is indicated by the fact that the group superiority disappears when individuals are artificially pooled in a manner approximating the group aggregation. On some efficiency measures, time for example, the individuals turned out to be superior when pooled in this way. Watson (1928) even suggested that group interaction, instead of providing an extra boost, resulted in

restraint.

A recent review by Lorge et al. (1958) provides a convenient summary of group-individual experiments. One should consult this review for studies not limited to ad hoc groups and eureka problems.

The above studies suggest that small, problem-solving groups have not received the attention they deserve in regard to formal models or the precise organization of group-individual solving data. The recommendation by Hays and Bush that existing models from other fields may be modified for use with small groups seems quite appropriate, considering the importance of the work of Bush and Mosteller (1955) to the developments in the next chapter.

SUMMARY

The study of group problem solving requires (a) an apt and relevant description of the problems used and how they are solved, (b) a systematic way of representing both the pooling of individual work and social interaction in group work, and (c) a sufficient volume of data to permit an exact comparison between groups and individuals. A review of previous work indicates that (a) There is no suitable framework for describing problems. (b) The pooling of individual accomplishments is clarified by the Lorge-Solomon model, which however

does not deal with social interaction. Some possibilities of social interaction are bracketed but not specified by the work of Hays and Bush on group learning. (c) Experimental comparisons of group and individual problem-solving are neither numerous nor, in the main, done on a sufficiently large scale to give a clear indication of whether, and to what degree, groups can outperform individuals.

CHAPTER II

THEORY

The problem solving process, which begins with the presentation of the problem, is terminated when the subject arrives at a correct solution, when he arrives at an incorrect solution he believes correct, or when the experimenter arbitarily ends the experimental session. Although the proportion of correct solutions is a standard means of describing experimental results in the area of problem solving, the inefficiency of such an approach is revealed by the question: What proportion of correct solutions occurs after what time interval? The choice of different time intervals (two seconds, three minutes, 12 hours, etc.) can yield different proportions. A distribution of solution times over some interval is more desirable. Consequently, enough time should be allowed so that the subjects themselves terminate the process. These more nearly complete data are advantageous for: (a) classifying problems, and (b) constructing a model for prediction of group problem solving.

In this chapter attention is focused on the distribution of the times when the problem solving process terminates, and the correctness of answers, when subjects work individually and when they work in

small *ad hoc* groups. When groups of subjects work together on the problem it is also possible to observe their statements, questions, etc., and to study the social structure which emerges during the problem solving process.

The central issue is the prediction of group performance from individual performance. It does not seem feasible to attempt to predict the performance of each individual or group separately, because one can collect only a small amount of information about each subject aside from his actual problem solving performance. Hence the problem may be approached through a stochastic model which attempts to account for the distribution of observations gathered. In this way the complete data, including mean, variance, and shape of the obtained distribution, can be brought to bear on the theoretical issues.

CLASSIFICATION: A MODEL FOR THE DISTRIBUTION OF SOLUTION TIMES

Under standard conditions a particular problem should consistently yield a characteristic distribution of solution times for samples of subjects drawn from the same parent population. The following discussion derives a theoretical distribution, using the approach of Bush and Mosteller (1955, Chap. 14). Consider first

that time is divided into short, equal intervals of duration h. In any interval the subject may either solve the problem or not. Imagine that there exists a probability p that the problem is solved in the N-th interval, given that it has not been solved previously. The probability p is taken to be constant over intervals and each interval represents an independent event.

The probability that the problem is solved in the first interval is p. The probability that the problem is first solved in interval 2 is (1-p)p, which is the joint probability that it is not solved in the first interval and is solved in the second. The probability that the problem is solved in the N-th interval is $(1-p)^{N-1}p$, the probability that it is not solved in any of the first N-1 intervals and then is solved in the N-th interval. Since each interval is of length h the solution time T, when the problem is solved in the N-th time interval, is

(1) $$T = Nh$$

Let the probability of solution at time T be called P(T). In this simple one-stage process,

(2) $$P(T) = p(1-p)^{N-1}$$

which is the geometric distribution. This is a positively skewed distribution with mean 1/p and variance $(1-p)/p^2$.

The problem solving processes dealt with here may

be expected to approximate the situation given above, except that solution of the problem may require the completion of more than one stage. Now consider that there are k stages to be accomplished, and that solution of the problem is recorded when all k stages have been completed. Assume further that each of the k stages has probability p of being completed in a given time interval of duration h, and assume that h is chosen so small that two stages are never completed in the same interval.

The probability that the k-th stage is completed at trial N is given by the negative binomial (Pascal) distribution,

(3) $$P(N) = \binom{N-1}{k-1} p^k (1-p)^{N-k}$$

An intuitive justification of this formula is given by Bush and Mosteller, and its derivation is also discussed by Feller (1957, pp. 155-157). (Note that (3) reduces to (2) when k = 1). It can be shown that the expected number of intervals, E(N) is

(4) $$E(N) = k/p$$

and, since each interval is of duration h, the expected time to solution is

(5) $$E(T) = h.E(N) = hk/p \quad .$$

Furthermore, the variance of the negative binomial distribution is

(6) $$\sigma^2(N) = k(1-p)/p^2 \quad .$$

The variance of the time scores is obtained by multiplying the variance of N by h^2, which gives

(7) $$\sigma^2(T) = h^2 k(1-p)/p^2$$

Other statistical properties of the Pascal distribution are discussed in standard sources on probability theory and mathematical statistics (Fisher, 1950; Feller, 1957).

The above model is not quite satisfactory because, in the problem solving process, time is measured as a continuous variable rather than as a succession of finite intervals. The above argument is applied to continuous time by letting the duration of an interval, h, approach zero as a limit while keeping the "rate," or probability per unit time p/h, constant. Let the first interval be designated 0, the second 1, etc.; since T = Nh, the geometric distribution of (2) becomes

(8) $$P(T) = (1-p)^N p = (1-p)^{T/h} p \quad .$$

Let $p/h = \lambda$ be a constant so that $p = \lambda h$ and

(9) $$P(T) = (1-\lambda h)^{T/h} \lambda h \quad .$$

If one takes the limit of p(T) as $h \to 0$, the term λh goes to 0 and, consequently P(T) approaches 0. Therefore let

(10) $$f(T; \lambda) = \lim_{h \to 0} \frac{(1-\lambda h)^{T/h} \lambda h}{h}$$

$$= \lambda \lim_{h \to 0} (1-\lambda h)^{T/h}$$

and it can be shown that $(1-\lambda h)^{T/h}$ approaches $e^{-\lambda T}$.

Hence,

(11) $$f(T; \lambda) = \lambda e^{-\lambda T} \quad .$$

If each individual stage has the distribution $f(T; \lambda) = \lambda e^{-\lambda T}$, the distribution of solution times to complete k stages in the gamma distribution

(12) $$g(T; \lambda, k) = \frac{\lambda}{(k-1)} e^{-\lambda T} (\lambda T)^{k-1} .$$

The gamma distribution $g(T; \lambda, k)$ may be readily obtained from $f(T; \lambda)$ by the use of moment generating functions, but such a development is beyond the scope of this thesis.

From formula (12),

(13) $$E(T) = k/\lambda$$

and

(14) $$\sigma^2(T) = k/\lambda^2 .$$

Figure 2.1 shows the graph of gamma distributions with k = 2 and k = 4. As k increases the graph of the function becomes more nearly symmetrical.

The parameter k depends on the structure of the problem and how it is solved by the subjects; k must be estimated from the data.

An estimate of k can easily be obtained using only the first two moments of the distribution. Recall that equations (13)
(14) gave the mean and variance as

$$E(T) = k/\lambda$$

and

$$\sigma^2(T) = k/\lambda^2$$

Therefore,

(15) $$\frac{[E(T)]^2}{\sigma^2(T)} = \frac{k^2/\lambda^2}{k/\lambda^2} - k$$

Replacing the population expectation and variance by estimates from the data, $\bar{T}$ and s_T^2 respectively, one has

(16) $$\hat{k} = \frac{(\bar{T})^2}{s_T^2} \quad .$$

PREDICTION OF GROUP FROM INDIVIDUAL PERFORMANCE: LORGE AND SOLOMON MODEL

Lorge and Solomon (1955) have suggested a non-interactional ability model, which however does not take into account the distribution of solutions over some experimental time interval. Their model provides the main example of an approach to the prediction problem. Lorge and Solomon submit the hypothesis that group superiority on eureka problems is not due to social interaction, but simply to the abilities of the members. "Such an hypothesis may be expressed in terms of two ability models: (A) group superiority is a function only of the ability of one or more of its members to solve the problem without taking account of the interpersonal rejection and acceptance of suggestions among its members; (B) group superiority is a function only of the pooled abilities of

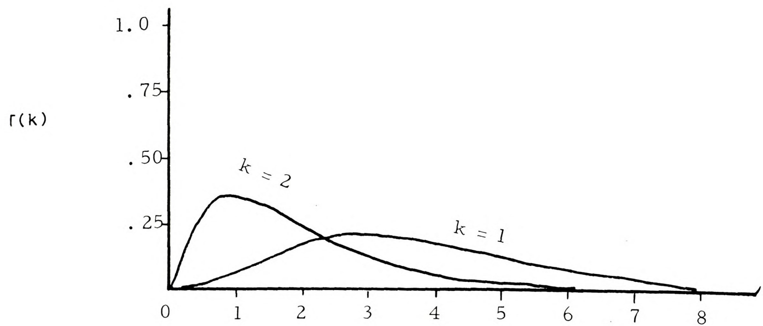

Figure 2.1 - Graph of the function $\frac{1}{(k-1)!} T^{(k-1)} e^{-T}$.

its members. The latter model, B, implies that any problem may be composed of, and solved in, two or more stages. Model B reduces to Model A for one-stage problems." (1955, p. 140)

Consider first the Lorge and Solomon Model A.

Let

P_G = probability that a group of size r solves the problem;

P_I = probability that an individual solves the problem.

Now $1-P_I$ is the probability that an individual will not solve the problem, and the probability that none of a group of r persons will solve is $(1-P_I)^r$. Consequently the probability that a group of size r will solve the problem is

(17) $$P_G = 1-(1-P_I)^r \quad .$$

Equivalently

(18) $$P_I = 1-(1-P_G)^{1/r} \quad .$$

Model B: Recall that a problem may be solved in two or more independent stages. Equations (17) and (18) may then be generalized to the k-stage case. Let P_{Ij} be the probability that an individual will solve stage j.

(19) $$P_G = \prod_{j=1}^{k} \left[1-(1-P_{Ij})^r\right] \;, \quad P_I = \prod_{j=1}^{k} P_{Ij}$$

where Model A applies at each stage j. If $P_{Ij} = P$ is assumed to be the same for all stages,

(20) $$P_G = \left[1-(1-P_I^{1/k})^r\right]^k$$

(The argument may be extended to the case where P_{Ij} is not constant from stage to stage).

The Lorge-Solomon model does not take into account the distribution of solution times, and the results may depend upon the time given the subjects.

PREDICTION: COMBINATION OF CONTRIBUTIONS MODEL

The Lorge-Solomon approach may permit prediction of the proportion of groups which solve a problem, but, as just indicated, such a proportion depends upon the time given the solvers. A satisfactory resolution of the theoretical problem involves predicting the distribution of group solution times from the distribution of individual solution times.

The problem will first be taken up for the idealized case in which all individual subjects and all groups solve the problem. It is necessary later to consider the consequences of having nonsolvers in the experiment, but this discussion will be postponed until later.

To evaluate the possibility of group facilitation or interference, one must first decide the performance which would result if the members of the group proceed without any important effect on one another's success.

In analyzing this situation it is most convenient to proceed at once to the model of continuous time. Recall that in the case of individual solvers the probability of solution per unit time was λ. If one has a group of r independent solvers contributing to the group, then the probability per unit time is $r\lambda$. This approach can be used, despite its great simplicity, because it is assumed in going to continuous time that the time interval h is very small. As h approaches zero the probability that two different subjects will both make a contribution at the same time becomes negligible. Hence, there exist none of the difficulties, discussed by Lorge and Solomon, of taking account of the possibility that a group may contain more than one solver. Since the concern here is with solution times, even if a group has more than one solver, one may neglect the infintesimal possibility that both solve at exactly the same time. Supposing, in the same spirit as the Lorge-Solomon model, that the group solves the problem immediately after the first subject solves, groups will differ from individuals only in the probability-per-unit-time parameter, which in groups is $r\lambda$.

The above argument obviously applies to the one-stage problem. However, it should be clear that it applies equally well to the k-stage problem, for the

number of subjects in the group can in no way change the number of stages in the problem. The prediction of group performance from individual performance, in this idealized case, is very simple. Let the gamma distribution for probability-per-unit-time of λ and k stages be called $g(T; \lambda, k)$.

Hypothesis: If performance of individual solvers is described by $g(T; \lambda, k)$, then performance of groups of r solvers will be described by $g(T; r\lambda, k)$.

As was mentioned at the beginning of this chapter, subjects can arrive at wrong answers. In the case of individuals working alone one can make quite adequate estimates of the proportion of subjects who will not arrive at the correct answer. When there is, say, one such person in a group of four, the group will very probably arrive at the correct answer. In fact, the nature of the experimental problems being what they are, it is unlikely that a whole group of size four will arrive at a wrong answer. However, it seems evident that a group member, who alone would arrive at a wrong answer (or fail to arrive at an answer at all), is not likely to aid his group in arriving at the correct answer.

One may formulate two hypotheses about the effects of these wrong answers and no answers, and these two hypotheses have different implications for the

distribution of success of groups. One asked of the data will be which of these two hypotheses is the more reasonable.

The first hypothesis is that non-solvers are non-functional in their groups. If so, a group with two solvers and two non-solvers will solve the problem just as rapidly as a group of two solvers alone. This would happen if solvers suppress non-solvers, forming an intellectual "hierarchy." The second hypothesis, which seems to square better with informal observation, is that the non-solvers consume some share of the groups working time, though they do not contribute to solution. This would reflect an equalitarian group structure. In this case, a group with two solvers and two non-solvers should take longer than a group made up solely of two solvers. If it is assumed that non-solvers consume their share of the group's time, the following proposition is suggested. Let A be the number of solvers and B be the number of non-solvers in a group, where $A + B = r$ is the number of persons in the group. The probability of solution-per-unit-time for each of the A solvers will be $[A/(A + B)]\lambda$. Thus, if there are A solvers in the group and $[A/(A + B)]\lambda$ is the rate at which each contributes, the group rate is $A \cdot [A/(A + B)]\lambda = [A^2/(A + B)]\lambda$.

In other words, if the distribution of solution

times of individual solvers is described by $g(T; \lambda, k)$, then the two hypotheses are:

Hypothesis 1, Hierarchical Model: Non-solvers have no effect, and the group's distribution of solution times will be described by $g(T; A\lambda, k)$;

Hypothesis 2, Equalitarian Model: Non-solvers consume time, and the group's distribution of solution times will be described by $g(T; [A^2/(A + B)]\lambda, k)$.

Except by extremely detailed and accurate protocols of group discussion, solvers are indistinguishable from non-solvers within a group. However, suppose that proportion a of the individuals solve the problem correctly, and the remainder, arrive at a false conclusion or no answer and are non-solvers. Since the groups are formed randomly, the probability that a group will have exactly A solvers and r-A = B non-solvers is given by the binomial distribution,

(21) $$P(A) = \binom{r}{A} a^A (1-a)^{r-A} .$$

The performance of a collection of groups can now be predicted. Hypothesis 1, that non-solvers have no effects, gives the proposition:

The distribution of solution times of a collection of groups will be described by (Hierarchical Model)

(22) $$W_H(T) = \sum_{A=0}^{r} \left[P(A) \; g(T; A\lambda, k) \right] .$$

Similarly, Hypothesis 2, that non-solvers consume time, leads to the proposition:

The distribution of solution times for a collection of groups will be described by (Equalitarian Model)

$$(23) \qquad W_E(T) = \sum_{A=0}^{r} \left[P\ (A) \quad g(T;\ \left[A^2/(A + B)\right]\lambda\ ,\ k\right) \ .$$

These last two distributions do not have any very simple explicit forms. However, all the parameters (a, λ, and k) can be estimated from the performance of individuals (by the method of moments). Then it is possible to compute the above distributions simply by taking the gamma distributions from a table (Pearson, 1922) for various values of k and T, and computing the weighted average, following the above equations through. With groups of size $r = 4$, this process is not overly laborious. The predicted distributions can be compared with the obtained results.

If, as in the present study, the fitted gamma distribution is very close to the obtained distribution of individual solution times, the predictions can be made directly from the distribution of solution times obtained from the individual solvers without going through the fitted gamma distribution.

A final note on the use of estimates in making predictions is in order. Some practical problems are associated with the use of $\hat{k}$ rather than the parameter itself. The sampling distributions of these estimates are not known and it is not possible to set up confidence

intervals in the customary fashion. Large samples (such as those used in the study to be reported in Chapter IV) will be of some assistance with this problem. If predictions are made from $\hat{k}$ instead of k, but the predictions are treated as fixed values, one fails to take account of the sampling variations of the predictions. Hence, any regular statistical test is overly stringent.

SUMMARY

The distribution of group solution times is predicted from λ (individual solver rate of solution), k (stage-structure of the problem), and the distribution of A (the manner in which groups are formed). The parameters k and λ are estimated from the first two moments of the sample of individual solvers. The proportion of all individuals who solved a problem, <u>a</u>, can be used with the binomial formula to give P (A).

The intuitively more difficult problem of predicting the distribution of group solution-times actually turns out to be easier than the problem of predicting the probability of group solution. Conceptually, this represents an improvement over the Lorge-Solomon formulation.

CHAPTER III

EXPERIMENTAL PROCEDURE

The quantitative models stated in Chapter II were submitted to experimental test. A large number of college students were given several Eureka problems to solve, and their times to completion and correctness of solution were observed. These data are the basis for evaluating the gamma-distribution hypothesis, and serve as the source of estimates of λ, k, and a. A number of *ad hoc* groups of four students each were also given the same problems under similar conditions except that the members of group cooperated in solving the problems. Comparison of group with individual performance were used to choose between the Hierarchical and Equalitarian models. Data on communication and social structure within the working groups were also collected.

SUBJECTS

The subjects were 267 volunteers from an introductory psychology course. The sign-up sheets had spaces for four subjects at each experimental session. When four subjects signed up and appeared on time, they were tested as a group. Whenever a group could not be formed, as when only two or three subjects signed up, a subject failed to appear, or some other accident occurred, the

subjects present were tested as individuals. This method of assignment to the Group and Individual Conditions was not random, but the assignment of any particular S who appeared on time at the laboratory depended solely on the behavior of other students who did or did not sign up, or who did or did not report on time. It is likely that the Group and Individual Conditions were composed of comparable samples of the population. This recruiting method was used to maximize the number of four-person groups obtained with the least waste of subjects. Subjects at any one experimental session, group or individuals, were of the same sex. There were twelve groups of men and ten groups of women. There were 98 men and 81 women tested as individuals.

APPARATUS AND MATERIALS

The room in which the experiment was conducted was a 40 ft. x 22 ft. seminar-type classroom. For the experiment, two tables were removed from their usual semi-circular arrangement and placed to one side. Four straight-backed chairs were placed around one of them; the other table, behind which E was seated during experimental sessions, was located six to eight feet away. At the experimenter's table was a large faced clock (from which elapsed time could easily be obtained), four cumulative frequency counters, and an electric

timer which emitted a soft click every six seconds.

During a group session, the four Ss were seated around the table and a carboard plaque, bearing the identifying letter A, B, C, or D, was placed in front of each. Individual sessions were conducted with Ss seated at the group table and/or other tables located about six feet away, depending on the number of individual Ss available for any one session.

The four eureka-type problems used were labeled "Sample," 3, 5, and 8. The magnitude of the numbers was meaningless and served only to bolster problem security; if Ss perceived the problems on which they labored as only three of a larger series, they would possibly be less inclined to relate them to naive classmates. Each problem was mimeographed on a separate sheet of paper.

The first or sample puzzle, used for practice, was developed for this study in colaboration with Dr. F. Restle, although it has no doubt appeared elsewhere.

SAMPLE

A drunk is walking down the sidewalk; his condition is such that he is unable to proceed without stumbling and staggering. He decides that he wants to cross to the other side of the street. Unfortunately, he is so inebriated that he staggers back and forth, from one side of the street to the other--wasting a great deal of effort. Finally, he is able to stay on the side of

the street he had been aiming for. He continues his journey on that side of the street.

Among the following, which is the best guess as to the number of times he crossed the street in either direction?

1. 4 4.8

2. 5 5.6

3. 2

Puzzle #3 (hereafter referred to as the "Rope Problem") was taken from Dudeney (1958) and subsequently modified.

3.

My dungeon did not lie beneath the moat, but was in one of the most high parts of the castle. So stout was the door, and so well locked and secured withal, that escape that way was not to be found. By hard work I did remove one of the bars from the narrow window, and was able to crush my body through the opening; but the distance to the courtyard below was so exceedingly great that it was certain death to drop thereto. I did find in the corner of the cell a rope that had been left there and lay hid in the darkness. But this rope had not length enough, and to drop in safety from the end was nowise possible. So I made haste to divide the rope in half and to tie the two parts thereof together again. It was then full long and did reach the ground, and I went down in safety. How could this have been? (Answer: Divide the rope lengthwise and tie the ends of the unraveled rope together).

Puzzle #5 (hereafter referred to as the Double Problem) was adapted from one given by Kaufman (1954).

5.

Here is a simple question. It happens to be couched in complex phraseology, but that shouldn't disturb you if you can find a way to reduce it to the fundamentals. It can be done by rephrasing the question in such simple terms that the answer is immediately apparent.

Now go ahead:

"If the puzzle you solved before you solved this one, was harder than the puzzle you solved after you solved the puzzle you solved before you solved this one, was the puzzle you solved before you solved this one hard than this one?" (Answer: Yes; problem says it was harder in the beginning and the central portion is a self-defeating qualification).

Answer

Yes__________ No__________ Why?

Puzzle #8 (hereafter called the Dust Problem) was constructed by the writer, using the familiar Luchins water jar problem as a guide.

8.

As a prospector you have been very successful. You have a container that holds exactly 163 ounces of gold dust. You must discharge a debt by paying exactly 77 ounces of gold fudy (no more, no less) to a friend. You have containers that hold exactly 14 ounces, 25 ounces, and 11 ounces. How would you remove the exact amount from the safe, using only these four containers? You must use all four containers. (Answer: Subtract the following from the 163 ounces container: 2(25) + 14 + 2(11) = 86 or 25 + 2(14) + 3(11) = 86).

The above problems were developed on the basis of a pilot study. This preliminary investigation suggested problems which seemed (a) to yield a high probability of solution within 12 minutes and (b) to represent, intuitively, differing numbers of stages. The number of stages to solution of a eureka problem may be likened, in the ideal case, to the number of moves to solution of a chess problem. The chess player may or

may not make an appropriate or an inappropriate move in any one interval of time. Obviously a eureka problem is less specifiable than a chess problem and hence less readily analyzed into sub-strategies or stages. In fact, it is a goal of this investigation to provide a framework within which the eureka problem may be analyzed into moves or stages.

Consider, intuitively, the Rope Problem. Once a subject hits upon the idea that the rope may be divided lengthwise, escape from the dungeon is possible. On this basis the Rope Problem would consist of one stage; only one strategy is involved. In the Double Problem, on the other hand, the subject must note that the problem states the previous puzzle was harder; then he must realize that the subsequent verbiage refers only to the present puzzle which is less difficult. The answer is self-evident with the completion of these two or three stages. The Gold Dust Problem requires perhaps three or four stages to completion. The subject must realize that he is to subtract each container in turn, leaving the original with the correct amount of gold dust. The second stage is reached when the subject decides the number of times the 25 oz. container is to be subtracted. Stages three and four would be the number of times the 14 oz. and 11 oz. containers, respectively, are filled and emptied.

Note that these stage-estimates are intuitive. Future investigation may show, for example, that the Rope Problem requires two stages: one for the lengthwise unraveling of the rope and another for the joining of the ends together. Or it may even be discovered that reading and comprehending the nature of the problem requires one or more stages. Whatever future research reveals the nature of these stages to be, the present investigation is concerned only with empirically estimating their number for a particular problem.

The questionnaire used to evaluate group choice structure arising from collective effort is given below:

QUESTIONNAIRE

____
Your Letter

____
Age

____
Sex

Your answers to the following questions will be kept in the strictest confidence.

Consider each question carefully before you answer. Your response should refer <u>only to the immediately preceding</u>

problem! Problems differ a great deal.

1. In a few weeks, new groups may be assembled for work on similar problems; participants would be paid and considerable prize money would be awarded to the best-performing group. Which members of this group would you like to have with you, if you were to decide to participate in such a contest? Consider performance on the immediately preceding problem only. Indicate your choice (s) by a check mark(s) before the appropriate letter. You may decide to check none, one, two, or three of the other members of the group.

Think carefully.

_______Person A

_______Person B

_______Person C

_______Person D

a. As a separate issue, rank the other members of this group--one through four on the basis of how much you feel that they contributed toward the solution of the immediately preceding problem. In your ranking, include individual X (a fictitious person) whom you know to be an "average solver" of problems like the one you have just considered. Rank 1 would be the person contributing the most; rank 2 the next greatest contributor, etc.

_______Person A

_______Person B

_______Person C

_______Person D

_______Person X (An average solver of such a problem)

PROCEDURE

In both the Group and Individual Conditions the Sample Problem was given first, with a time limit of five minutes. The experimenter answered any questions and

then proceeded to the first of the three test problems; twelve minutes were allowed for each of these problems. For any experimental session, the order of presentation of the three problems was randomized. After E had distributed a copy of the problem to every S in the room, he retired to the table where the timing apparatus was located and gave the signal to commence work. The experimenter then recorded his observations.

1. Group Condition.

Only one four-person group was tested at an experimental session. When the Ss had been seated around the square table, E placed the cardboard plaque, bearing the identifying letter, in front of each. The experimenter then gave the following verbal instructions to the group:

"I would like to thank you for coming here this afternoon (evening). In a few moments I will give you a kind of word puzzle or word problem on which you are to work as a group. Talk freely among yourselves and try to 'think out loud' as much as possible."

"Try to work as rapidly as you can and still do the best work possible. It is very important, if this experiment is to be successful at all, that I have an accurate record of how long it took you, as a group, to solve

the problem. Please do not arrive at an answer you feel is correct and then wait until the alloted time has elapsed before you write down the group's answer.

"When you, as a group, have decided on a solution, one of you (anyone will do) should write it down immediately and then jot down the time remaining as given by this clock. You read the clock this way ...

"You will have three problems in all. A maximum of 12 minutes will be alloted for each problem. I will tell you when to begin and when to stop, if that is necessary. Are there any questions--I will not be able to answer questions after we begin."

After collecting the practice problem, the E answered any further questions and corrected the procedures if necessary. The E continued with the following:
"Here is a copy of one of the 'real' problems: you must stop and start when I signal: Begin."

Following each problem, E advised the group, "Here is a very short questionnaire, pertaining to the problem you have just finished; you will have three minutes in which to complete it. I know that you may be reluctant to answer at first, but I would like to have your honest opinion anyway."

E closed the session as follows: "I would like to thank you for your time and cooperation. It is very important to the success of this study that you do not reveal to others the nature of these problems; some of your classmates may receive some of the same problems when they serve as subjects later on. In fact, the best-performing group will win a prize of $8.00--about $2.00 per individual."

The experimenter emphasized certain aspects of the proceedings with informal asides. For example, it was stressed that the questionnaire referred only to the immediately preceding problem, and just before handing out the Double Problem it was casually pointed out that each problem was complete in itself and referred only to itself.

One measure of problem-solving activity was the frequency with which each S talked. During the interaction, E recorded the frequency with which each member communicated to another member or to the group as a whole. The recording was accomplished by means of four hand-tally counters--one for each member of the group. The clock and timer shielded this activity from the Ss and none of them gave any indication they were aware of the fact that they were being observed in this way. Every time a member addressed a communication to

another member or to the group he was counted as having "talked." Every six seconds, at the soft click of the timing device, E registered a count on one (or more) counters depending on which members were talking at that time. If no one was talking, no entry was made. This yielded a systematic time-sample of verbal activity of the Ss sufficient to indicate which one talked most often and longest, prerequisite to the description of a communication structure. The six-second interval was chosen as the shortest which could be recorded with precision by E under the conditions of observation and the other activities he had to maintain. Ideally, the total speech of a S would be recorded as a deflection on a kymograph and the area circumscribed by the base line and deflection would be taken as the communication measure.

Another important datum was, who did most to solve the problem? Right after completion of the problem, E made an intuitive estimate as to the major contributor to the solution. Furthermore, any unusual events which occur during problem-solving may lead to deviations from the overall theoretical predictions. Consequently, observations of unusual happenings were recorded--e.g., angry exchanges, pre-group friendships, etc.

2. Individual Condition

One, two, or three Ss were tested individually at an experimental session and were seated in the same general area as that occupied by groups. The experimenter gave the following verbal instructions:

"I would like to thank you for coming here this afternoon (evening). In a few moments I will give you a kind of word puzzle on which you are to work individually. Please do not talk among yourselves now or at any time during the course of the experiment.

"Try to work as rapidly as you can and still do the best work possible. It is very important, if this experiment is to be successful at all, that I have an accurate record of how long it took you to solve the problem. Please do not arrive at an answer you feel is correct and then wait until the alloted time has elapsed before you write it down.

"When you have decided on a solution, you should write it down immediately and then jot down the time remaining as given by this clock. You read the clock this way...

"You will have three problems in all. A maximum of 12 minutes will be alloted for each problem. I will tell

you when to begin and when to stop, if that is necessary. Are there any questions--I will not be able to answer questions after we begin."

After collecting the practice problem, E answered any further questions and corrected the procedures if necessary. E continued with the following: "Here is a copy of one of the 'real' problems; you must stop and start when I signal...begin."

Following each problem, E advised the individuals, collectively: "I would like for you to list the names of three of your acquaintances who, in your opinion, would do well on the kind of problem you have just finished. It is preferred that those named be students at MSU--but that is not necessary; they could be anyone at all."

The E closed the session as follows: "I would like to thank you for your time and cooperation. It is very important to the success of this study that you do not reveal to others the nature of these problems; some of your classmates may receive some of the same problems when they serve as subjects later on. In fact, the best-performing individual will win a prize of $2.00."

Again, E emphasized certain aspects of the proceedings with informal asides. For example, it was emphasized

that a particular name-listing applied only to the immediately preceding problem. And just before handing out the Double Problem it was pointed out that each problem was complete in itself and referred only to itself.

Each S received the same problem. After he had written down the solution, he recorded the time in minutes and seconds that had elapsed; the time was obtained from the large-faced clock.

When it became apparent that all Ss had finished, E waited a couple of minutes longer, to avoid creating a sense of urgency in the Ss finishing late, and then stopped the clock. Following the collection of papers, the Ss received a blank sheet of paper and were requested to name three of their acquaintances, any three at all, they felt might do well on the immediately preceding problem. This activity was designed as a counterbalance to the questionnaire of the Group Condition. The same data were collected for each problem.

SUMMARY

Each group and each individual was tested on the sample problem and three experimental problems, the experimental problems being given in different orders for different Ss.

Solution time was recorded for each individual on each problem. In the groups, a count of frequency with which each member talked, their choices of future partners in such problem-solving work, observations of contributions of each, and any unusual events, were also recorded.

CHAPTER IV

RESULTS

In direct comparison, group problem solving was found to be clearly superior to individual problem solving. As is shown in Table 4.1, a higher proportion of groups than individuals solved each of the three problems correctly. Although more groups solve the problems, groups and individuals do not seem to differ systematically on mean time to a right or wrong answer; see Table 4.2. In fact, standard t-tests applied to the differences between individual and group mean times on each problem do not show significant differences. However, mean solution time does differ from problem for both groups and individuals.

The distributions of answer-times (right or wrong) throughout the 12-minute experimental sessions are represented by the frequency polygons in Figures 4.1 (Individual Condition) and 4.2 (Group Condition). To facilitate graphical representation, subsequent discussion will deal with cumulative relative frequency distributions.

The statistical models discussed in the previous chapter use the idea that stages of problems are solved at randomly distributed times, so that total time to solution should have a gamma distribution, and the parameter k of the distribution should be the

Table 4.1 Individuals Vs. Groups on each experimental problem, T = 720 seconds.

	Individual Condition			Group Condition			Individual Vs. Group**
Problems	No. of Cases*	No. of Solvers	Prop. of Solvers	No. of Cases	No. of Solvers	Prop. of Solvers	(two tail tests)
3	178	133	.735	22	21	.955	$z = 2.48$ ($p < .013$)
5	177	89	.503	22	17	.773	$z = 2.70$ ($p < .007$)
8	163	71	.436	22	17	.773	$z = 3.30$ ($p < .001$)

*The relative frequency of correct answers for the Individual condition were based on different numbers of subjects. Some few subjects were discarded because they did not comply with the directions, found it necessary to leave early, etc. There was no evidence to indicate that any systematic bias affected the results.

** The statistical tests reported in Table 4.1 make use of normal curve theory where

$$z = \frac{p_1 - p_2}{\left[p(1-p)(N_1+N_2)/N_1N_2\right]^{\frac{1}{2}}}$$

and p_1 is the proportion answering correctly among N_1 individuals; p_2 is the proportion of correct answers among N_2 groups.

Table 4.2 Mean ± Standard Deviation of Solution Times in Seconds for Groups and Individuals.

Outcome	Problem	Individuals	Groups
Correct Solution	3	194.1 ± 115.5	156.7 ± 124.7
	5	325.8 ± 154.1	330.0 ± 216.2
	8	414.9 ± 166.7	419.4 ± 219.9
Wrong or No solution	3	249.1 ± 151.5	(Only one case)
	5	276.8 ± 169.0	376.0 ± 227.7
	8	462.5 ± 346.2	486.0 ± 217.2
All Subjects	3	208.0 ± 127.4	172.3 + 142.0
	5	301.5 ± 163.0	340.4 ± 214.1
	8	441.8 ± 282.6	434.5 ± 215.0

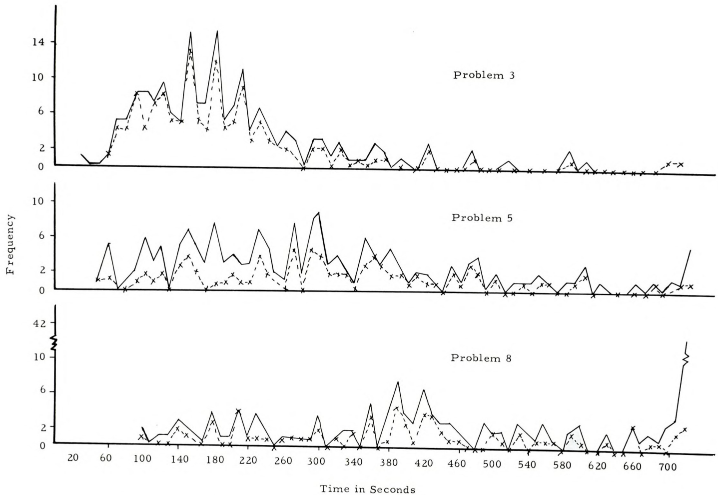

Figure 4.1 - Solid lines represent distributions of solution times for all individual subjects. The broken lines represent the distribution of solution times for individual solvers only.

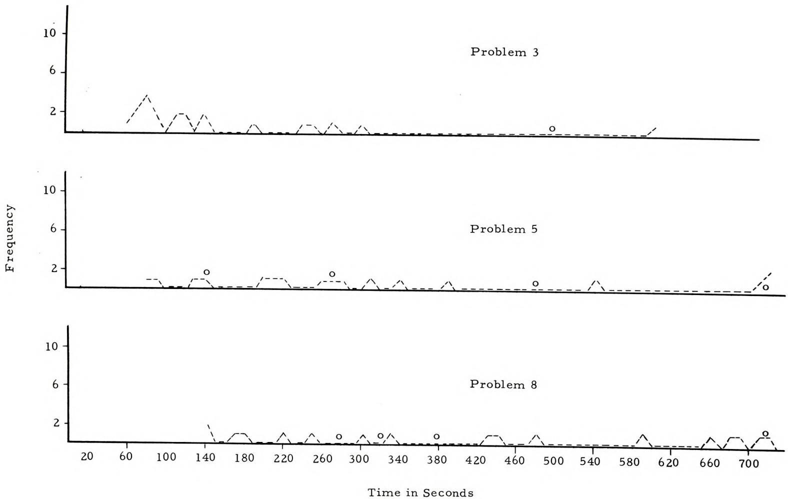

Figure 4. 2 - Broken lines represent the distributions of group solution times (solvers only). The superimposed circles indicate non-solving groups.

number of stages in the problem.

CLASSIFICATION: THE MODEL FOR THE DISTRIBUTION OF SOLUTION TIMES

Each problem was classified by its distribution of individual solution times. Only the distribution of solvers was used for this purpose; non-solvers (those who failed to solve or gave a wrong answer) were excluded from this analysis, since a person who fails to solve the problem may have blundered on any one of the k stages; it is not possible to determine from the data the stage on which this might have occurred. Consequently, data composed of persons who failed at different and unknown stages cannot be expected to provide an estimate of the number of stages in the problem.

Hypothetically, this distribution of solution times is a gamma distribution, and k is the number of stages in the problem. The parameter k was estimated by the method of moments, equation (16).

$$\hat{k} = \frac{(\bar{T})^2}{s_T^2} \quad .$$

An estimate of λ, the rate parameter, was obtained by the method of moments making use of equations (13 and

14),

$$\hat{\lambda} = \frac{\bar{T}}{s_T^2}$$

Table 4.3 shows the estimated number of stages and the estimates of rate of solution of each stage (λ) for each problem. As might be expected, $\hat{k}_i$ did not turn out to be an integer for any one of the three problems. Theoretically, k should be an integer, but it is not surprising that the estimates, which are subject to error, do not happen to be integers.

The obtained estimates of k for the three problems in order were similar to the intuitive estimates established from pre-experimental inspection of the problems. Moreover, the small number of stages estimated is in fair accord with intuition. For example the estimate of 4.47 stages (Table 4.3) for the Double problem is similar to the three stage pre-experimental speculation.

The sampling distribution of k is unknown under the estimation method used. To provide a check on the estimates, the theoretical curve, determined by a particular $\hat{k}_i$, was fitted to the observed distribution of solutions for each problem. The appropriate cumulative relative frequency curves are shown in Figures 4.3, 4.4, and 4.5. The obtained distributions are strikingly similar to gamma distributions, as

hypothesized. The theoretical curves were taken from tables of the incomplete gamma function (Pearson, 1922). Transformations of the time scale, etc., necessary for the use of the tables are discussed in Appendix 1.

Goodness of fit was tested by the Kolmogorov-Smirnov one-sample test. (See Walker and Lev, 1953; Siegel, 1956). For each problem, the statistical hypothesis tested was whether the sample of individual solvers could reasonably have been thought to come from a population having the theoretical distribution. This hypothesis was regarded as tenable for each problem. That is to say, the theoretical curves determined by $\hat{k}_3 = 2.8$, $\hat{k}_5 = 4.5$, and $\hat{k}_8 = 6.2$ gave acceptable fits to the data for problems 3, 5, and 8. See Table 4.4.

<u>Reading Time</u>. The experimental time interval for any one problem has heretofore been treated as a uniform block of time. However, one can distinguish between the relatively-invariant time required to read a problem and the highly variable time to solve it. It would be desirable to remove the reading time.

It was assumed as a rough estimate, that the first solver would solve the problem about 10 seconds after reading time was completed. This gave estimates of reading time for the three problems of 50, 40, and 90 seconds respectively. When these solution-times were subtracted from all individual time-scores, the

Table 4.3 Method of moments estimates of k and λ for each problem. (Individual solvers)

Problem	$\hat{k} = (\bar{T})^2 / s_T^2$	$\hat{\lambda} = \bar{T} / s_T^2$
3 (Rope)	2.83	.146
5 (Double)	4.47	.137
8 (Gold Dust)	6.21	.150

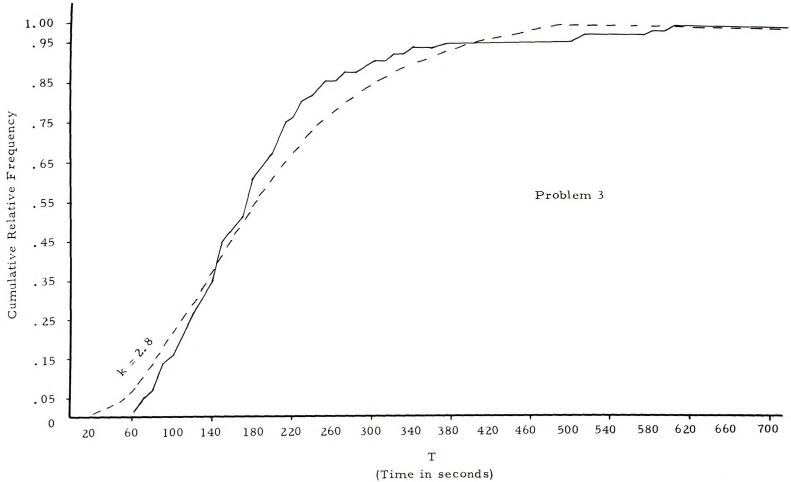

Figure 4.3 - Cumulative relative frequency of individual solvers on Problem 3 (solid line) and fitted gamma distribution with k = 2.8 (broken line).

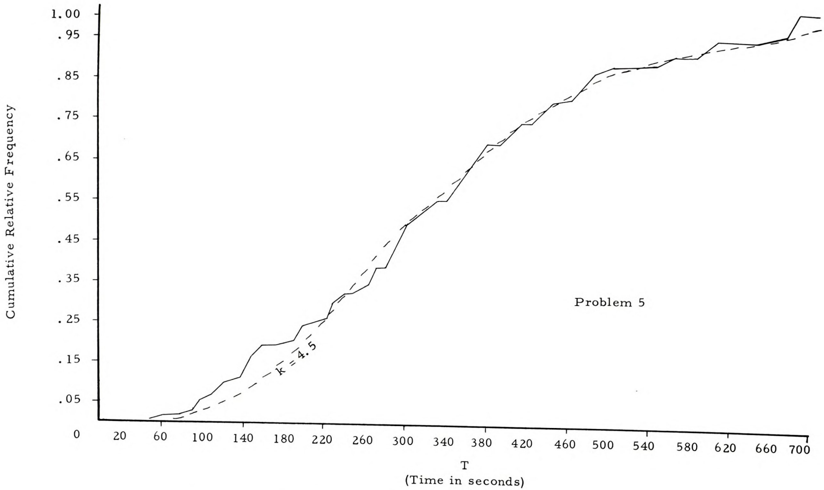

Figure 4.4 - Same as Figure 4.3, but for Problem 5. Fitted gamma distribution has k = 4.5. Note that distribution is more nearly symmetrical than in Figure 4.3.

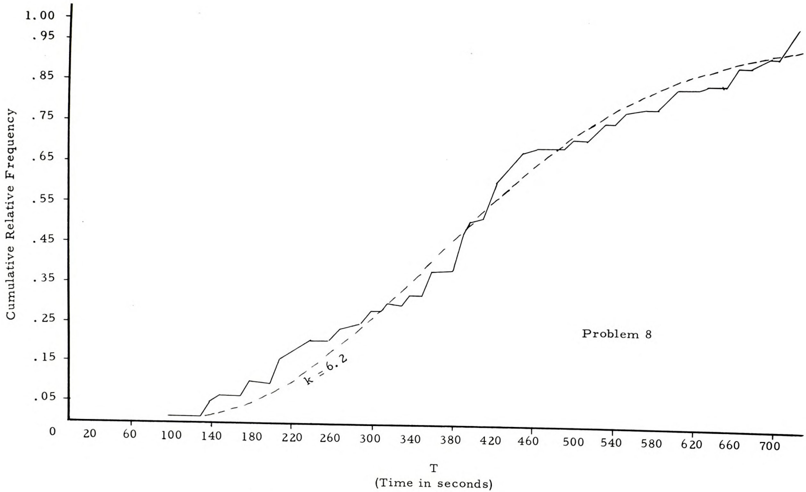

Figure 4.5 - Same as Figures 4.3 and 4.4; data for Problem 8, and gamma distribution with k = 6.2.

Table 4.4 Summary of Kolmogorov-Smirnov one-sample, goodness-of-fit tests for theoretical curves determined by estimates of k.

Problem	No. of Solvers	$\hat{k}$	D (maximum|theoret.-obs.|)	Sig. Values of D: .05 level
3 (Rope)	133	2.8	.110	.118
5 (Double)	89	4.5	.086	.144
8 (Gold Dust)	71	6.2	.082	.161

resulting net solution times also fit gamma distributions, with $\hat{k}' = 1.6$, 3.4, and 3.8 respectively, where $\hat{k}'$ is the estimate from net solution times. Note that this is a sharp drop from the estimates of k obtained before; 2.8, 4.5, and 6.2.

Subsequent to the completion of the main experiment, students in a regular class meeting of elementary psychology at Miami University (Ohio) were given the same three problems with essentially the same instructions, except that they were directed to record the time when they finished reading the problem, as well as when they had written an answer. The mean reading times for problems 3, 5, and 8, as recorded by the Miami students, were 62.7, 61.0, and 41.9 sec. (See Table 4.5).

The Miami University students were probably quite similar to the Michigan State University students, in reading time of such simple problems. If the Miami reading times are subtracted from the times-to-solution in Table 4.2 above, new estimates of the number of stages in these problems can be obtained. These new estimates, with reading time removed, are 1.3, 3.0, and 5.0 for problems 3, 5, and 8 respectively. The writer's judgments of number of stages were 1, 2 or 3, and 3 or 4 (see Chapter III above).

The Miami students were also asked to estimate the

number of stages in each of the problems, discounting reading time. The definition of a stage was essentially as given above in this dissertation, and the subjects made their judgments after having solved the problems in question. The mean judgments of number of stages, for problems, 3, 5, and 8, were 2.8, 4.0, and 6.3. These mean judgments differ from the best estimates (using the gamma-distribution theory) by 1.5, 1.0, and 1.3 stages, with the students' judgments being higher than the theoretical estimates, a very striking correlation. It should be borne in mind that the theoretical estimates come from the distributions of solution times, which were entirely unknown to the judges; and the theoretical estimates in no way used the judgments of the students. Hence, the two estimates were entirely independent and could have been completely unrelated.

Hereafter "solution time" will refer to net time, with reading time (estimated by the rough method of the original study) subtracted. Using the k'_i suggested by table 4.5, new theoretical functions were selected from Pearson's tables (1922) and fitted to the new distributions of individual solution times. By inspection of Figures 4.6, 4.7, and 4.8, it is evident that high agreement exists between the data and the theoretical curves. Results of the Kolmogorov-Smirnov one sample goodness to fit test appear in Table 4.6.

Table 4.5 Estimates of k and λ with reading time included and with reading time removed, where reading time was estimated by two different methods.

Reading Time Included			Reading Time Removed			
			Original Estimate removed		Second Estimade Removed	
Problem	$\hat{k}$	$\hat{\lambda}$	$\hat{k}'$	$\hat{\lambda}'$	$\hat{k}''$	$\hat{\lambda}''$
3	2.8	.146	1.6	.108	1.3	.099
5	4.5	.137	3.4	.120	3.0	.111
8	6.2	.150	3.8	.117	5.0	.134

PREDICTION: THE LORGE AND SOLOMON MODEL

As indicated in earlier discussion, the Lorge and Solomon Model does not deal with the question of the time interval allowed. In assessing the predictive usefullness of their model it was, therefore, necessary to apply it at various arbitrarily selected points in time.

The probability of an individual solution at every fifth time unit was taken to be the cumulative relative frequency of solutions by T = 50, 100, 150,..., 720. (Transformations of the time scale to take reading time into account were unnecessary since they do not affect the calculations.)

One first considers the case of a one-stage problem (Model A). Of the 178 individuals, .118 had solved Problem 3 by T = 100. Applying Model A (equation (18)) to this observed value of P_I', the predicted cumulative relative frequency for the group is

$$P_G' = 1 - (1 - .118)^4 = .395 \quad .$$

By T = 150 the probability of a correct solution had become .332 and the predicted group value is

$$P_G' = 1 - (1 - .332)^4 = .800 \quad .$$

The same procedure was repeated for all the points selected on the individual cumulative relative frequency distribution. (P_I' is the sample estimate of P_I).

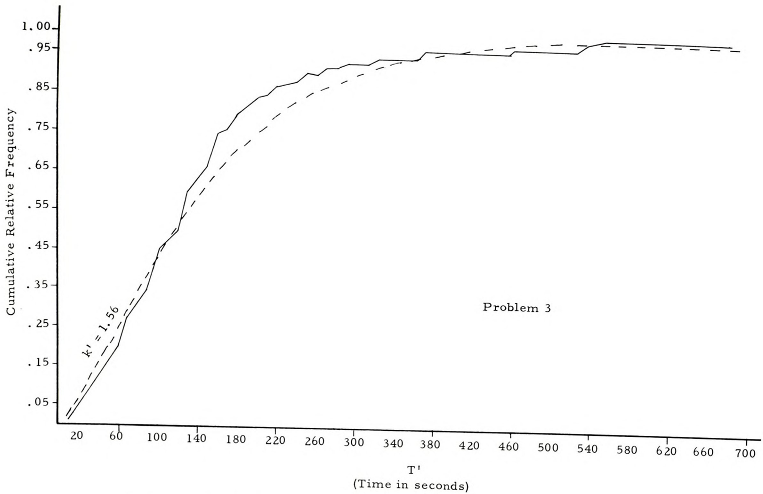

Figure 4.6 - The solid line represents cumulative relative frequency for individual solvers with reading time removed. The fitted theoretical curve ($k_3' = 1.56$) is described by a broken line.

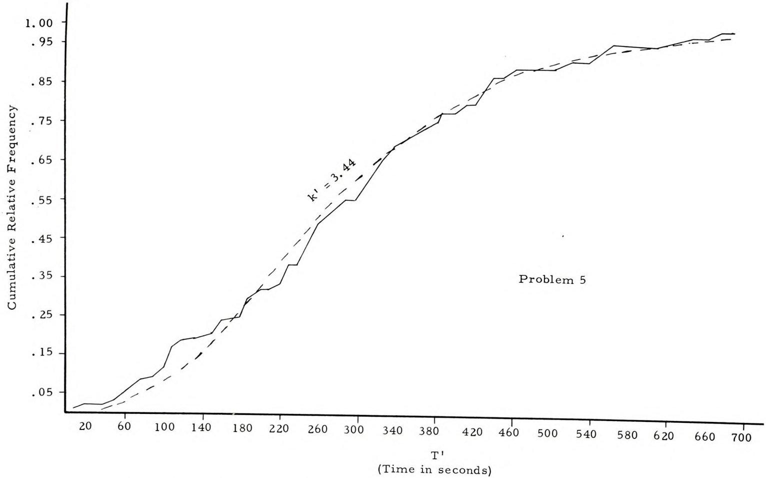

Figure 4.7 - The solid line represents cumulative relative frequency for individual solvers with reading time removed. The fitted theoretical curve ($k_5' = 3.44$) is described by a broken line.

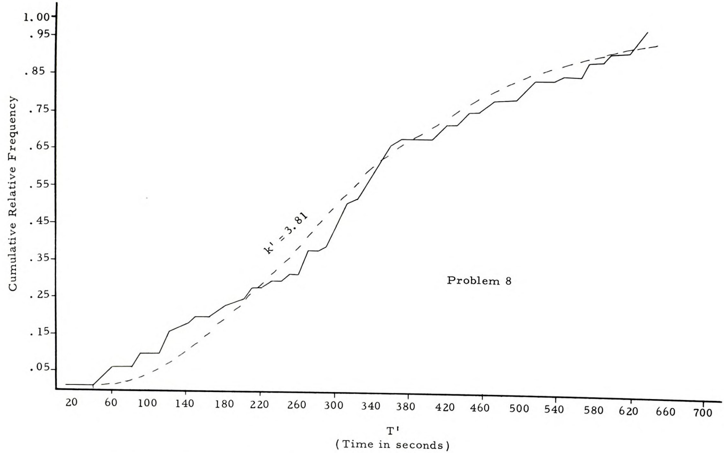

Figure 4.8 - The solid line represents cumulative relative frequency for individual solvers with reading time removed. The fitted theoretical curve ($k_8' = 3.81$) is described by a broken line. The subtraction of reading time has not altered the shapes of the theoretical curves for Problems 3, 5, and 8. But the origin in each case has been moved nearer to the first solution. (See Figures 4.3, 4.4, and 4.5).

Table 4.6 Summary of Kolmogorov-Smirnov goodness-of-fit tests for curves determined by $\hat{k}'_i$ (i = 3, 5, 8).

Problem	No. of Cases	$\hat{k}'$	D (maximum| theoret. obs. |)	Sig. Values of D: .05 level
3	133	1.6	.092	.118
5	89	3.4	.080	.144
8	71	3.8	.098	.161

Under the assumption of two equally probable stages, equation (21), Model B, was used to predict the group probabilities. For example, on Problem 3 the probability of an individual solution by T = 100 was taken to be .118 and by T = 150 this probability was .332, just as before. For k = 2, the group predictions were

$$P_G' = \left[1-(1-(.118)^{\frac{1}{2}})^4\right]^2 = .664 \text{ and}$$
$$P_G' = \left[1-(1-(.332)^{\frac{1}{2}})^4\right]^2 = .927 \quad .$$

Figures 4.9, 4.10, and 4.11 show the cumulative relative frequencies observed for groups and individuals (at every 50 seconds) along with the values predicted by the Lorge and Solomon Model (k = 1 and k = 2). Predictions of group performance are poor on each problem, and the predictions become worse with the larger value of k. Only for Problem 3 (The Rope Problem) would one regard as tenable the hypothesis that the observed distribution of group solution times was drawn from the theoretical distribution represented by the Lorge-Solomon predictions. On Problem 3, this null hypothesis was tenable for both k = 1 and k = 2. However, the predictions were everywhere uncomfortably larger than the observed proportions of group solutions. A summary of the Kolmogorov-Smirnov significance tests on each problem appears in Table 4.7.

The estimates of k from the classification section

are decidedly inappropriate with the Lorge-Solomon Model, regardless of their usefulness in choosing descriptive functions. Inspection of predicted and observed curves (Figures 4.9, 4.10, and 4.11) suggests that k must be taken less than one to provide good agreement between the two throughout the experimental interval. Any $k < 1$ is uninterpretable psychologically; so it is concluded that the Lorge-Solomon Model as applied here is unsuitable.

In essence, the Lorge-Solomon Model predicts the probability of a group solution to be the probability of solution by any of four persons working separately. Groups in this study were found to fall below this standard. That is, the real groups did not do as well as was expected on the basis of a simple pooling of individual results by means of the Lorge-Solomon Model.

PREDICTION: THE COMBINATION-OF-CONTRIBUTIONS MODELS

Two combination-of-contributions Models were proposed in Chapter II: (1) The Hierarchical Model; and (2) The Equalitarian Model.

The Hierarchical Model

Suppose that non-solvers are non-functional in their groups, as when the "solvers" in the group suppress the contributions of non-solvers and form a functional hierarchy. A group with two solvers and two non-solvers

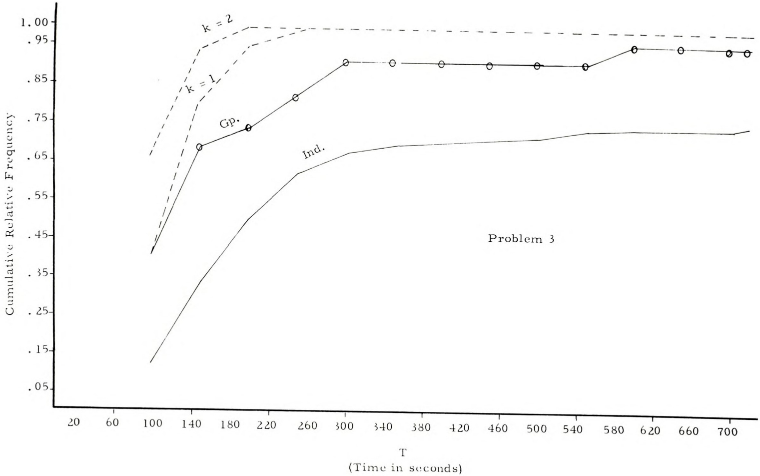

Figure 4.9 - The solid line represents the cumulative relative frequency of individual solutions. The solid line with open points is the same for groups of size 4. The broken lines are the cumulative relative frequency curves predicted by the Lorge and Solomon Model (k = 1, k = 2).

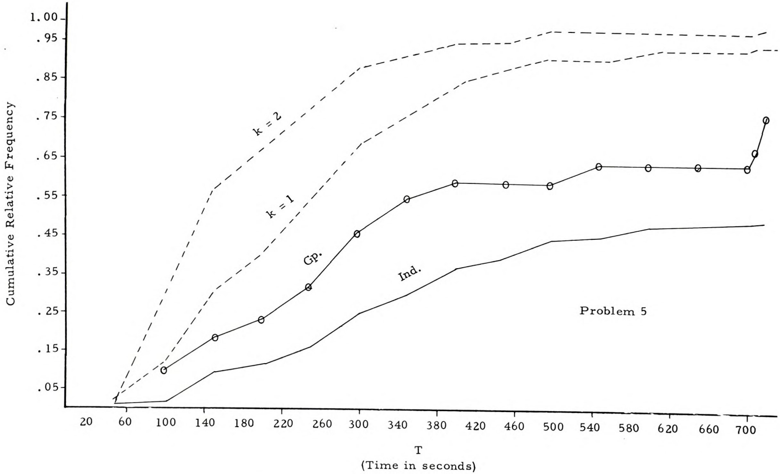

Figure 4.10 - The solid line represents the cumulative relative frequency of individual solutions. The solid line with open points is the same for groups of size 4. The broken lines are the cumulative relative frequency curves predicted by the Lorge and Solomon Model (k = 1, k = 2).

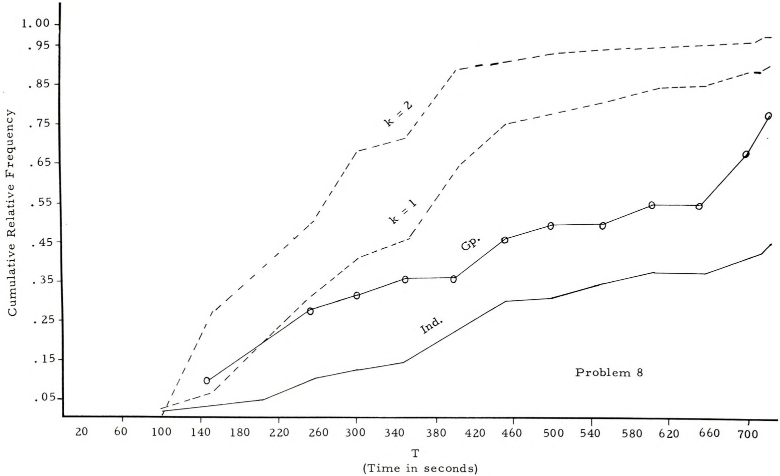

Figure 4.11 - The solid line represents the cumulative relative frequency of individual solutions. The solid line with open points is the same for groups of size 4. The broken lines are the cumulative relative frequency curves predicted by the Lorge and Solomon Model (k = 1, k = 2).

Table 4.7 Kolmogorov-Smirnov goodness-of-fit tests of predicted group solutions (Lorge-Solomon Model) with the observed data.

Problem	No. of Cases	D (maximum | theoret.-obs. |) k = 1	k = 2	Sig. Values of D for 25 df: .05 level	.01 level
3	22	.211	.245	.270	.320
5	22	.314	.385	.270	.320
8	22	.306	.521	.270	.320

would be expected to solve a problem just as rapidly as a group of two solvers alone. Then the probability of a group solution over time may be written as

(24) $$W_H(T') = \sum_{A=0}^{4} \left[Pr(A) \; g(T'; \; A\lambda, k)\right] \quad ,$$

where A is the number of solvers in the group. (The more general form was given in Chapter II, equation (22). Here T' time after reading is completed, has been substituted for T, and r = 4). Computationally more convenient is

(25) $$W_H(T') = \sum_{A=0}^{4} \left[Pr(A) \; g(AT'; \lambda, k)\right] \quad .$$

It is proved in Appendix 2 that the distribution $g(T'; \; A\lambda, k)$ is equal to $g(AT'; \lambda, k)$.

The excellent fits obtained with the classification model (Figures 4.6, 4.7, and 4.8) suggested that prediction may proceed directly from the obtained distribution of individual solvers, rather than going through the fitted gamma distribution, $g(AT'; \lambda, k)$. Let $h(T')$ by the obtained cumulative relative frequency distribution of individual solvers on a particular problem. Using the obtained data, the theoretical probabilities, $W_H(T')$, may be obtained from

(26) $$W_H(T') = \sum_{A=0}^{4} \left[Pr(A) \; h(AT')\right] \quad .$$

An example of the computations involved in the group predictions is given below. From Table 4.1, a = .735 of all subjects solved Problem 3 and .265 were

non-solvers. Equation (21) becomes

$$\Pr(A) = \sum_{A=0}^{4} \binom{4}{A} (.735)^{A} (.265)^{4-A}$$

Thus, for Problem 3 one has

(27) $W_H(T') = .005\ h\ (0T') + .052\ h\ (T') + .222\ h\ (2T') + .421\ h\ (3T') + .300\ h\ (4T').$

For example, when $T' = 70$,

$$W_H(70) =).005)\ (0) + (.052)\ (.271) + (.222)\ (.346) + (.421)\ (.850) + (.300)\ (.917) = .787\ .$$

Here, h(0), h(70), h(140), h(210), and h(280) are obtained from Figure 4.6, with $T' = T-50$. The curve of predicted group solutions under the Hierarchical Model was obtained by applying equation (27) at each time T'.

Similarly, for Problem 5, with $\underline{a} = .503$, h obtained from Figure 4.7, and $T' = T-40$,

(28) $W_H(T') = .063\ h\ (0T') + .250\ h\ (T') + .375\ h\ (2T') + .250\ h\ (3T') + .063\ h\ (4T')$;

and for Problem 8, with $\underline{a} = .436$, h from Figure 4.8, and $T' = T-90$,

(29) $W_H(T') = .099\ h\ (0T') + .308\ h\ (T') + .366\ h\ (2T') + .192\ h\ (3T') + .037\ h\ (4T')$.

The obtained cumulative relative frequency distributions of group solutions appear in Figures 4.12, 4.13,

and 4.14 for Problems 3, 5, and 8 respectively. The obtained distributions are solid lines, the predictions from the Hierarchical Model are dotted lines. Large discrepancies between the observed and predicted probabilities are evident. Kolmogorov-Smirnov one-sample, goodness of fit tests are summarized on Table 4.8.

The prediction was statistically tenable for Problem 3, but could be rejected confidently for Problems 5 and 8. These results indicate the Hierarchical Model to be unsuitable--at least with regard to Problem 5 and Problem 8. A comparison of Figures 4.9, 4.10, 4.11, with 4.12, 4.13, 4.14 suggests that the predictions under the Hierarchical Model are close to, and about as good as, those attained with the Lorge-Solomon Model A (k-1).

The Equalitarian Model

The Equalitarian Model says that non-solvers consume their share of the group's working time, though they do not contribute to the solution. That is, A solvers work on the problem; B non-solvers take their share of the group's time. Consequently, if A solvers are working, A/(A + B) is the rate at which they work.

The earlier statement of the Equalitarian Model (equation 23) may be rewritten as

$$W_E(T') = \sum_{A=0}^{4} \left[\Pr(A)\ g\left(\frac{A^2}{A+B}\ T';\ \lambda, k\right) \right] \tag{30}$$

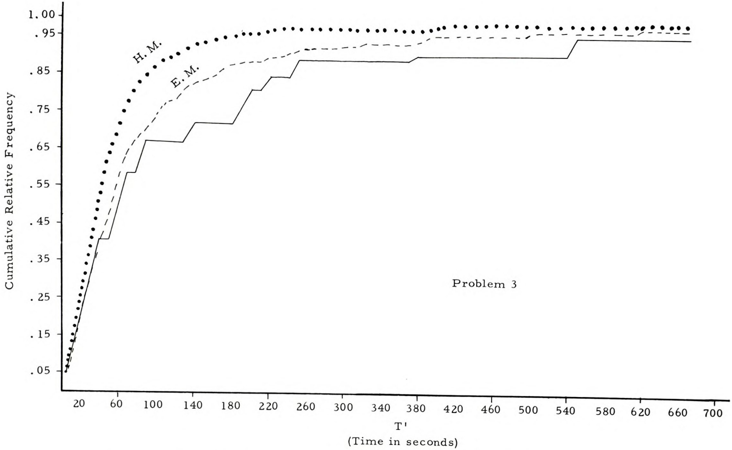

Figure 4.12 - Observed cumulative relative frequency of group solutions is shown as a solid line. Predictions under the Hierarchical Model are shown as a dotted line. Predictions under the Equalitarian Model are shown as a broken line. The Hierarchical Model provides an unacceptable fit to the observed group solutions; the Equalitarian Model provides acceptable predictions.

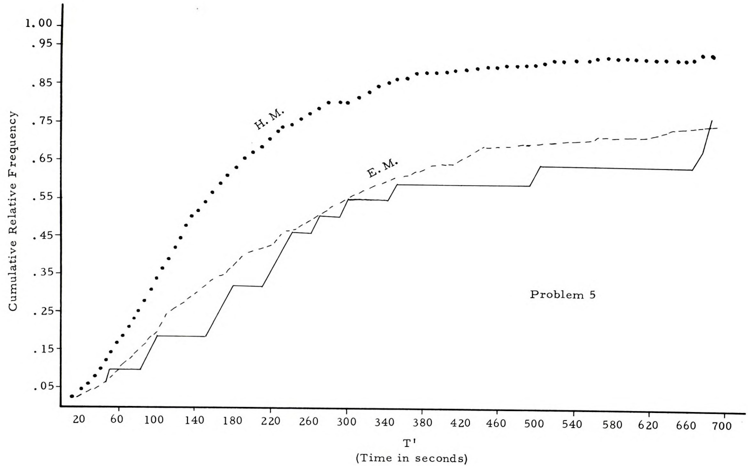

Figure 4.13 - Observed cumulative relative frequency of group solutions is shown as a solid line. Predictions under the Hierarchical Model are shown as a dotted line. Predictions under the Equalitarian Model are shown as a broken line. As with Problem 3, the Hierarchical Model does not successfully account for the distribution of group solutions.

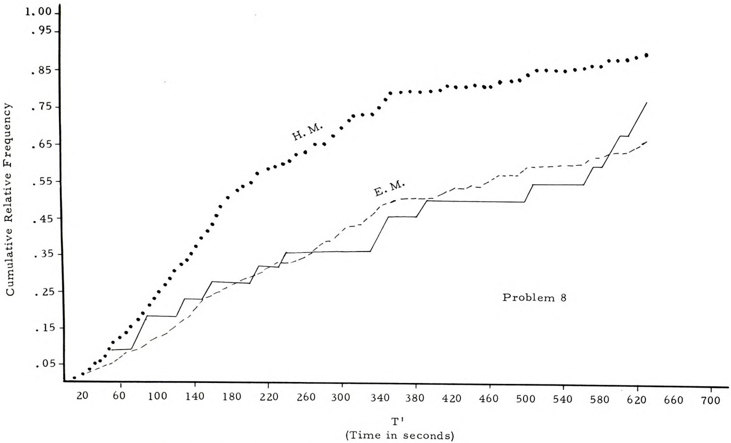

Figure 4.14 - Observed cumulative relative frequency of group solutions is shown as a solid line. Predictions under the Hierarchical Model are shown as a dotted line. Predictions under the Equalitarian Model are shown as a broken line. As with Problems 3 and 5, the Hierarchical Model yields predictions significantly different (too high) from observation. The Equalitarian Model agrees satisfactorily with the observed distribution of group solution.

Table 4.8 Kolmogorov-Smirnov one-sample, goodness-of-fit tests for observed and predicted group solutions under the Hierarchical Model.

Problem	No. of Cases	D (Maximum | theoret.-obs.|)	Sig. Values of D for 25 df: .05 level	.01 level
3	22	.242	.270	.320
5	22	.371	.270	.320
8	22	.379	.270	.320

which is convenient for computations. Again, predictions were made from the obtained distribution of individual solvers, $h_i(T')$, rather than the fitted gamma distribution, $g(T';\lambda,k)$. The predictions are

(31) $$W_E(T') = \sum_{A=0}^{4} \left[Pr(A)\ h\left(\frac{A^2}{A+B}\ T'\right)\right] .$$

Sample computations for the group predictions under the Equalitarian Model are given below. As in the case of the Hierarchical Model, one has for Problem 3

$$Pr(A) = \sum_{A=0}^{4} \binom{4}{A} (.735)^A (.265)^{4-A} ,$$

and thus

(32) $$W_E(T') = .005\ h\ (0T') + .052\ h\ (.25T') + .222\ h\ (T') + .421\ h\ (2.25T') + .300\ h\ (4T').$$

For example, when $T' = 70$

$$W_E(70) = (.005)\ ()) + (.052)\ (.038) + (.222)\ (.271) + (.421)\ (.737) + (.300)\ (.917) = .647.$$

The curve of predicted group solutions under the Equalitarian Model was obtained by applying equation (33) at each time, T'.

Similarly, for Problem 5 with h from Figure 4.7,

(33) $$W_E(T') = .063\ h\ (0T') + .250\ h\ (.25T') + .375\ h\ (T') + .250\ h\ (2.25T') + .063\ h\ (4T'),$$

and for Problem 8, with h from Figure 4.8,

(34) $$W_E(T') = .099\ h\ (0T') + .308\ h\ (.25T') + .366\ h\ (T') + .192\ h\ (2.25T' + .037\ h\ (4T') \quad .$$

Predictions from the Equalitarian Model were compared with the data from problem solving groups. The predicted distributions of group solutions appear in Figures 4.12, 4.13, and 4.14 as broken lines. On all three problems the correspondance between prediction and observation is striking; the predicted distributions of group solution times are very close to the observed distributions of group solution times. Kolmogorov-Smirnov goodness-of-fit tests provide no evidence against the hypothesis that the sample observations have been drawn from a theoretical population having the predicted distribution. See Table 4.9 for a summary of these significance tests.

These results agree with the proposition that non-solvers, those members of a group who are for some reason on the wrong track, take a full part in group discussion and consume their share of the group's time. The successful Equalitarian Model has a variety of implications about group structure. (a) All members of the group contribute about equally to the discussion, whether they are helpful or not. (b) The group does not form a hierarchical structure with a particular member

Table 4.9 Kolmogorov-Smirnov one sample, goodness-of-fit tests for observed and predicted group solutions under the Equalitarian Model.

Problem	No. of Cases	D (maximum | theoret.-obs. |)	Sig. Values of D for 25 df: .05 level	.01 level
3	22	.159	.270	.320
5	22	.138	.270	.320
8	22	.093	.270	.320

or members taking over, but is equalitarian. (c) The superiority of groups which solve over groups which fail to solve is explained merely on the basis of accidents of individual successes, not on the basis of the evolution of superior social organization.

The big question is whether non-solvers consume as much time as solvers in the group discussion. Even though the group process was observed in some detail, this question remains unanswered because the observer was unable to decide who the solvers and who the non-solvers were. A variety of other observations were made successfully, however, and can be used to test some of the consequences of the mathematical model.

EMERGENT SOCIAL STRUCTURE

The Equalitarian Model implies that the group members equally share the group's working time. In structural terms, the implication is that an equalitarian organization should emerge during the social interaction attending the attempt to solve the problem. The companion-choice data, gathered for each group after each problem were cast in the form of square matrices with binary entries--a 1 represented a choice and a 0 absence of choice. These 4 X 4 arrays are called "sociomatrices" by Glanzer and Glaser (1959). The three sets of post-problem choices generated three

sociomatrices for each group. A typical sociomatrix is given in Figure 4.15. Rows represent "choosers" and columns represent those chosen.

Each column of the sociomatrix was summed to give the total choices received by each group member. The analysis proceeded on the column totals. An estimate of the uncertainty, H, of each set of column totals was computed. Uncertainty, in information theory, is defined by the formula:

$$H = -\sum_{i-1}^{r} P_i \log_2 P_i \quad , \tag{35}$$

where an event has r possible outcomes, the i-th of which has a probability P_i, and i = 1, ..., r. (Miller, 1955). The sample estimate, $\hat{H}$, may be written as

$$\hat{H} = -\sum_{i=1}^{r} p_i \log_2 pi \quad , \tag{36}$$

where r is group size (4 in this case) and p_i is the obtained relative frequency of the i-th column. Uncertainty is at a maximum, 2.0 when the four categories are equiprobable (P_i = .250). Minimum uncertainty is, of course, H = 0.0. Thus for the present case $0.0<H<2.0$.

Generally speaking, the estimates $\hat{H}$ are biased. For example, suppose that the true probabilities $P_1 = P_2 = P_3 = P_4 = .25$, hence the true H = 2. If a small sample of observations are recorded, the values of the sample relative frequencies, p_1, p_2, p_3, and

chosen

chooser		A	B	C	D	
	A	0	1	0	1	
	B	1	0	0	0	
	C	1	0	0	0	
	D	1	1	0	0	
Σ		3	2	0	1	6
p		.500	.333	.000	.167	
-p log p		.5000	.5283	.0000	.4312	

$$\hat{H} = -\sum_{i=A}^{D} p_i \log_2 p_i$$

$$= .5000 + .5283 + .4312$$

$$\hat{H} = 1.4595 \text{ bits}$$

Figure 4.15 - A typical sociomatrix where 1 represents a choice and 0 is the absence of choice. The uncertainty computed from column totals is shown above right.

p_4 make up a multinomial distribution. In general, the sample relative frequencies will be unequal. These random inequalities of the p's will always reduce $\hat{H}$, with the result that all values of $\hat{H}$ are less than or equal to the true H = 2, and the mean $E(\hat{H}) < H$. With samples as small as those in the choice matrices studied here, the bias is considerable.

The four categories of the sociomatrix were considered equally probable under the hypothesis of an equalitarian choice structure. Given this hypothesis it was possible to ascertain the expected value of the estimated uncertainty for a particular grand total of choices, M, in a sociomatrix and a particular number of categories, r = 4. Making use of tables prepared by Rogers and Green (1955), the $E(\hat{H})$ was compared with $\hat{H}$ computed for the sociomatrix of each group, where $E(\hat{H})$ is the first moment of sample information (uncertainty) and $\hat{H}$ is the sample information as computed from (36).

Like the more traditional indices of group structure, the sampling distribution of $\hat{H}$ is unknown. Consequently, it was not possible to perform a significance test in connection with the difference $(\hat{H}-E(\hat{H}))$, noted for a particular matrix. (Unlike most indices of group structure, however, it was at least possible to decide what value of uncertainty

might be expected for a given number of choices and a given number of equiprobably categories). In the absence of a suitable probabilistic statement, the comparisons of $\hat{H}$ and $E(\hat{H})$ were scrutinized and the general correspondence was deemed satisfactory.

However, interest was mainly in the central tendency that characterized the structural properties of the experimental groups. Consequently, the mean of the various values of $\hat{H}$ was computed (i.e., $\bar{\hat{H}} = (1/n) \sum_{i=1}^{n} \hat{H}_i$), and the mean of the various values of E(H) as well (i.e., $\overline{E(\hat{H})} = (1/n) \sum_{i=1}^{n} E(\hat{H})_i$). These data are summarized in Table 4.10. Inspection of this table reveals rather striking similarities between the mean uncertainty and the mean of the uncertainties expected for each group under the equalitarian hypothesis. In fact, t-tests (solvers only) performed on the difference $(\bar{\hat{H}} - \overline{E(\hat{H})})$ for each of the three problems were not significant.

The inclusion of non-solving groups produced almost no change in the mean uncertainties. In fact the means of the non-solving groups behave in about the same way as the solving groups. Although the mean uncertainties of the non-solving groups are somewhat lower than those of the solving groups (on both problems 5 and 8), the differences do not appear so great as to lead one to speculate that non-solving groups

developed a different structure.

Additional support for the hypothesis of equalitarian structure was developed by considering the possible consequences should a non-equalitarian structure emerge from the problem-solving interaction. An equalitarian structure might emerge when the problem has a larger number of stages; more stages would, presumably, require more and diverse contributions from the members. Assume further that these talkers or contributors were recognized by others in the group. Important talkers or contributors would be reflected in the companion-choice structure; this structure would become more equalitarian as the number of stages increased.

To summarize, assume that successful problem solvers 1) talk more and 2) are recognized by others and become "leaders" or "stars" of the group. Both assumptions are, of course, denied by the Equalitarian Model. Inspection of the mean expectancies in Table 4.10, which are very similar for each of the three problems, does not suggest the emergence of a non-equalitarian structure as the number of stages increases. In the analysis of variance, reported in Table 4.11, the mean square for problems was not significant. In this analysis (performed on groups which solved all three problems), the criterion variable

Table 4.10 Mean uncertainty, $\bar{\hat{H}}$, of member companion choices--means and standard deviations for problems 3, 5, and 8.

Problem	Solvers				Solvers & Non-Solvers			
	n	$\bar{\hat{H}}$	$s_{\hat{H}}$	$\bar{E}(\hat{H})$	n	$\bar{\hat{H}}$	$s_{\hat{H}}$	$\bar{E}(\hat{H})$
3 (Rope)	21	1.6094	.3505	1.5482	22	1.6178	.3444	1.5487
5 (Double)	17	1.5542	.3991	1.5720	22	1.5417	.4042	1.5719
8 (Gold Dust)	17	1.4403	.4628	1.5457	22	1.4161	.4633	1.5388
	Non-Solvers							
3 (Rope)	1	---	---	---				
5 (Double)	5	1.4992	.5211	1.5718				
8 (Gold Dust)	5	1.3340	.5092	1.5152				

Tests of the hypothesis the population value is $\bar{E}(\hat{H})$ for solving groups.

Problem	t	df	Probability
3 (Rope)	0.800	20	<.50
5 (Double)	-0.184	16	<.50
8 (Gold Dust)	-0.939	16	<.40

was sample uncertainty, computed from the choice matrices, and the results are in accord with the equalitarian hypothesis.

Communication structure was analyzed in similar fashion to provide support for the dispersion assumption of the Equalitarian Model. As expected, a group member addressed most of his communications to the group at large rather than another member. It was possible to assign a set of talking frequencies, weighted for length of communication, to each group member. A communication structure was obtained for each group in this way. Each group's communication structure on each problem was summarized by computing the uncertainty, $\hat{H}$, for each set of four frequencies. Uncertainty was computed from equation (36), as before, after the frequencies had been converted to relative frequencies by dividing them by the total communications of all group members.

The equalitarian structure proposition may be restated in terms of intermember communication; uncertainty approaches 2.0 when members talked equally often, equally long. Conversely, if one member monopolizes the conversation, $\hat{H}$ is near 0.0.

As before, the first moment of the sampling distribution of $\hat{H}$ was approximated for each set of $r = 4$ equiprobable categories and the group total of M

Table 4.11 Analysis of variance where criterion variable is sample uncertainty, computed from the sociomatrices of choices.

Source	Sum of Squares	df	Mean Square	F
Groups	2.4849	13	.1912	1.996 N.S.
Problems	.1185	2	.0593	--
Error	2.4902	26	.0958	
Total	5.0936	41		

Analysis of variance where criterion variable is sample communication uncertainty.

Source	Sum of Squares	df	Mean Square	F
Groups	.4854	13	.0373	1.719 N.S.
Problems	.0515	2	.0258	1.189 N.S.
Error	.5653	26	.0217	
Total	1.1022	41		

communications. The Rogers and Green (1955) tables for $E(\hat{H})$ do not extend beyond the case where M = 10. It was, therefore, necessary to make use of the Miller-Madow approximation presented by Rogers and Green. The formula was as follows:

$$(37) \quad E(\hat{H}) = \log_2 r - \frac{(\log_2 e)(r^2 + 6M(r-1)-1)}{12M^2}, \quad M > r .$$

Again, ignorance of the sampling distribution of $\hat{H}$ precluded testing the difference, $\hat{H}-E(\hat{H})$ for significance in each of the experimental groups. The difference of central interest, however, was $\bar{\hat{H}} - \bar{E}(\hat{H})$. Table 4.1 summarizes these data for communication structure. Significant differences (t-tests) were found on each of the three problems between the observed mean uncertainty of solving groups, $\bar{\hat{H}}$, and the hypothesized value. Unfortunately, it was not possible, experimentally, to distinguish between solvers and non-solvers and one may only conclude that some people talked more than others. This by no means controverts the idea that non-solvers talk as much as solvers; the result is neutral on that issue. However, the (in fact) quite high estimates of H for communication indicate that contributions were actually widely spread. To make a more complete statement one would have to know something about individual differences. Concentrate for the moment on the absolute

Table 4.12 Uncertainty, $\bar{\hat{H}}$, of member communications--means and standard deviations for Problems 3, 5, and 8.

Problem	Solvers				Solvers & Non-Solvers			
	N	$\bar{\hat{H}}$	$s_{\hat{H}}$	$\overline{E(\hat{H})}$	N	$\bar{\hat{H}}$	$s_{\hat{H}}$	$\overline{E(\hat{H})}$
3 (Rope)	21	1.8085	.1601	1.9945	22	1.8145	.1586	1.9946
5 (Double)	17	1.7951	.1755	1.9949	22	1.8134	.1774	1.9952
8 (Gold Dust)	17	1.7848	.1585	1.9956	22	1.7714	.1685	1.9957

Problem	Non-Solvers			
3 (Rope)	1	----	---	---
5 (Double)	5	1.8755	.1888	1.9965
8 (Gold Dust)	5	1.7256	.2133	1.9962

Test of the hypothesis the population value is $\overline{E(\hat{H})}$ for solving groups.

Problem	t	df	Probability
3 (Rope)	-5.296	20	.001
5 (Double)	-4.694	16	.001
8 (Gold Dust)	-5 .484	16	.001

value of mean H; if all four subjects talk with equal frequency, H = 2.0. Note that if three-and one-half subjects talked with equal frequency, one would have the uncertainty 1.80 which is about the $\bar{\hat{H}}$ (communication) observed on each of the three problems. This is close to no organization, regardless of the decision required by the statistical test.

As before, (p. 81, paragraphs 1 and 2) comparisons among the mean communication uncertainties, $\bar{\hat{H}}$, on the three problems revealed them to be quite similar. In the analysis of variance reported in the bottom half of Table 4.11, the mean square for problems was not significant. In contrast to the just preceding, indecisive analysis of communication structure, the analysis of variance was interpreted as offering further indirect support of the equalitarian structure proposition.

SUMMARY OF RESULTS

The major points of this chapter may be summarized as follows:

1) It was found that word problems could be classified, without resorting to subjective decisions, by means of rational curve-fitting in accord with the mathematical model. Gamma distributions were successfully fitted to distributions of solution times.

Estimates of the number of stages in each problem were close to judgments of number of stages obtained in subsidiary observations.

2) A mathematical model was developed for predicting the distributions of group solution times from the distributions of individual solution-times; the theoretical predictions (the Equalitarian Model) were very close to the obtained data on all three of the problems employed.

3) Observations of the groups' social interaction and emergent structure were found to be in general accord with subsidiary assumptions of the successful Equalitarian Model.

4) Group performance was inferior to individual performance when the latter was pooled according to the non-interactional model of Lorge and Solomon.

CHAPTER V

DISCUSSION AND CONCLUSIONS

Suppose that the mean times to completion of two processes are equal at, say, 10 minutes. Imagine the first process yields highly variable data with a standard deviation of 10 minutes and a J-shaped distribution. This would be interpreted as a one-stage process in which the probability of the event is very low, the rate parameter being 0.1 per minute. Suppose that the other process has a mean of 10 minutes but much smaller variance, having a standard deviation equal to 1 minute and a more nearly symmetrical, bell-shaped distribution. One would interpret this as a 100-stage process, each stage of which has a rate (probability per unit time) 100 times greater than in the first process. Thus, while solving a eureka problem and reading an essay may take the same (average) time, the problem solving process, having one or two low-probability stages, would yield very variable data; and the reading, being made up of a great many high-probability stages, would yield much less variable data. These general relationships seem intuitively compelling, and the waiting-time theory used in this study is merely a mathematical idealization of this line of argument.

CLASSIFICATION MODEL

The mathematical model takes advantage of a great deal of information; the distribution of individual solution times as well as the proportion of right and wrong answers. The theoretical model begins with simple assumptions; that the probability of solving a stage is a constant over time until solution, and that the times to solve the several stages of a problem are independent and have equal means. From these assumptions, and the concept of continuous time, precise logical argument leads to the gamma distribution. The derivation is, in fact, a fairly routine application of the methods of probability theory.

This model had two empirical hurdles to surmount. First, it claimed that times to solution would have a gamma distribution. The obtained distributions were close to the fitted gamma and did not differ in any significant way. Second, it was possible to estimate k, the number of equally-difficult stages, for each of the three problems. The writer's intention and intuition was that the three problems differ in number of stages, and have approximately 1, 2 or 3, and 3 or 4 stages, respectively. Miami University students judged that the problems would have about 3, 4, and 6 stages. The best estimates from the data are 1.3, 3.0, and 5.0, which progress in the direction intended

and are of about the magnitude expected. Thus, the estimates of k from the data (depending on the squared mean divided by the variance) have some apparent validity. Only further investigations of stages in problem-solving will determine whether this is an accident. At present, except for an intuitive judgment of the number of distinct "ideas" required to obtain the solution, there is no adequate description of the stages involved.

The number of stages appears to be the number of distinct ideas which must be <u>conjoined</u> to obtain the correct answer. The subject must have idea 1 <u>and</u> idea 2 <u>and</u> so forth, if he is to complete stage 1 and stage 2, etc. If there are several answers to the problem, so that the subject may have idea 1 <u>or</u> idea 2 <u>or</u> ... and so forth, this should make the problem easier, increasing the rate λ. This <u>disjunction</u> might very well correspond to a single stage. Thus, to some degree, the notion of stages is related to the logical requirements of the problem, at least to the logical interrelationships of the ideas involved. This does not yield a complete description of the hypothetical process, however. Given any set of ideas one could specify the various compounds which would solve the problem. However, we do not know for sure what ideas the subjects entertained, so that we know the logical

structure and not the elements embedded in that structure. Only further empirical separation of stages can fulfill the final requirement.

While the concept of "number of stages," k, has some value in beginning an analysis of problem-solving, it should be emphasized that the number of stages estimated by the present method is only an approximation. One would expect that in most problems with several stages, the stages differ in difficulty and have different values of λ. Given any set of stages and their λ-values one can construct the distribution of total times. However, any empirically-determined distribution of total times can be closely approximated by various combinations of stages and λ-values. The only workable <u>unique</u> answer to the question, How many stages does this problem involve? is attained by first assuming that all the component stages have equal values of λ. The resulting estimate is something like the information-theory idea that the information transmitted through a given system, if it is 2 bits, is "equivalent" to a system which can discriminate exactly four equiprobable categories. The system may actually be discriminating nine unequally probable categories, and the fact that information transmitted equals 2 bits does not reveal the fine grain. In a similar way, the estimate of 3 stages says that the subject's performance is

"equivalent" to three equiprobable stages, though it might actually consist of 5 or 6 stages which differ widely in probability or rate, λ.

Experimental separation of stages is needed before definite conclusions can be drawn. A first step in that direction was taken in the follow-up study, in which reading-times were estimated and subtracted from problem-solving data. Intuitively, reading the problem is a multi-stage process in which each stage has a high value of λ. Compared with problem-solving, reading the problem consumes a relatively constant amount of time. Similarly, it may be presumed that time to write the solution is relatively constant, and since it can easily be measured it should be separated from pure problem-solving time. Possibly, by the use of protocols, it is possible to determine when the subject overcomes various stages of the problem, such as "understanding" the problem, obtaining an hypothesis, etc. Such logical or psychological analysis may lead to an understanding of the stages of problem-solving, but each such analysis should be adopted with caution. The data of this study suggest that some problems involve compounds of "part problems", and that each part (even though it may involve formulation of the problem, discovery of an hypothesis, devising a test of the hypothesis, etc.) is mastered in a single, all-

or-nothing stage. This view is supported by the close correlation of estimated number of stages with what appear to be the number of parts into which the problem can be divided; particularly, it is interesting to note that the gold-dust problem has answers which involve about six measurements, and is estimated to involve about six stages.

The method of classifying problems according to the number of stages involved can, in principle, be applied to other than Eureka problems. For example, a geometrical proof might involve roughly as many stages as there are steps in the proof. A cross-word puzzle might involve as many stages as there are difficult words to discover. A task which yields relatively invariable times, such as manipulating a large number of blocks in a routine way, would presumably have a very large number of stages. One way of investigating the present model is by studying tasks which, prima facie, differ widely in number of stages.

This discussion has resolutely used the phrase "number of stages" in place of the more natural adjective, "complexity." In the model, the difficulty of a problem (that is, for example, mean time to solution) depends upon both k, the number of stages, and λ, the probability of accomplishing a given stage per unit time. The mean is, in fact, $k(1/\lambda)$. The

term "complexity" refers to both of these factors, for a problem may be complex because it involves many stages to completion, or complex because, at the one stage required, there are many alternative possibilities. The number of wrong alternatives does not affect k, but when there are many alternatives the probability of the correct one is presumably low, hence λ would be low.

Though the present data give no such indication, it is plausible to suppose that the emergent structure of a permanent problem-solving group might be affected by the kind of problems they work on. If each problem is mainly of a single stage, the group might develop a hierarchical structure, attempting to let the best problem-solver remain in control. Multi-stage problems, on the other hand, would tend to favor an equalitarian structure so that various members can contribute to the solution.

GROUPS VS. INDIVIDUALS

Although the present data are more extensive, and analyzed in more detail, than previous results, the general relationship between individual and group performance agrees well with earlier findings. Groups outperformed individuals, as has been almost universally true in earlier studies. Furthermore, the groups were less efficient than would be expected by pooling the

abilities or accomplishments of individuals. The observation that actual groups lie between individuals alone and pooled individuals agrees with results of Husband (1940) and Taylor and Faust (1952), and the observation of Watson (1928) though Faust (1959) found a slight superiority of real groups over pooled individuals. The present data indicate that real groups, while as sure as pooled individuals, are comparatively slow. In Figs. 4.9, 4.10, and 4.11, in which group data are compared with Lorge-Solomon predictions, the largest discrepancy is at some intermediate time, roughly at the mean time-to-solution. With small numbers of groups it is possible that various results would occur depending on the total time given the subjects; the usual procedure gives only one point of the cumulative distribution.

In the present study, groups failed to reach the level predicted by the Lorge-Solomon pooling model, or by the Hierarchical Model which is, in effect, the continuous analog of the pooling model. The interpretation offered, which is embedded in the assumptions of the Equalitarian Model, is that group members who are on the wrong track consume as much group time as do members who are approaching solution. The pooling models pool the abilities or accomplishments but do not pool the disabilities and errors of the members in the

group. That is, the Lorge-Solomon model combines the wheat and leaves behind the chaff. The Equalitarian model which fits the present data, pools both wheat and chaff.

In a sense, the Lorge-Solomon model is impossible, for until the problem is solved no member of the group knows who is on the right and who on the wrong track. Before solution there is no possible mechanism within the group for singling out the wheat from the chaff. The Lorge-Solomon model, and the Hierarchical Model formulated here, both set an unrealistic standard for groups.

The data and the Equalitarian Model indicate that, in terms of man-hours, it is cheaper to have individuals solve problems separately than to use groups. Four individuals in separate rooms will, on the average, solve a problem sooner than the same four individuals in a face to face group. On this criterion, groups are clearly inefficient problem-solvers. Yet this conclusion is seriously misleading.

First, the Equalitarian Model differs from the Hierarchical Model only when the probability of solution, $\underline{a}$, is less than one. If all subjects would solve the problem, there would be no difference between the models. What is more, in that case a group of size r would have a rate of $r\lambda$, if individuals have a rate

of λ; whence groups would be precisely as efficient as individuals. It is only the probability of an individual failing to solve the problem, which slows down the group average. In the case of individual failures, the group has an advantage over individuals quite aside from speed--for the group yields a single answer, and the individuals working separately may offer several different answers. It seems improbable that a group of size 4, for example, could attain agreement on a wrong answer in problems like those use in this study. Hence it is very probable that the group will yield the correct answer. Individuals, on the same problem, may have a considerable probability of giving a wrong answer.

In an experiment, one tends to ignore the individual failures for the experimenter knows the right answer to the problem. In any real life application, there would be no experimenter with the correct answer written down. If there were, there would be no need for either the individuals in separate rooms or the group.

Though the individuals will, on the average, obtain a correct answer before a group, they may also obtain some wrong answers. Working separately, four individuals would not separate the correct from the wrong answers, and as a result the first individual

solution (which on the average will be before the group solution) cannot be accepted with any confidence. The group solution has a probability $1 - (1-a)^r$ of being correct, when a is the probability that an individual is correct and the group is of size r. The first individual solution has only probability a of being correct. In this sense the group is slower but surer.

It should be remarked that the present model makes sense only if the problem is one of searching for a unique correct answer. In a "human relations" problem (Lorge, et al., 1958) the task may be to assemble a large number of ideas none of which is the "answer." In such a task, mutual stimulation may be an important advantage, whereas the drawback of groups, that non-solvers take up as much time as solvers, may not be found if virtually all ideas are accepted as valuable.

Groups may have advantages, in real life, which are not reflected in the experimental conditions studied here. For example, it is possible that a group will persevere longer on a problem than the individuals would separately, thus enhancing the probability of an eventual solution. Work in a group may be conducive to higher or more appropriate motivation than work alone. The exigencies of group management may lead to such benefits as regular work hours, orderly arrangement of

materials, etc., which would give group workers an advantage over individuals. By the same reasoning, of course, members of groups may have the disadvantage of working at other than their own best time, of struggling with an inappropriate organization of material, etc. Such factors are probably more characteristic of stable than of the ad hoc groups discussed in this study, and there is no reason to believe that the simple models employed here would apply to very different situations with new properties.

The observations of groups in this study agreed quite well with the supposition that they were organized in an equalitarian structure. The success of the Equalitarian Model suggests that, at any rate, members who are on the wrong track contribute their share to the discussions. The data on choice of future partners failed to reveal any noticeable trace of stable leadership within groups. An equalitarian structure is, as was discussed above, relatively slow and sure for it devotes time to members who contribute only confusion, but at the same time it provides a check against error. An hierarchical structure should, according to the present analysis, increase speed but also increase the probability of accepting a wrong solution. The conditions under which a hierarchy would arise may be at least imagined; if the group were under

extreme time pressure and could not afford to "suffer fools gladly," it might be induced to adopt a hierarchical structure. The experimental conditions employed here, giving enough time for almost all individuals to solve the problems or arrive at a satisfying error, are clearly not those which would most probably lead to a hierarchy.

THE EFFECTS OF INDIVIDUAL DIFFERENCES

If there are large and stable individual differences between subjects in their abilities to solve these problems, then the number of stages in the problems is underestimated. Individual differences would inflate the variances of time scores, and since k is estimated by the squared mean divided by the variance, this would decrease the estimate of k below what it should be.

Large and stable differences should, however, produce discernable hierarchical structures in groups, and should lead groups to prefer to choose one bright member over the others. The lack of any such hierarchy suggests that any differences between subjects were imperceptible to the subjects themselves.

Furthermore, an investigation was made of solvers and non-solvers among the individuals working on the three problems. While the form of the data makes it

difficult to perform a suitable statistical test, the indication was that there was little correlation between success on the three problems. This is not unexpected when it is noted that the subjects were all drawn from an elementary psychology course and were thus of a restricted range of talent, and when it is recalled that the present three-problem experiment would constitute a very short and unreliable intelligence test.

CONCLUSIONS

It was found that the distributions of solution times on three problems were adequately described by gamma distributions. Gamma distributions arise from a rational consideration of the stochastic theory of waiting times. Estimates of the number of equally-difficult stages involved in the three problems, agreed in substance with the intentions of the experimenter and with judgments by a group of students. Furthermore, inspection of the logical structure of the problems suggested that the number of stages is close to the number of ideas which must be conjoined to arrive at the solution.

Groups of four subjects solved a significantly higher proportion of problems than did individuals. However, groups which solved took about as long as

individuals who solved. It was found that group performance could not be predicted accurately from individual performance on the assumption that accomplishments were simply pooled. If, however, it is also assumed that group members on the wrong track consume their share of group time, so that both accomplishments and failures of the individuals are pooled, the entire distribution of group solution-times could be predicted accurately.

The Equalitarian Model, which accurately predicted group performance, led to the expectation that there would be no hierarchical structures within groups. Sociometric choices made after each problem supported this hypothesis, since they were in agreement with an hypothesis of random, equiprobable choice between group members.

APPENDIX I

Use of Pearson's Tables of the Incomplete Γ Function

In this study $\Gamma(k) = (k-1)!$ (k an integer).

Let $\lambda T = x$. If

(1) $$\Gamma(k) = \int_0^\infty e^{-x} x^{k-1} dx$$

(see Mood, 1950), let

(2) $$\Gamma_x(k) = \int_0^x e^{-x} x^{k-1} dx \quad .$$

The ratio of these two functions is the density

(3) $$\frac{\Gamma_x(k)}{\Gamma(k)} \quad .$$

Pearson (1922) defines

(4) $$\Gamma_x(p+1) = \int_0^x e^{-x} x^p dx \quad .$$

In this notation, equation (3) may be rewritten as

(5) $$\frac{\Gamma_x(p)}{\Gamma(p)} \quad .$$

Pearson defines

(6) $$I(x, p) = \frac{\Gamma_x(p+1)}{\Gamma(p+1)} \quad .$$

Due to computational difficulties, Pearson tables not I(x,p) but the auxillary function I(u,p) where $u = x/(p+1)^{\frac{1}{2}}$ or $x = u(p+1)^{\frac{1}{2}}$

If $p + 1$ is taken equal to k, then by substitution $u = x/k^{\frac{1}{2}}$.

Furthermore $p = k - 1$.

Recall $x = \lambda T$. Hence

$$u = x/k^{\frac{1}{2}} = T\frac{\lambda}{k^{\frac{1}{2}}} \tag{7}$$

Equation (7) takes T into u, and it is the u argument which has been tabled extensively by Pearson (1922).

However, the transformation to u is facilitated by noting that $k = \left[E(T)\right]^2/\sigma^2(T)$ and $\lambda = E(T)/\sigma^2(T)$. Substituting these expressions in (7)

$$u = T\left[\frac{\frac{E(T)}{\sigma^2(T)}}{\left(\frac{\left[E(T)\right]^2}{\sigma^2(T)}\right)^{\frac{1}{2}}}\right] = \frac{T}{\sigma(T)} \quad . \tag{8}$$

Using the sample value s, equation (8) may be written as

$$u = T/s \; .$$

The tables are entered with appropriate u and $p = \hat{k}-1$, where $\hat{k}$ is the sample estimate of k.

APPENDIX 2

Proof that $g(T;\alpha\lambda,k) = g(\alpha T;\lambda,k)$

The function $W_i(T) = \sum_{A=0}^{4} \left[Pr(A)\ g(T;\alpha\lambda,k)\right]$ (generalized form) is not convenient for computational purposes. In the second term of $W_i(T)$ it is more advantageous for the coefficient α to appear with the variable T rather than the parameter λ. Therefore, it is necessary to prove:

(1) $$g(T;\alpha\lambda,k) = g(\alpha T;\lambda,k) \quad .$$

Recall that

(2) $$g(T;\lambda,k) = \frac{\lambda}{(k-1)!} e^{-\lambda T}(\lambda T)^{k-1} .$$

Thus,

(3) $$g(T;\alpha\lambda,k) = \frac{\alpha\lambda}{(k-1)!} e^{-\alpha\lambda T}(\alpha\lambda T)^{k-1}$$

and

(4) $$g(\alpha T;\lambda,k) = \frac{\lambda}{(k-1)!} e^{-\alpha\lambda T}(\alpha\lambda T)^{k-1} .$$

In order to abbreviate notation, let $G(T) = G(T;\alpha\lambda,k)$ be the cumulative distribution of $g(T;\alpha\lambda,k)$ and $H(T)$ be the cumulative distribution of $g(\alpha T;\lambda,k) = q(T)$. It is sufficient to prove that $G(T) = H(T)$.

It follows that

(5) $$\frac{d\ G(T)}{dT} = \alpha . q(T) \text{ and } \frac{d\ H(T)}{d(\alpha T)} = q(T) .$$

Let $\alpha T = z$; then

(6) $$\frac{d \quad G(T)}{dT} = \alpha\left[\frac{d \quad H(T)}{dz}\right].$$

Multiplying both sides of (6) by $1/\alpha$,

(7) $$(1/\alpha)\left[\frac{d \quad G(T)}{dT}\right] = \frac{d \quad H(T)}{dz}$$

Obviously $G(T) = H(T)$ if $\frac{d \quad G(T)}{dT} = \frac{d \quad H(T)}{dT}$, since $G(0) = H(0) = 0$ and $G(\infty) = H(\infty) = 1$.

Now one may write

(8) $$\frac{d \quad H(T)}{dT} = \frac{d \quad H(T)}{dz} \cdot \frac{dz}{dT},$$

but from (7)

$$\frac{d \quad H(T)}{dz} = (1/\alpha)\left[\frac{d \quad G(T)}{dT}\right].$$

Hence,

(9) $$\frac{d \quad H(T)}{dT} = (1/\alpha)\left[\frac{d \quad G(T)}{dT}\right] \cdot \frac{dz}{dT}$$

Since $z = \alpha T$, it follows that

(10) $$\frac{dz}{dT} = \alpha.$$

Substituting (10) in (9), one has

(11) $$\frac{d \quad H(T)}{dT} = (1/\alpha)\left[\frac{d \quad G(T)}{dT}\right]\alpha = \frac{d \quad G(T)}{dT}.$$ Q.E.D.

REFERENCES

Augenstine, L. The use of ILLIAC in determining distributions for information functionals. In H. Quastler (Ed.), _Information theory in psychology_. Glencoe, Ill.: Free Press, 1955.

Battig, W. F. Some factors affecting performance on a word-formation problem. _J. exp. Psychol._, 1957, 96-104.

Bavelas, A. A mathematical model for group structures. _Appl. Anthrop._, 1948, 16-30.

Benedetti, D. T. A situational determiner of the Einstellung-effect. _J. gen. Psychol._, 1956, 271-278.

Bush, R. R., and Mosteller, F. A stochastic model with applications to learning. _Annals of math. Stat._, 1953, 559-585.

Bush, R. R., and Mosteller, F. _Stochastic models for learning_. New York: Wiley, 1955.

Cartwright, D., and Zander, A. _Group dynamics_. Evanston, Ill.: Row, Peterson, 1953.

Cobb, H. V., and Brenneise, S. H. Solutions of the Meier string problem as a function of the method of problem presentation. _Proc. S. Dak. Acad. Sci_, 1952, 138-142. In Duncan, C. P. Recent research on human problem solving. _Psychol. Bull._, 1959, 397-429.

Dudeney, H. H. _The Canterbury Puzzles_. New York: Dover Publications, 1958.

Duncan, C. P. Recent research on human problem solving. _Psychol. Bull._, 56, 1959, 397-429.

Ekman, Gosta. The four effects of cooperation. _J. of soc. Psychol._, 1955, 149-162.

Faust, W. L. Group versus individual problem solving. _J. abnorm. soc. Psychol._, 1959, 68-72.

Feller, W. _An introduction to probability theory and its applications_. Vol. 1, New York: Wiley, 1957.

Fisher, R. A. _Contributions to mathematical statistics._ New York: Wiley, 1950.

Gibb, E. G. Children's thinking in the process of subtraction. J. exp. Educ., 1956, 71-80.

Glanzer, M., and Glaser, R. Techniques for the study of team structure and behavior Part I: analysis of structure. Tech. Report, Amer. Instit. for Research, 1957a.

Glanzer, M. and Glaser, R. Techniques for the study of team structure and behavior Part II: empirical studies of the effects of structure. Tech. Report, Amer. Instit. for Research, 1957b.

Glanzer, M. and Glaser, R. Techniques for the study of group structure and behavior: I. Analysis of structure. Psychol. Bull., 56, 5, 1959, 317-332.

Hays, D. G. and Bush, R. R. A study of group acting. Amer. sociol. Rev., 1954, 693-701.

Homans, G.C. The human group. New York: Harcourt, Brace, 1950.

Husband, R. W. Cooperative versus solitary problem solution. J. soc. Psychol., 1940, 405-409.

Jennings, Helen H. Leadership and isolation: A study of personality in interpersonal relations. New York: Longmans, 1943.

Johnson, D.M. The psychology of thought and judgement. New York: Harper, 1955.

Judson, A.J. and Coffer, C.N. Reasoning as an associative process: I. "Direction" in a simple verbal problem. Psychol. Rep., 1956, 469-476.

Katz, D. Blendungsphanomene and konnektive Hemmung bei Denkprozessen. Theoria, 1949, 141-154. Summarized in Duncan, C.P. Recent research on human problem solving. Psychol. Bull., 1959, 397-429.

Kaufman, G. L. The book of modern puzzles. New York: Dover Publications, 1954.

Kay, H. The effects of position in a display upon problem solving. Quart. J. Psychol., 1954, 155-169.

Kelley, H.H., and Thibaut, J.W. Experimental studies of group problem solving and process. In G. Lindzey (Ed.) Handbook of Social Psychology. Vol. II. Cambridge, Massachusetts: Addison-Wesley, 1954, 735-785.

Klemmer, E.T. Tables for computing informational measures. In Quastler (Ed.) Information theory in psychology. Glencoe, Illinois: Free Press, 1955.

Lindzey, G., and Borgatta, E.F. Sociometric measurement. In G. Lindzey (Ed.) Handbook of social psychology. Vol. 1 Cambridge, Massachusetts: Addison-Wesley, 1954, 405-448.

Lorge, I., Fox, D., Davitz, J., and Brenner, M. A survey of studies contrasting the quality of group performance and individual performance, 1920-1957. Psychol. Bull., 1958, 337-372.

Lorge, I. and Solomon, H. Two models of group behavior in the solution of eureka-type problems. Psychometrika, 20, 2, 1955, 139-148.

Lorge, I., Tuckman, J., Aikman, L., Spiegel, J., and Moss, Gilda. Solutions by teams and by individuals to a field problem at different levels of reality. J. educ. Psychol., 1955, 17-24 (a).

Lorge, I., Tuckman, J., Aikman, L., Spiegel, J., and Moss, G. Problem-solving by teams and by individuals. J. educ. Psychol., 1955, 1960-166 (b).

Marks, M.R. Problem solving as a function of the situation. J. exp. Psychol., 1951, 74-80.

Marquart, Dorothy I. Group problem solving. J. soc. Psychol., 1955, 103-113.

Miller, G.A. Note on the bias of information estimates. In Quastler (Ed.) Information theory in psychology. Glencoe, Illinois: Free Press, 1955.

Mood, A.M. Introduction to the theory of statistics. New York: McGraw-Hill, 1950.

Pearson, K. Tables of the incomplete gamma function. London: Cambridge University Press, 1922.

Proctor, C.H., and Loomis, C.P. Analysis of sociometric data. In Marie Jahoda, M. Deutach, and S.W. Cook (Eds.) Research methods in social relations: With special reference to prejudice. Part Two. New York: Dryden, 1951, 561-585.

Quastler, H. (Ed.) Information theory in psychology: problems and methods. Glencoe, Illinois: Free Press, 1955.

Ray, W.S. Complex tasks for use in human problem-solving research. Psychol. Bull., 1955, 134-149.

Roby, T. B., and Lanzetta, J.T. Considerations in the analysis of group tasks. Psychol. Bull., 1958, 88-101.

Rogers, M.S. and Green, B.F. The moments of sample information when the alternatives are equally likely. In Quastler (Ed.) Information theory in Psychology. Glencoe, Illinois: Free Press, 1955.

Saugstad, P. An analysis of Maier's pendulum problem. J. exp. Psychol., 1957, 168-179.

Shaw, Majorie, E. Comparison of individuals and small groups in the rational solution of complex problems. Amer. J. of Psychol., 1932, 491-504.

Sherif, M. and Sherif, Carolyn, S. An outline of social psychology. New York: Harper, 1956.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Simon, H. A. A formal theory of interaction in social groups. Amer. sociol. Rev., 1952, 202-211.

Solley, C.M. Problem solving difficulty as a function of deviation of "meaning" of physical cues from expected "meaning." J. of gen. Psychol., 1957, 165-171.

Taylor, D.W. and Faust, W.L. Twenty questions: Efficiency in problem solving as a function of size of group. J. exp. Psychol., 1952, 360-368.

Thorndike, R.L. On what type of task will a group do well: J. abnorm. soc. Psychol., 1938, 408-412.

Walker, Helen M. and Lev, J. Statistical Inference. New York: Holt, 1953.

Watson, G.B. Do groups think more efficiently than individuals? J. abnorm. soc. Psychol., 1928, 328-336.