AN EXPERIMENTAL STUDY OF CASE-DISCUSSION LEARNING

by

Leonard A. Ostlund
A.B., Colgate University, 1948
A.M., Clark University, 1950

Submitted to the Department of Psychology and the Faculty of the Graduate School of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Advisory Committee:

Redacted Signature

Redacted Signature

Redacted Signature

July 28, 1953
ACKNOWLEDGMENTS

The writer is indebted to members of the administration, staff, and student body of the University of Kansas, whose cooperation made this study possible.

The Thesis Committee gave generous assistance. They were: Dr. Fritz Heider, Dr. Marston McCluggage, Dr. Hilden Gibson, and the Chairman of the Thesis Committee, Dr. Anthony Smith, who contributed valuable insights throughout the study.

The study was expedited considerably by the clerical, stenographical, and statistical assistance of my wife, who also typed the manuscript.
## CONTENTS

### PART I

**THE PROBLEM AND THE EXPERIMENT**

<table>
<thead>
<tr>
<th>Chapter I - The Problem and Related Research:</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Problem:</td>
<td>1</td>
</tr>
<tr>
<td>&quot;Case-Discussion&quot; Method Teaching:</td>
<td>2</td>
</tr>
<tr>
<td>Representative Experimentation:</td>
<td>8</td>
</tr>
<tr>
<td>Review of Related Research:</td>
<td>14</td>
</tr>
<tr>
<td>General Predictions:</td>
<td>25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter II - The Experimental Design and The Sample:</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Experimental Design:</td>
<td>26</td>
</tr>
<tr>
<td>The Experimental Groups:</td>
<td>28</td>
</tr>
<tr>
<td>The Control Groups:</td>
<td>28</td>
</tr>
<tr>
<td>The Number of Subjects Per Group:</td>
<td>29</td>
</tr>
<tr>
<td>The Scheduling of Groups Per Semester:</td>
<td>30</td>
</tr>
<tr>
<td>The Matching Criteria:</td>
<td>31</td>
</tr>
<tr>
<td>Method of Contacting S.'s:</td>
<td>33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter III - The Experimental Situation and Techniques:</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Experimental Setting:</td>
<td>38</td>
</tr>
<tr>
<td>The Seating Arrangement:</td>
<td>38</td>
</tr>
<tr>
<td>The Use of Recording Equipment:</td>
<td>40</td>
</tr>
<tr>
<td>The Notation Technique:</td>
<td>41</td>
</tr>
<tr>
<td>The Leaderless Group Discussion Technique:</td>
<td>42</td>
</tr>
<tr>
<td>The Selection of the Discussion Problems:</td>
<td>45</td>
</tr>
<tr>
<td>The Criteria for the Discussion Problems:</td>
<td>46</td>
</tr>
<tr>
<td>The Pilot Study:</td>
<td>47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter IV - The Experimental Procedure:</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preliminary Considerations:</td>
<td>49</td>
</tr>
<tr>
<td>The Experimental Procedure:</td>
<td>50</td>
</tr>
</tbody>
</table>

### PART II

**THE ANALYSIS OF THE DATA**

<table>
<thead>
<tr>
<th>Chapter V - Some Qualitative Observations:</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapport:</td>
<td>57</td>
</tr>
<tr>
<td>Solutions: That Occurred Within the Time Limit.</td>
<td>59</td>
</tr>
<tr>
<td>The S.'s Opinions Concerning the Problems:</td>
<td>61</td>
</tr>
<tr>
<td>Charting Dimensions of Groups:</td>
<td>65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter VI - The Index of Participation:</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equality of Participation:</td>
<td>67</td>
</tr>
<tr>
<td>The Problem of Choosing Appropriate Units.</td>
<td>68</td>
</tr>
<tr>
<td>Hypothesis and Results:</td>
<td>71</td>
</tr>
</tbody>
</table>
Chapter VII - Extent and Agreement Concerning Perceived Communication Channels:
The Communication Data Form.......................... 73
Extent of Communication Channels........................ 74
Hypothesis and Results................................. 74
Agreement Concerning Perceived Communication Channels...................... 75
Hypothesis and Results................................. 75
Agreement on Perceived Communication Channels - Successes Only............. 76
Hypothesis and Results................................. 76
Conclusion................................................ 77

Chapter VIII - Group Agreement Concerning "Leadership" Hierarchy:
Introduction................................................ 78
The Rating Scale........................................ 78
Procedure.................................................. 79
Hypothesis and Results................................. 79

Chapter IX - The Written Solutions:
"The Tennis Problem".......................................... 83
Rationale for the "Centering" Category........................ 84
Criteria Concerning Inadequate Solutions.......................... 86
Rationale for the "Means" Category................................ 86
Criteria Concerning "Means".................................. 87
Reliability of the Judges..................................... 89
"The Tennis Problem" - Judges' Agreement..................... 91
"The Assembly-Line Problem - Judges' Agreement.................. 93
Conclusion................................................. 94
Tabulation of the Judges' Ratings................................ 95
Hypothesis and Results..................................... 98
Conclusion................................................. 98
The Problem Solving Behavior of "The Average College Student"........... 98
"The Assembly-Line Problem".................................. 100
Criteria....................................................... 102
Criteria Concerning Concrete Action Proposals........................ 104
Criteria Concerning Centering.................................. 105
Criteria Concerning Inadequate Solutions........................... 106
Criteria Concerning Means.................................... 106
Tabulation of the Judges' Ratings................................ 106
Hypotheses Concerning Concrete Action Proposals............................. 107
Hypothesis Concerning "Inadequate Solutions"............................... 113
Hypothesis Concerning "Centering".................................. 113
Hypothesis Concerning "Means".................................. 114
Conclusion................................................. 114

Chapter X - The Group Discussion Transcripts:
Introduction.................................................. 116
Transcripts as Scientific Behavioral Data.............................. 117
The Transcript Judgment Design..................................... 118
Reliability of the Judges...................................... 123
Chapter X - The Group Discussion Transcripts, Cont'd.:
Agreements Concerning Type of Group.............123
Agreements Concerning Solution and Adequacy....126
Agreements Concerning Continua Ratings...........126
The Tabulation of the Data..........................128
The Criteria..........................................129
Hypotheses and Results...............................132
Solutions and Adequacy...............................132
Type of Group........................................136
Continua Ratings.....................................137
The General Pattern of Both Sessions.............141

Chapter XI - Discussion:
Introduction..........................................144
Adequacy of Experimental Tools and Techniques..144
Adequacy of Experimental Design....................144
Incidental Findings..................................145
Major Findings.......................................148
Recommendations for Curriculum Change............153
Recommendations for Further Investigation........154
Conclusion............................................156

Chapter XII - Summary:
..........................................................158

*    *    *

Appendix..............................................164-203
Bibliography.........................................204-211
INDEX OF TABLES

Table Number | Page
-------------|------
I. The Experimental Design | 27
II. Table of Summary Data | 56
III. Written Solutions: Number of Instances in Which Both Judges Agreed | 90
IV. Judges Agreement Concerning Written Solutions to "The Tennis Problem" | 92
V. Judges Agreement Concerning Written Solutions to "The Assembly-Line Problem" | 96
VI. Frequencies of Forms of "Centering" in "The Tennis Problem" in the Experimental and Control Groups | 97
VII. Frequencies of "Adequate and Inadequate Solutions" to "The Tennis Problem" in the Experimental and Control Groups | 97
VIII. Frequencies of "Mutual Means" for "The Tennis Problem" in the Experimental and Control Groups | 97
IX. Frequencies of "Concrete Action Proposals to "The Assembly-Line Problem" in the Experimental and Control Groups | 108
X. Frequencies of Forms of "Centering" in "The Assembly-Line Problem" in the Experimental and Control Groups | 108
XI. Frequencies of "Adequate and Inadequate Solutions" to "The Assembly-Line Problem" in the Experimental and Control Groups | 109
INDEX OF TABLES
(continued)

Table Number | Description                                                                 | Page
---|---|---
XII. | Frequencies of "Use Joe Productively" Means for "The Assembly-Line Problem" in the Experimental and Control Groups | 109
XIII. | The Transcript Judging Design | 120
XIV. | Ability to Judge Type of Group from Discussion Transcript | 125
XV. | Reliability of Judges Concerning Presence or Absence of Group Solution | 127
XVI. | Frequency of Group Solutions | 133
XVII. | Tabulation of the Transcript Ratings | 139
XVIII. | Index of Participation Formula | 171
XIX. | Index of Participation | 172
XX. | Summary Chart for Communication Data | 173
XXI. | Summary Chart for Communication Data: Subjects - A Group as a Whole - B | 174
XXII. | Agreement in Perception of Communication Channels | 175
XXIII. | Agreement in Perception of Communication Channels (Successes Only) | 176
XXIV. | Correlation Coefficients and Z Transformations, A, B, C, D | 177-180
XXV. | Range on Leadership Scale | 181
PART I

THE PROBLEM AND THE EXPERIMENT
The Problem

A perusal of educational and psychological literature indicates that in recent years there has been an unmistakable trend toward the use of the case-discussion method of teaching. Because of the far-reaching effects of educational processes, the validation of new educational techniques becomes imperative.

In its broadest terms, the problem investigated concerned the consequences of this new educational procedure. An experimental methodology was designed that would measure differences in behavior between experimental groups of students currently taking a "case-discussion" course and control groups of students who were not taking a case-discussion course. In order to study the problem experimentally, the writer observed the behavior during two sessions of twelve leaderless discussion groups, composed of six members each, who had been drawn from "case-discussion" and non-case-discussion courses. Each experimental (case-discussion) group was matched with a control (non-case-discussion) group. Each group met for an hour at the beginning of the semester and again at the end of the semester. Their task was the discussion and solution
of problems in interpersonal relations. The present study was conducted at the University of Kansas during the academic year of 1952-53. It was essentially an attempt to evaluate the extent to which certain goals had been achieved by the "case-discussion" method of teaching. The nature of these goals will be made clear later in this study.

"Case-Discussion" Method Teaching

"Case-discussion" method teaching at the University of Kansas owes its principal stimulus to the influence of Wallace B. Donham, former Dean of the Harvard Business School, and the group associated with Mr. Donham at that Institution. Professor Carroll D. Clark, Chairman of the Department of Sociology was the first member of the University of Kansas faculty to participate in the training program instituted by Professor Donham in the fall of 1945. He was a participant in the program during the academic year of 1945-46. During the ensuing five years, five other members of the University of Kansas faculty spent either a half or a whole year as participants in the Harvard program.*

*These participants were: Professor Leonard Axe, Dean of the School of Business, Professor Hilden Gibson, Chairman of the Department of Human Relations, Professor Marston M. McCluggage, Professor Edward G. Nelson, and Professor Edwin C. Stene.
Upon their return to their own campus, these men became the nucleus from which has spread an interest in the "case-discussion" method of teaching. In addition to the Department of Human Relations, courses taught largely by the "case-discussion" method are now offered in the School of Business, the Department of Sociology, and the Department of Political Science.

Were this a purely local phenomenon, it might scarcely merit systematic investigation. However, even a cursory study will reveal that the Harvard influence has spread to many institutions other than the University of Kansas. An annual Human Relations Conference is held and is attended by representatives of many institutions who use the "case-discussion" method of teaching in a wide variety of courses. Especially prominent among these institutions are Colgate University, where the writer himself had experience with "case-discussion" teaching, Ohio University, and the University of Washington.

The content of this type of course consists of problems in interpersonal relations which occur in the family, college, and industry. Students are presented with factual accounts of incidents in which there has been some disturbance in the relationships between groups or between individuals. Students probe the nature of the difficulties and are expected to develop some understanding of the basic motivations and attitudes of the individuals involved. Student-
student participation is expected and students are encouraged to assume responsibility for the readings and for participation in class discussion. Students are provided with case-books and readings, but in general they are not often examined concerning the readings. Instead, they are expected to demonstrate some appropriate applications of the readings. Grading is determined largely on the basis of classroom participation, individual growth, and upon examinations which consist of analyses of cases.

The teaching technique is largely non-directive. Ordinarily, the teacher does not lecture, though upon occasion, he may spell out some points from the readings which seem to be particularly relevant to the task at hand. The teacher attempts to set a permissive atmosphere in which each student will feel free to participate. It was from courses of this kind that the experimental S.'s were selected.

In recent years, a growing number of publications has been concerned with this type of course. The first of these was by Donham, whose book, Education for Responsible Living (42) pointed to the need for the course. In his experience in working with business executives and college students, he became convinced that the colleges were graduating competent technicians, but that these individuals were woefully lacking in interpersonal skills. An article by this same author describes the underlying philosophy
and methodology of the course (41). An article by Berrien does likewise and points out the need for this course in the psychology curriculum. It is stressed that the usual elementary course doesn't help students to solve everyday problems of human relations. Further, it is stated that the goal of the course is to:

"...develop an attitude and a way of thinking rather than a technical vocabulary" (21, p. 148).

Roethlisberger writes that an important value of the "case-discussion" method of teaching is that it functions to improve communication skills. He severely criticizes contemporary college teaching as follows:

"Although our colleges and universities try to get students to appreciate intellectually points of view different from their own, very little is done to help them to implement this general intellectual appreciation in a simple face-to-face relationship - at the level of a skill. Most universities train their students to be logical, lucid, and clear. Very little is done to help them to listen more skillfully" (76, p. 93).

An article by Beé, one of the teachers who cooperated in the present study, explains how "case-discussion" course principles and methodology have been used in a course entitled, "Cases in Marriage and the Family." Criteria for evaluation are also included in this article (18).

Representative textbooks have also appeared. Comments
and Cases on Human Relations, by Berrien (22) has been used at Colgate University for several years. At Harvard University and at the University of Kansas, The Administrator, by Glover and Hower (49) has been used as a textbook in the School of Business for several years. The cases in this book are all industrially oriented, whereas those in Berrien involve school, family, and industrial situations. Cabot and Kahl's book, Human Relations (31) is a compendium of readings oriented toward a Human Relations Course.

Moreover, the Harvard influence is not the only one in this direction. While terms applied to this method of teaching vary, it has become fairly widespread in recent years, and is held by its exponents to be fruitful for a variety of subjects. Their work will be discussed shortly.

Other educators stress discussion as an important aspect of their teaching procedure. Maier has several publications concerning the discussion of problems in interpersonal relationships in an industrial setting. They have played an important part in a management training program at the University of Michigan (67), (68). He has recently published a textbook entitled Principles of Human Relations (67) which is used in this course. Another educator who stresses discussion is Cantor. In his book, The Dynamics of Learning (33), he stresses the creation of a permissive atmosphere in the classroom, after which the teacher should force the students to
accept and face the responsibility for their decisions.

It is important to note that this viewpoint is not shared by some who teach by the case method. Opponents of Cantor's view favor a permissive, Rogerian approach in which the teacher reflects, clarifies, and summarizes, but does not challenge the students bluntly as Cantor does. The teacher does not interject his ideas as Cantor does, but allows the group more freedom in choosing their direction.

In a recent article, Cantor has joined the issue, and stated his point of view as follows:

"The final responsibility for determining goals lies not with the group as a whole, but with the leader and the group" (32, p. 382).

A discussion course which has been described in the literature as somewhat similar to the "case-discussion" method courses previously outlined is entitled, "Culture and Personality," and is taught at the University of Chicago by David Riesman. In this course, selected readings from books are used as the media for classroom discussion. An article by this teacher describes the course and mentions problems common to those faced by teachers in "case-discussion" method courses. For example, a common problem in "case-discussion" teaching is the "silent" student. Regarding this problem, Riesman writes:
"Silence in class is hardly more welcome than in a broadcasting studio.... The teacher must be able to endure the silence....and the hostility or uneasiness of the class because of the silence. He must be able to wait...."

While we thus have considerable evidence to demonstrate that discussion course akin to "case-discussion" method courses are being taught in many colleges and universities, there is also a trend to introduce occasional case-discussion in more traditional, lecture-type courses. However, the use of cases as a medium for instruction is not new, it has been prevalent in medicine and law for centuries. It is only within the past decade that cases have become widely utilized in the social sciences, whereas medical cases of Hippocrates and legal cases of the Romans have long been classics of instruction.

Representative Experimentation

In general, research pertaining to the "case-discussion" method of teaching falls into two classifications. The first comprises investigations which focused on the relationship between performance in the classroom, such as the amount of participation, and other criteria, such as grades or test scores. The second classification consists of studies which measured behavior changes, such as an increased awareness of the complexities of interpersonal relations, attributable to case method teaching.

Under the first classification, an investigation by Wetherell involved the study of four sections of a Human
Relations Course in order to investigate the assumption that the more an individual participated, the more he would learn. Participation was defined in terms of the number of words spoken by each individual and learning was defined in terms of the individual's final examination score. The correlation was low and statistically insignificant, indicating that quantity of participation was not necessarily associated with the type of learning which was called for on final examinations (84).

Butler studied a Human Relations Course at Ohio University and investigated the discussion by the use of Bales' Interaction Process Analysis (30). In studying the correlation between the pattern of the individual's reactions as a member of the classroom group and course grade, no significant relationship was found. The correlation between pattern of interaction and mental ability, as measured by a standardized test - The Ohio State Psychological Examination - likewise displayed no significant relationship (30, pp. 99-102). Both these studies agree that achievement as measured by grades is not positively related to the more obvious aspects of classroom participation.

We now turn to studies that have been primarily concerned with changes in behavior which may be attributed to the course. Berrien states that one appropriate criterion for "case-discussion" method courses concerns the progressive refinement of ideas. He reproduces a portion of a
recorded classroom discussion, and by a content analysis, indicates that the group, in discussing the background to the problem, began with vague generalizations, which were progressively refined into more meaningful hypotheses with which the problem was approached. The author concluded:

"This movement of the discussion from generalized labels to interpretations of behavior is a mark of insight" (22, p. 478).

Kelley and Pepitone studied a class in Human Relations by giving the students problems in interpersonal relations and twenty minutes was given for the analysis of each problem. The answers were then categorized according to the degree of insight into human relations. Answers were classified into two broad categories, "interactional-conceptual treatment" and "superficial treatment". Students were divided into three groups and each group was given this task at a different period during the term. The results indicate that the later the group performed this task, the more insightful the results. From these findings, the authors conclude that the course produced significant changes on both the intellectual and attitudinal levels. (59).

Goedicke studied a pre-college group, composed of high school seniors who belonged to a young peoples' organization in a church. She conducted a series of group discussions in the same manner as the Human Relations Course. At the conclusion of the experiment the 3.'s
filled out questionnaires concerning the effect of the course on their approach to problems in interpersonal relations and related topics. The results indicated a trend, in that the students felt that they had benefitted by the case-discussion experience for it had made them more aware of problems in interpersonal relations. Moreover, the course was received very favorably and this was also indicated by the high rate of attendance which was voluntary (50).

Castore's study also involved changes attributed to the course which were ascertained by means of a questionnaire. He also correlated class behavior with other measures. He reported a study of three sections of a Human Relations Course. The methodology involved a six item attitude blank which was filled out and returned by students at the end of each class session. The items were:

"(1) This discussion was constructive.  
(2) Problems were clarified as the case advanced.  
(3) Instructor influenced discussion.  
(4) I thought deeply about the case today.  
(5) I participated in class discussion.  
(6) Outside affairs reduced my interest."

Each item was followed by seven rating categories ranging from 'not at all' to 'very much.'" (37, p. 202).

One important behavior change that was attributed to the course was some evidence that training in the use of the "case-discussion" method tended to place the burden
of responsibility on the students. In comparing classroom behavior with other measures, it was found that those students who were judged by others as having made the most sensible remarks tended to receive the higher grades. Moreover, those who were judged as having made the most sensible remarks tended to have a higher frequency of class participation (37, pp. 212-213).

Another study probed the dip in morale, a recurrent problem in "case-discussion" courses, which usually occurs early in the second half of the term. The students were encouraged to discuss this problem which was very real to them, in lieu of the usual case-book problem assignment. Results indicated that the class had demonstrated some ability to use their "case-discussion" experience in dealing with their own problem. Moreover, the experience of discussing their own problem had a salutary effect upon class morale (36).

The last study to be cited concerning behavior changes, was done by the writer at the University of Kansas. A "case-discussion" course entitled "Business Administration Practices," which was composed of upperclassmen was investigated by means of sociometric techniques. Students were asked to state their present leisure-time associates and their acquaintances at the beginning of the semester, as well as their choice concerning leisure-time associate and work-associate. The data was analyzed by means of
sociometric statistics and correlation statistics. The findings indicated that a high degree of group integration had emerged as a result of their experiences in this particular classroom situation (73).

The present study was also an experimental exploration of behavior changes that may be attributed to the "case-discussion" method of teaching. However, the present study differed from all of the foregoing studies in that it attempted to investigate the transfer of case-discussion learning from the classroom to a different situation, and that a control group was used as a yardstick for comparison.

All of the foregoing studies had one important procedure in common. The S.'s met together in a classroom at the beginning of the semester, and were investigated some time later in the same classroom situation. In the present study, S.'s were drawn from classrooms into a different situation, an experimentally created setting. This difference in procedures is stressed, because under the latter circumstances, behavior changes which occurred in the classroom would have to be transferred into the new situation. This was felt important, for it may be argued that in the previous studies cited, such behavior changes as noted might have been anchored to the specific classroom situation in which they had been detected.

It should be noted that the experimental research
just cited concerns the "case-discussion" method of teaching which was instigated by Professor Donham at The Harvard School of Business. However, it is probably apparent that this type of course has much in common with other case-discussion courses which attempt to productively use the resources of the classroom group both for content learning and for the formation of positive attitudes and interpersonal skills. Though the "case-discussion" method of teaching under investigation in the present study originated in the curriculum of the Harvard School of Business, it has spread to other departments and colleges, and gained the attention of other social scientists. Moreover, the techniques inherent in this type of course come under the general heading of what has been called "student-centered teaching." It is therefore appropriate to briefly review the literature concerning experimentation in this broader field which is so closely related to the "case-discussion" method of teaching which was instigated at Harvard.

Review of Related Research

A survey of the literature reveals that many psychologists have been interested in similar aspects of classroom teaching and learning. In fact, quite a few psychologists have used their classes in psychology for experimental purposes, just as have the "case-discussion" method teachers. Therefore, it seemed to the writer that the problems under investigation in the present study were
not necessarily encapsulated within the orientation, as described. They were directly related to disciplines that emphasize the constructive usage of classroom group dynamics. Within the past five years a considerable amount of research has been carried out in this area, of which some representative studies follow.

We may begin this review of related research with a consideration of the importance of the classroom as a center for interpersonal activity. Most American children go to school for many hours per day over a period of twelve years. Therefore, the interpersonal relationships between students, and between students and their teacher, and the effect of these relationships on the learning process has been the focus of many investigations. Wright, et al. emphasized the importance of social interaction in the classroom. Their study of children involves the unusual technique of having a team of observers follow a child around from morning till night (7), (8). In the course of their intensive observation of Raymond Birch they determined that twenty-three percent of the episodes in his day occurred in the classroom. In approximately seventy percent of these, Raymond interacted with one or more persons. The authors conclude:

"These figures call attention to the significance of the classroom as a facility for the study of social interaction and they call attention to the practical need for information
that will enable us to guide better the multitudinous social events of the behavior setting" (88, p. 200).

Thelen has also stressed the influence of the classroom situation on adolescent personality formation in a consideration of factors that operate to enhance adjustment. One of the principles which he considers important, in understanding the impact of the classroom on the individual, is that the potency for the pupil of the classroom group depends upon the extent to which it enables him to meet his needs. (82). Jenkins has also been concerned with the impact of the classroom situation on the personality of the student, though on the college level. In an article that presented a review of some long-range research in this area, he echoed Thelen's viewpoint, and carried it one step further stating:

"The greater the potency of the classroom group (for meeting the needs of the individual member), the greater will be his personal involvement in what occurs in the classroom and consequently, the greater the learning" (57, p. 138).

Mensch was also concerned about the influence of the classroom environment on student behavior. He experimented with contrasting environmental conditions which had been experimentally created. He found that students in a "progressive" school environment differed significantly in their reactions to socially frustrating situations from students who had been trained under "authoritarian leadership." He concluded:
"... environments in which less frustration occurs tend to provide conditions under which the individual is most likely to utilize his own initiative in developing his capabilities" (72, p. 286).

These findings serve to recall the earlier and similar findings of Lippitt and White concerning the experimentally contrived social climates (65).

Closely related to these studies are three studies of attitudes in differently structured classroom situations. One of these was by Deutsch. He used experimentally created groups chosen from college students enrolled in an introductory psychology course in order to study the relative merits of cooperative and competitive grading systems. The groups were so structured that the members of the cooperative groups all shared in rewards, whereas members of competitive groups achieved rewards individually. The results revealed no evidence to indicate that either grading system produced greater student interest or task involvement, or that either system resulted in greater learning. However, the author found interesting differences between the groups which he summarized as follows:

"Our results suggest that the intercommunication of ideas, the coordination of efforts, the friendliness and pride in one's group which are basic to group harmony and effectiveness appear to be disrupted when students see themselves to be competing for mutually exclusive goals. There is some indication that competitiveness produces greater personal insecurity (expectation of hostility from others), than does cooperation."
In addition, it is evident that greater group productivity will result when members of a group are cooperative rather than competitive in their relationships" (40, p. 152).

Another study in this area was conducted by Bills. He sought to determine if students taught by a traditional lecture-discussion method would differ in their understanding of course content in general psychology from a group of students taught by the student-centered method. Though the author does not spell out what he means by "student-centered" method, he states that these classes were conducted in a Rogerian non-directive fashion. It may be safely assumed that this entailed a group discussion situation of the type previously described in other studies. Evaluation was done by objective examinations which were administered to both groups. Bills concluded:

"...the students learned an equal amount of text book material from the two methods, but the attitudes of the student-centered class toward psychology was significantly more positive than were those of the lecture-discussion class. The opinions of the students in the student-centered class revealed that they believed the course was of personal value, but the students in the lecture-discussion class did not concur in this opinion" (23, p. 317).

McKeachie, experimenting in the same vein, reported a study in which a general psychology course was sub-
divided into groups which were conducted according to three teaching methods, and students were asked to indicate their preferences. The following is a summary of the three teaching methods used:

1. "discussion," in which the instructor acted as chairman, summarizer, stimulator, and informant. Maximum student participation was encouraged.

2. "study-tutorial," in which the students were expected to work at their own speed. The instructors provided additional readings in the general areas being studied. The students consulted the instructor individually.

3. "recitation," in which the instructor gave brief lectures, presented demonstrations, and followed the practice of conducting a specific question-and-answer routine.

One of the conclusive findings was that students preferred the recitation section. This was not surprising when one considers that their reason was logical - this section provided optimum preparation for the objective-type final examinations (66, p. 154).

It is interesting to note that this same finding has been reported by Wispe, who studied eight sections of an elementary course in Social Relations. These sections were matched and taught by different methods for one semester. Four were described as directive, highly-structured, and subject-matter centered, four were permissive, unstructured, and student-centered. The investigator reports:
"The directive sections were preferred by most of the students, because they were clearly defined, and for their presumed value in preparing for examinations; although the permissive sections were enjoyed more." (86, p. 184).

However, other findings from this same experiment confirm results cited by other investigators, with regard to the attitudes of students, which has been summed up as follows:

"The permissive sections had more interest, humor, student-participation, student-instructor interaction, and commending, while the directive sections offered more course relevant information formally presented." (86, p. 183).

Two further studies concerning attitudes of students may be mentioned. However, in these studies, there was no attempt to create different kinds of classroom situations. Marsh experimented with an introductory psychology course by giving them a questionnaire concerning the realization of course goals. The results indicated that the class felt that objectives relating to personality and emotions had been more successfully attained than others. (70, p. 384).

In a further ramification of this research, the same experimenter gave another class a questionnaire of this same nature. At the conclusion of their first course in psychology, the students were asked for their opinions pertaining to how important fifteen stated objectives were for them and to specify how well these same objectives had been attained. The objective which received the
highest average rating was "A genuine effort toward self-improvement psychologically." The objective rated lowest was, "A development of interest in psychological issues shown in reading and thinking beyond the required assignment." In conclusion, the experimenter stated:

"Results indicated that students considered the objectives to be of substantial importance but that there was only a fair degree of relationship between how important an objective was judged to be and how well it had been attained in the course" (71, p. 142).

An experiment by Karn is worthy of mention since it is somewhat related to the present study. Karn tested a non-case-discussion psychology class in order to measure the learning of supervisory skills. The instrument used was the File-Remmers Test, "How Supervise?" a paper and pencil test of situations in which a choice of action is to be made. The test was administered at the beginning and at the conclusion of a course of which the learning of supervisory skills was a primary goal. As a control group, Karn selected a group of students which comprised a class in English. The psychology group made significant gains (53).

The writer would like to point out that Karn's study is similar to the present study in that both studies employed control groups, of a different population from the one which was the focus of the study and that the control groups were tested before and after the learning exper-
ience which was given only to the experimental group. However, there is an important difference, namely, that Karn's control group was of a special population - English students - whereas the writer's control group was of a more general population, which may be called the "average non-case-discussion student." Another important difference is that Karn's measurement was by means of a paper and pencil test, whereas one of the measurements used in the present study was the rating of actual behavior in the experimental setting. Paper and pencil tests are open to the criticism that "changes" may be on the verbal or formal level, and thus do not constitute a truly valid measure of how an individual would react if faced with the actual situation. Furthermore, when such paper and pencil tests are administered in the classroom situation by the teacher, it is always possible that the students may perceive this test as a further evaluation of their course progress, and they may tend to produce answers biased by what they know of the teacher's predilections. It should be pointed out that the present study was conducted by E. who had no authority status, and that the experimental situation was not presented as a "test," nor was it held in the classroom situation.

In concluding this section concerning research related to the "case-discussion" method of teaching - research which has attempted to explore the effectiveness of teaching methods - we shall refer to a timely summary
of research in this area by Wispe. This author traces the problem of instructor versus student-centered teaching to the influence of John Dewey and those who were concerned with what has been called "progressive education." In general, their interest was:

"...primarily with the way individuals met and solved problems, with the habits they developed in adjusting to their environment, and with the implications of these for democratic living" (87, p. 147).

Wispe goes on to explain that this controversy spread into college teaching and became known as the "lecture method" versus the "discussion method." The "discussion method" flourished since educators believed that it was more important to teach students how to think than what to think. However, the terminology was vague and ill-defined. More recently, the controversy has become known as "instructor-centered" versus "student-centered" teaching, with conceptual and methodological improvement, notably the Lewinian concept of the "interdependence" of needs, and Rogerian non-directive concepts. While further refinements of conceptual formulations are needed, Wispe's review makes it clear that the contributions of both of these contemporary schools of psychology have proved beneficial; stating:
"Essentially this restatement in Le-\nwinian and Rogerian terms has served \nto integrate the teaching problem \nwith the fields of group dynamics \nand psychotherapy, from which many \nimportant insights into the discus-\nsion and the lecture process have \nbeen gained" (87, p. 147).

Thus a survey of research in the related general \narea of the effect of different teaching methods upon \nthe learning process indicates that the traditional lec-\nture method of teaching is useful in mastery of content, \nand that non-lecture methods apparently do equally well. \nConcerning favorable changes in attitudes and interper-\nsonal skills, there is a striking agreement among these \nstudies. Non-lecture methods of the type that stress \nfree participation and student responsibility for class-\nroom discussion prove superior to traditional lecture \nmethods. The non-lecture method effects changes of atti-\ntudes in the direction of a positive valence for subject \nmatter and also tends to improve interpersonal skills.

Because of these considerations, further investiga-\ntion of non-lecture teaching, as practiced in the Harvard \ninfluenced "case-discussion" method courses and in relat-\ned courses, seems appropriate. As will be explained in \nthe next chapter, the present study involved a rigorous \nexperimental design and a methodology which required the \nS.'s to demonstrate a transfer of skills. This approach \nmade it possible to investigate important issues which \nhave not been covered in previous investigations.
The experimental methodology was designed to elicit behavior amenable to qualitative and quantitative analysis, so that differences between groups could be determined. It was predicted that the experimental group, because of their "case-discussion" method of training, would excel the control group who lacked this training. The following predictions were framed in general terms. They will be treated in greater detail as specific hypotheses in Part II-The Analysis of the Data.

General Predictions

With regard to group interaction, it was predicted that the experimental group would surpass the control group in communication skills. Pertaining to leadership functions and participation, it was predicted that they would tend to share these aspects of the discussion process more equally than the control group.

In the matter of group problem-solving, it was predicted that the experimental group would demonstrate more insight into the basic issues involved in the problem, show fewer inadequacies in their solutions, and, in general, use the most constructive approach in view of the possibilities and limitations of the total situation.
CHAPTER II.
THE EXPERIMENTAL DESIGN AND THE SAMPLE

The Experimental Design

The experimental design involved experimental and control groups which were matched according to certain criteria which will be discussed shortly. Each group was tested at the beginning of a semester and at the conclusion of the same semester.

The experimental S.'s were drawn from case-discussion courses, whereas the control S.'s were drawn from non-case-discussion courses. The experimental S.'s, during the period between tests, were having experience in the case-method of instruction which stresses classroom discussion, whereas the control group was not. Therefore, it would be appropriate to attribute changes which occurred in the experimental group to case-discussion learning, provided these changes did not occur in the non-case-discussion learning group. Thus the control group functions as a constant frame of reference against which changes in the experimental group may be evaluated. Since the control group has not had the experience of this special educational training, we cannot expect any changes in their behavior.

The control group thus provides a base-line, as it were, against which we may evaluate the progress of the experimental group. We can determine whether they have
been able to transfer their case-discussion learning to
this new situation. This experimentally created situation
is purposely unstructured so that each S. is provided with
an equal opportunity to demonstrate any general increase
in maturity of approach to the discussion and solution of
problems in interpersonal relationships, which has occurred
over this three-month interval. The experimental design
used in the present experiment has been reproduced below:
and salient aspects have been pointed out.

<table>
<thead>
<tr>
<th>TABLE I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>THE EXPERIMENTAL DESIGN</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Control Groups</th>
<th>Experimental Groups</th>
<th>Total Meetings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>After</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Total Meetings</td>
<td>12</td>
<td>12</td>
<td>24</td>
</tr>
</tbody>
</table>

The number of S.'s per group was held constant at six
for each meeting. There were six experimental ("case-discus-
sion" method) groups which met for two sessions, one at
the beginning and one at the end of the semester. Six con-
trol (non-case-discussion method) groups did likewise. The
"after" sessions occurred three months later than the "be-
fore" sessions. The same S.'s came to the "before" and
the "after" sessions and the personnel for each group was
kept the same for both of these meetings. With few excep-
tions which were occasioned by drop-outs and refusals, almost all of the seventy-two S.'s each came to two, one-hour meetings.

The Experimental Groups

The experimental groups were chosen from four courses taught by the "case-discussion" method: Cases in Human Relations, Personnel Management, Cases in Marriage and the Family, and Business Administration Practices. The cases used in these courses all involve problems of interpersonal relations. While the teachers differ in regard to age, experience, and classroom technique, these courses spring from a common source - the seminars given at the Harvard Business School by Dean Wallace B. Donham - and adhere to a common set of underlying principles and methodology.

There were six groups of experimental subjects and each group met at the beginning and at the end of a semester. Each group consisted of six subjects who were drawn from different sections of the same course or from different "case-discussion" method courses, depending upon availability of the subjects or their suitability in terms of the matching criteria.

The Control Groups

The control group subjects were drawn from classes whose students were registered in five different schools of the University in order to gain an adequate sample of
the "average non-case-discussion course student." These schools included the College of Liberal Arts, the School of Engineering, the School of Journalism, the School of Education, and the School of Business.

It was felt that balancing the personnel within a control group of six non-case-discussion students by selecting them from different schools, the control group would prove more representative of "the average non-case-discussion student." Therefore, each control group was arranged so that no more than two members, in each group of six subjects, were from the same school. Moreover, these two individuals were chosen from different courses within the school. This arrangement was possible because there was a wide range of non-case-discussion courses in each of these five schools.

There were six groups of control subjects, and each group of the same subjects met at the beginning and at the end of a semester. Each group consisted of six subjects who were drawn from different sections of the same course or from different non-case-discussion courses.

The Number of Subjects Per Group

The choice of six subjects per group was not an arbitrary one, but was governed by important considerations. Practical considerations involved the fact that the size of the group could not be too large, since not enough case-discussion courses were available from which subjects could be drawn so that the subjects would come from a different
section or course. Then too, the group could not be too large so that E., working alone, might not be able to observe all the interactions. However, if the groups were too small, the results would not be amenable to statistical treatment.

Thelen, in discussing the size of the group, stresses the following principle:

"The size of group should be the smallest group in which it is possible to have represented at a functional level all the socialization and achievement skills required for the particular learning activity at hand." (20, p. 133).

He goes on to emphasize that if the group is larger than needed, there will be duplication of skills and less need for individuals to assume responsibility. If the group is smaller, there will be gaps. In view of these considerations, six subjects per group seemed an optimum number.

Groups of six subjects have been used by Bales (5), Stephan (79), and French (48). However, a survey of the literature revealed considerable variety in the number of S.'s per group used by different investigators. Groups of two (34), three (54), five (62), (19), and eight (15) have been reported. According to one writer, experimentation indicated that the group should consist of between six and eight for the judging of industrial personnel in a leaderless group discussion (69, p. 211).

The Scheduling of Groups Per Semester

It was necessary to start the "before" sessions as early
as possible in the semester so that no learning would have taken place. Because the clerical work involved in matching was so complicated and time-consuming, and because many personal contacts were necessary in order to arrange each group, three experimental and three control groups were used in a semester. This made it necessary to run the experiment over two semesters.

The Matching Criteria

Each S. in the experimental group was matched with an S. in the control group according to the characteristics described below.

1. Friendship - The underlying assumption was that friendships between S.'s would influence group interaction. In order to minimize friendships every group was composed of six individuals each of whom was from a different course or section of a course. E. also attempted to control this variable in the experimental groups by bringing together no more than two students in any group who were enrolled in the same school.

When the initial telephone contact was made, E. did not read out a list of the other member's names and ask the S. if he knew any of them. An easier remedy was available. As will presently be disclosed in detail, supernumeraries were invited in order to ensure a full complement of six S.'s. As each S. entered the experimental room, E. noticed whether or not he greeted anyone. Because of the previously mentioned method of selecting members from different
courses, sections, and schools, such cases were rare. When
this occurred, E. rapidly reshuffled the group according to
the matching criteria so that one of these individuals was
transferred to the supernumerary group. This device proved
satisfactory. Other investigators have considered it im-
portant to experiment with groups of individuals not pre-
viously acquainted, for example, French (49), Bass (17),
Bell and French (19), and Steinzor (78).

2. Grade Point Average - It was felt that unless this
characteristic were strictly controlled, disparities between
groups as to college achievement might mitigate the results.
Therefore, individuals were matched according to grade point
standing by consulting college records with the permission
of the Registrar and the Deans.*

3. Class Standing - It was felt that experience in col-
lege might be reflected in group discussion or problem-solv-
ing. Therefore, students were matched according to this
characteristic. However, no group was entirely composed of
any one class. The S.'s were upperclassmen, only a few
sophomores were used.

4. Previous Experience in "Case-Discussion" Courses -
No member of the control groups was currently enrolled in

*The writer acknowledges the assistance of the following
individuals: Kenneth Anderson, Dean, School of Education;
Leonard H. Axe, Dean, School of Business; T. De Witt Carr,
Dean, School of Engineering; James K. Hitt, Registrar; Paul
B. Lawson, Dean, College of Liberal Arts and Sciences; Burton
W. Marvin, Dean, School of Journalism.
any of the case-discussion courses from whom experimental S.'s had been selected, nor had they previously taken any of these courses. Of the experimental groups, no member had previously taken any of these courses, nor was he currently enrolled in any case-discussion course other than the one from which he was selected. Concerning case-discussion courses other than those under study, few appear in the curriculum, and all S.'s were questioned concerning these courses.

5. Miscellaneous Factors — The following characteristics were eliminated from both groups because it was felt that matching might prove difficult, or because the inclusion of such characteristics might unduly bias the results. These characteristics included: women, foreign students, Negroes, campus celebrities, students not regularly enrolled in the course, graduate students, freshmen, and E.'s friends, students, or counselors. E. made every effort to closely equate each control S. with an experimental S. However, some mismatches occurred due to the failure of an S. to appear for the experiment, and to a few substitutions for drop-outs. However, such instances were negligible.

Method of Contacting S.'s

During registration week in September, E. contacted members of the faculty, acquainting them with his study and asking for their assistance in contacting students to be used as S.'s. It was agreed that the teachers would
announce to the students that their cooperation was desired for a research project. E. felt that it was important for all S.'s to be similarly oriented, and presented each teacher with a mimeographed explanation of the research which was to be read to the class, and information forms to be filled in.* Despite this precaution E. learned that some teachers had explained the study in their own words. In order to correct any misconceptions and to present the study in a standard manner to all S.'s, E. explained the study to each S. during the initial telephone contact and again at the group meeting.

At the beginning of the Fall semester, seventeen teachers distributed forms to students in twenty-eight classes. At the beginning of the Spring semester, twenty-two teachers distributed forms to students in thirty-one classes.** A total of over two thousand forms were distributed by the teachers, filled out by the students, and returned to E. Cooperation on the part of all individuals concerned was splendid and many completed forms were returned at the end of the first class meeting. The lengthy and complicated process of matching then began.

As the first step, the groups described earlier - women, foreign students, etc. - were eliminated and the re-

*A copy of this form, which was titled "Confidential Data Sheet," appears in the Appendix, page 164.

**The writer is indebted to these many members of the faculty for their generous cooperation.
mainder were classified according to a coding system devised by E.*

E. was fortunate in gaining the cooperation of members of the administration who made official records available so that grade point averages could be computed for matching. Hundreds of grades were computed according to the official formula. The final step in matching S.'s on paper was the scheduling of two groups of matched individuals at times stated as free on the information form.

E. next telephoned each S. and explained that the purpose of the research was to study the group problem-solving discussion process scientifically, that it would prove an interesting and valuable experience to the S.'s, as well as being a contribution to research. E. stressed that participation was voluntary and that this was a year-long, campus-wide research which had the approval of the Deans and the Teachers. E. also stated that all data would be kept strictly confidential and that the results were to be reported in general terms. The S.'s were also told that the full purpose of the research and the procedures would be explained to them at the conclusion of the study. An average of thirty telephone calls was necessary in order to schedule one group meeting. Individuals without telephones were contacted personally between classes, so that

*The coding system appears in the Appendix, p. 165.
having access to a telephone would not be a "hidden" variable. Individuals who commuted were also contacted personally and included in the study.

About ten percent of the students contacted eliminated themselves from the study through lack of interest or competing activities. Of those who promised to attend a scheduled meeting about fifteen percent failed to appear. However, of those who attended the "before" session, only five percent failed to attend the "after" session due to drop-outs, refusals, and changed schedules. Every effort short of pressure was used by E. to convince the s.'s that it would be a worthwhile and interesting experience. E. felt that unless this was done, a selective factor might bias the results - that E. would have the more compliant or the better students. Apparently this biasing effect did not occur. The students as a whole showed considerable initiative and independence in their behavior, and asked pertinent questions concerning some aspects of the proceedings which indicated that they realized that they had some power in the situation. A survey of the grade point averages of all s.'s permitted E. to discount any bias on the grounds that only superior students were used - as defined by grade point average - since many students used in the study did not have high grades.

Throughout the year the teachers were kept informed of the progress of the experiments, and those who were interested were given a general report when the experiments
had concluded. After the last meeting in May, E. held a special meeting of all participants as he had promised. At this meeting, E. explained the goals and techniques of the research and answered questions. As the meeting concluded, five numbers were chosen at random by those present and the lucky winners each received a reward. E. had planned to surprise the S.'s with these rewards and therefore had not previously mentioned the matter.

E. did not ask the S.'s for their comments and opinions concerning the research procedures, although the S.'s were asked to indicate whether they found the problems interesting and to explain this. However, unsolicited reports indicated that in general the experiments had been well received by the S.'s. Several teachers reported favorable reactions, and in one class, several students contacted the teacher and demanded that they be permitted to participate. The fact that only five percent of the S.'s failed to return for the second meeting seemed to indicate good rapport. Attendance at the final general explanation meeting was good, and E. was frequently stopped and questioned by the S.'s concerning both the methodology and the progress of the research.
CHAPTER III.
THE EXPERIMENTAL SITUATION AND TECHNIQUES

The Experimental Setting

Barker and Wright have called attention to the coerciveness of the behavior setting - the place in which behavior takes place (7, p. 17). With this principle in mind, E. felt that discussion would be enhanced if it took place in an appropriate physical setting, one which looked as though it were suited to discussion. E. therefore chose to hold all meetings in a room which was regularly used for seminars and conferences. The room was furnished with tables, desks, chairs, and filing cabinets. It was adequately ventilated and lighted. All books and extraneous materials were removed before the S.'s appeared. S.'s within the room could not see outside, nor could outsiders look in. All meetings were held between one and five o'clock in the afternoon week days, and the same room was used for all twenty-four meetings.

The Seating Arrangement

E. sat at one end of the table, slightly removed from the groups so that physical distance might cause the S.'s to ignore E. However, the distance was near enough so that E. could hear and see each S. Distance also functioned to conceal E.'s kit of apparatus, which was placed on a chair below the level of the table and out of sight of the S.'s. This was done so that sight of the kit, which contained
pencils, forms, writing boards, etc., would not arouse speculation or suspicion concerning the use of these items. All articles of apparatus were placed in the order of use so that fumbling or loss of time would be minimized.

The S.'s were arranged across the table opposite each other, three facing three. This arrangement has been found optimum by several investigators who have studied leaderless group discussion. Mandell states:

"...no candidate should sit at the head of the table as this position may tend to give him an advantage" (69, p. 211).

Presumably the position at the head of the table carries implications of social status, or since it is a central location, it may prove a focus of attention, and these aspects may affect the discussion. Bass, who is one of the foremost American experimenters in the field of leaderless group discussion, has used the seating arrangement used by E. He experimented concerning the relationship between leadership rating given by members of the group and the seat occupied, in experimental situations in which groups of S.'s sat opposite each other at a table. His conclusion was:

"...the variations among individuals occupying a seat of a given type were so great that, despite the fact that the 150 persons occupying end seats on the average attained leadership scores approximately 5 per cent higher than the 317 occupying middle seats, the F. for the 7 analyses combined was only 1.6 - a value to be expected on a chance basis" (14, p. 725).
It was considered important that the S.'s should be able to identify each other, should they desire to refer to each others' communications, yet all were strangers. Faced with a similar situation, experimenters have used various methods of identification. French identified his S.'s by having them wear colored armbands (48). Leavitt also used a color scheme for this purpose (62). Festinger had his S.'s print their first names on cards (45). However, none of these methods were satisfactory to E. It was felt that armbands might prove disconcerting to upper-class college men. E. did not want the S.'s to print their names, since identical first names might prove confusing and since knowledge of names might provoke conversations between the "before" and "after" sessions. E. believed it most satisfactory to set up a number on the table in front of each S. The S.'s were therefore numbered clockwise around the table, from one to six.

The Use of Recording Equipment

On the basis of past experience, E. felt that knowledge that the meeting was being recorded need not necessarily alter the discussion. It is true that some individuals may remain silent, but inhibiting effects in general are most pronounced at the outset. As time passes, S.'s

"The recording equipment used was a Gray Audograph Disc Recorder and an Ekotape Tape Recorder. Both proved satisfactory."
become accustomed to being recorded and if the topic is interesting, they become so absorbed in the discussion that they apparently forget that they are being recorded. This diminution of inhibiting effect is akin to that noted by Barker and Wright. Their observers, who followed children around and noted their behavior, became almost ignored by the S.'s (7, p. 39). For this reason, there was no attempt made to conceal the recording equipment. Instead, the S.'s were given a five-minute "warmup" problem which was discussed in order to accustom them to speaking in the presence of a microphone. The fifteen-minute discussion problem followed immediately.

The Notation Technique

There are various methods of tracing statements to their source. One way involves playing back the recording to the group and asking the speakers to identify themselves. Another manner is for the E. to familiarize himself with the group so that he readily recognizes all voices. However, under the present experimental conditions, neither of these approaches was feasible. Another possibility would be to induce each S. to announce his name or number previous to the discussion. To the writer, this would be time consuming and a rather artificial procedure which might prove detrimental to rapport. Moreover, it would seem of dubious validity, for under conditions of intense involvement, changes in pitch and rapidity of speech often render familiar voices indistinguishable.
After preliminary exploration and experience in observing "case-discussion" method courses, E. had refined a simple, but satisfactory system. It involved the careful sequential notation of each individual who spoke by his identifying number, and a key word or phrase. When the recording was played back, this system proved highly reliable and it became a matter of patient listening in order to complete each statement. Moreover, with each group of individuals always arranged clockwise in numerical order, the possibility of mistaken notations was virtually eliminated.

The Leaderless Group Discussion Technique

The leaderless group discussion technique involves a group of S.'s who are seated around a table and given a problem to discuss as a group. No member is designated as a leader, moderator, etc., the group is not formally or informally structured in any way before the discussion begins. Thus, the group evolves its own interaction structure in the course of the discussion. According to Ansbacher, this technique was first used in Germany by military psychologists and was originated by J. B. Rieffert in 1925 (1, p. 385). It has been called Schlusskolloquium or Rundgespräch. It was used intensively by German military psychologists during World War II in screening officer candidates. Harris explains in detail how this technique was used by British military psychologists in choos-
ing key personnel for important positions (53). Another writer gives a detailed résumé of its use by British, Australian, and American civilian and military establishments and the various contributions to the leaderless group discussion technique which these experimentalists have made (12). These three writers agree that generally a team of experts, composed of psychologists, psychiatrists, military officers, or company officials observes these leaderless group discussions and assesses each candidate. The criteria varies according to the position for which the candidates are being chosen. However, the following criteria are usually mentioned: ability to communicate clearly, ability to reconcile differences, ability to draw out non-contributing members, responsiveness to feelings of individuals, clarity of reasoning, and appropriateness of solution. It has been generally agreed that the leaderless group technique is a promising method of assessing the individual's social skills in interpersonal relations. Admittedly, paper and pencil tests have not proved adequate to measure such skills.

Another problem involved in the use of leaderless discussion groups is the factor of previous acquaintance of the members. It will be recalled that E. used individuals not previously acquainted. Bass has experimented with leaderless discussion groups composed of individuals previously acquainted and individuals who were strangers.
Bass reports:

"...leaderless group discussion among strangers is as effective for predicting future outside leadership status as is the leaderless group discussion among friends" (17, p. 131).

One may well ask how the S.'s react to leaderless group discussion. The writer has read many descriptions of leaderless group discussion experiments and has found universally favorable comments by the participants, as well as by observers who are often prospective employers. Bass sums up this view by stating:

"One of the peculiar qualities of the leaderless group-discussion technique seems to be that everyone likes it" (11, p. 206).

E. had several important reasons for using the leaderless group discussion technique. In the first place, the "case-discussion" method S.'s were accustomed to the teacher playing different roles, such as moderator, clarifier, summarizer, etc. Had E. assumed any of these roles, these S.'s would have perceived this situation as the familiar classroom pattern carried into a different setting. These S.'s would have felt more at ease, and this probably would have influenced their discussion. On the other hand, the non-case-discussion S.'s would perceive this as a new situation and react differently, since they would not be able to anticipate or understand E.'s role-playing behavior. Thus, because of different past experiences, each group would have perceived the situation in a different manner.
and this would have confounded the results. Moreover, any participation in discussion by E. would have necessitated a standardized procedure, the same remarks would have had to be made in the same sequence for all groups. This would have been impossible to manage, due to the limitations imposed by the methodology, for E. was busy taking notes throughout the discussion.

The Selection of the Discussion Problems

At the "before" sessions, a "warmup" case was used in order to accustom the group to discussion and the presence of a microphone. This problem was called, "The Coach's Dilemma." It was written four years ago by E. and is the story of a star basketball player who disobeyed a new coach. E. had found this problem useful for promoting general class discussion in introductory psychology courses.

The second problem used at the "before" sessions was Wertheimer's, "Two Boys Play At Badminton" (83, pp. 124-136). This problem involves a game between two boys. The older and better skilled player is constantly defeating the younger and less skilled opponent. Wertheimer reported that he had used this case as a medium for exploring group problem-solving in the classroom. E. changed the story slightly, substituting tennis for badminton, as the former sport is more popular on the campus.

*The problems appear in the Appendix, pp. 166-168.*
At the "after" sessions, only one problem was used. This problem originated with N.R.F. Maier. It involves an industrial situation in which the failing dexterity of an elderly and faithful worker has caused a bottleneck in production which threatens the entire plant. E. called it "The Assembly-Line Problem." Maier has used this case repeatedly in his industrial management training program, as a medium for group discussion with and without leaders (67, pp. 301-326).

The Criteria for the Discussion Problems

The three cases which were used for discussion were carefully chosen to conform to the following criteria:

a. The problems must be capable of solution in accordance with the principles followed in the "case-discussion" method courses. These problems could be approached in the same manner as cases used in these courses.

b. The problems must involve a disturbance in interpersonal relations, for this is typical of the cases used in "case-discussion" method courses. These problems involved this aspect.

c. The problems should be of sufficient intrinsic interest to generate and insure continuance of the discussion. This criterion was apparently met, for only a few groups spontaneously decided that they had solved the problem before the fifteen-minute discussion period ended.

d. The problems should not be too narrow or encapsulated, but should involve social issues that go beyond the specific problem situation. These problems all contained far-reaching social implications.
e. The problems should be concise so that they could be retained mentally without reference to the printed problem, yet contain material sufficient to support a fifteen minute discussion. This criterion was apparently met, since the S.'s rarely consulted the problem except when in disagreement over a fact.

The Pilot Study

After the experimental procedure had been formulated, E. tried it out in a pilot study in order to become familiar with the routine and to discover whether the methodology would prove adequate. In early September, E. visited a freshman dorm, announced he was formulating a research project and asked for volunteers. A six-man group was formed and a complete group discussion was conducted. E. chose freshmen for the pilot study because no freshmen were to be used in the forthcoming experiments. Even if these S.'s discussed the experiment despite having promised not to do so, there would be little danger of contaminating potential participants.

The meeting proceeded according to the complete and detailed methodology that had been formulated. The entire meeting was recorded so that E. could play it back and check his presentation of the material as well as test the efficacy of his notation technique, as a means of tracing each statement to its source. The pilot study proved useful in that it demonstrated the adequacy of some procedures and led to some fruitful modifications. Apparently, motivation to discuss these problems as a lead-
erless group had been high, for the discussion occupied
the efforts of the S.'s for the full fifteen minutes. The
S.'s ran true to form concerning the inhibiting influence
of the microphone, as this wore off rapidly. It was found
that ten minutes proved ample for the independently written
solutions. The notation technique proved adequate for trac-
ing statements to their source.

However, E. had to rearrange the tools in his experi-
mental kit, so that problems, writing boards, and forms
could be produced with aplomb. E. corrected one glaring
error which came to light in the pilot study. In playing
back the recording, E. was surprised to note that in read-
ing a problem aloud, he had unintentionally emphasized the
viewpoint of one of the persons involved in the problem.
Evidently this had structured the discussion, for the S.'s
began the discussion by concentrating on the viewpoint of
this individual. Therefore, for future meetings, E. had
the problems typewritten and read by the S.'s themselves,
to rule out any biasing factors of this nature.
CHAPTER IV.
THE EXPERIMENTAL PROCEDURE

Preliminary Considerations

As was previously stated, the present study attempted to determine whether skills learned in a "case-discussion" method course could be transferred. For this reason, the experiments were structured as "Problem-Solving Conferences," in order to avoid reference to the classroom on the verbal level, so that the meetings would not be perceived as a variation of classroom routine. It is true that the present study did not entirely achieve the ideal transfer situation, for the experiments were carried out on the campus, and this probably influenced the S.'s behavior to some extent. However, this was a necessary limitation.

Even though the meetings were held on the campus, the physical setting - the experimental room - had no resemblance to any classroom from which the S.'s had been drawn, since the meetings were held in a conference room. Moreover, E. attempted to create a different psychological atmosphere from that prevalent in the classroom. For example, one minor, but important consideration was that in the experimental situation, the S.'s were permitted to smoke. Regulations forbid smoking in classrooms.

E. also attempted to avoid forming any associations for the "case-discussion" students between the conference
room and the classroom. E. did this by carefully avoiding the lingo of "case-discussion." For example, "loaded" words such as "case" or "case-discussion" were not mentioned. Moreover, from the very beginning the experiments were described to everyone as "Problem-Solving Conferences" as an attempt to forestall such associations. It was for this reason that E. read all his directions during the experimental sessions, since E. was in the habit of using these "loaded" words. However, despite these precautions E. does not doubt that some S.'s made the association and saw the experimental situation as an extension of class routine.

As has been previously mentioned, E. attempted to structure the situation in advance by having the teachers present the study to all S.'s in the same manner, but did not succeed. Experience in counseling had demonstrated to E. that the kind of referral an S. has had, largely determines his perception and expectation of the situation. E. therefore structured the situation to each S. in the same manner during the initial telephone contact and again at the group meetings.

The Experimental Procedure

E. sat at the head of the table. As each S. entered, E. said, "Hello, I'm Mr. Ostlund, your name is --?"

"Jack Blank."

"Thanks, I want to check you on the list. Here's a magazine to look at till the others arrive. Sit wherever you like, smoke if you wish, just make yourself at home."
Each S. was given a copy of a popular news magazine to make time pass quickly and pleasantly, and to discourage conversation with other S.'s, since it was desirable that they remain strangers. E. also read a magazine in order to inhibit conversation. E. did not want to create any experimental relationship with the S.'s, as this might lead some other S. to believe that E. was acquainted with the conversing S. Even between sessions, E. avoided conversation with S.'s.

As has been previously mentioned, supernumeraries were invited in order to insure a full complement of six S.'s. When the matched group had arrived, E. read out the names of the supernumeraries, and asked them to accompany him to another room. Before leaving the room, E. explained to everyone that both groups were going to work on the same problems, but separately. E. felt that if the supernumeraries had been dismissed, they would have become disappointed, and would have communicated their feelings to other S.'s to the detriment of the study. Their data were not used.

Returning to the group, E. asked for the magazines, placed a "No Admittance" sign on the door and seated the S.'s at the table according to the plan previously discussed. S.'s one, two, and three were seated opposite S.'s four, five, and six. E. said, "Smoke if you like, make yourselves at home. The purpose of these problem-solving conferences is to study the process whereby a group of people get together and solve problems. Unfortunately, not much is known about
about this process, but you can all appreciate the importance of problem-solving in school, business, and everyday life."

"The results of these conferences will be reported in general terms, your names will be mentioned only as cooperating participants, what we say and do here is strictly confidential. Our procedure is this - I give you a problem to read, then you discuss it together. The discussion is being recorded because I can't write as quickly as you can talk. I will just take brief notes to keep track of the discussion."

"Here is the recorder. I am placing this cardboard carton in front of it, not to conceal it, but because the light reflecting off the revolving disc is distracting."

(Actually, this was a ruse, E. did not want to record the "warmup" case.)

E. then handed out white cards numbered from one to six, and said, "Just place the number on the table in front of you. These numbers are for identification. In case you want to directly refer to anyone, you can say, 'I agree with you, Number Six.' They're for your convenience, you don't have to use them. Speak freely, say anything you please, no censorship. Keep working on the problem and ignore me."

E. then handed out "The Coach's Dilemma," and said, "Here is a problem, read it."

While they read it, E. pretended to turn on the re-
corder. When all appeared to have read the problem, E. said,

"O.K., let me have the problems. You can begin now. You will have five minutes on this one."

E. set a timer to ring at four minutes. When it rang, E. said,

"You have one minute to go."

When time was up, E. said,

"That was fine, here is another problem."

When all had read the problem, E. said,

"You can pass the problems to me, you'll have fifteen minutes on this problem, you may begin now."

E. turned on the recording apparatus, set the timer to ring at thirteen minutes, and took notes during the discussion. When the timer rang, E. said,

"You have two minutes to go."

When time was up, E. said,

"I guess time ran out. That was a fine discussion. Now, pull your chairs out, turn around, and write what you consider the best solution to the problem."

E. handed out pencils and writing boards with papers attached to them. When the solutions had been written, E. said,

"Now write out briefly which problem you found most interesting and why."

When this had been done, E. said,

"Unfasten the staple and open the paper. This is a
rating scale." I'll read it aloud as you read it over, and illustrate."

When the rating scales had been filled out, they were handed to E. who checked them for completion, and said, "This group has done very well and we appreciate your cooperation. You all understand that we are to meet again at the end of the semester in order to complete this research. Here is a notice of the meeting. However, since other groups suggested it, I will call each one of you two days before the meeting to remind you. In the meantime, I must ask you not to discuss this among yourselves or with anyone else, as many other students are participating and prior knowledge would invalidate the results. Let me repeat, this material is confidential."

The usual elapsed time for a group meeting was approximately fifty minutes, distributed as follows:

- Introductory remarks: 5 minutes
- Problem I: 5 minutes
- Problem II: 15 minutes
- Written solution: 10 minutes
- Rating scale: 10 minutes
- Concluding remarks: 5 minutes

Three months later, the same S.'s were experimentally tested again. The location was the same, the procedure modified in several aspects. The session began with a fifteen-minute discussion of a different case, "The Assem-

*A copy of the Rating Scale appears in the Appendix, P. 169.
At the conclusion of this discussion, the S.'s wrote individually what they considered the best solution. Then they were asked to state whether they considered the problem interesting or not and why. The next procedure involved the introduction of a new response measure, a "Communication Data" form. The S.'s next filled out the "Rating Scale" as at the previous session. At the conclusion of the meeting, E, thanked them and invited them to attend a meeting of all S.'s at which time the experiment would be explained.

None of the twenty-four meetings was interrupted or interfered with in any way. Once begun, all groups proceeded according to the standard methodology. The first groups met in September, 1952, the last in May, 1953. A summary of all the data which were collected appears on the next page.

---

*A copy of this problem appears in the Appendix, p. 168.*

**A copy of this form, "Communication Data" appears in the Appendix, p. 170.**

***A copy of the "Rating Scale" appears in the Appendix, p. 169."
## TABLE II.
### SUMMARY OF THE DATA

A. For each S.  

1. An information form, including age, class, etc. ........................................... 72
2. An independently written solution, following the discussion of "The Tennis Problem" ("before" session) ...... 72
3. An independently written solution, following the discussion of "The Assembly-Line Problem" ("after" session) .................................................. 72
4. A "Communication Data" form (used at the "after" session only) ................. 72
5. A "Leadership" scale (used at both sessions) ........................................... 144
6. An opinion concerning "The Tennis Problem" versus "The Coach's Dilemma." .................................................. 72
7. An opinion concerning "The Assembly-Line Problem." ........................................... 72

B. For each Group.  

8. A fifteen-minute transcript of discussion ........................................... 24

*The opinions were requested during the experiment as was related on p. 53 and p. 55.*
PART II

THE ANALYSIS OF THE DATA
CHAPTER V
SOME QUALITATIVE OBSERVATIONS

A. Rapport

It is pertinent to inquire about the degree of rapport established between E. and the s.'s. Rapport usually enhances experimental results, for properly motivated s.'s are more apt to do their best, and thus the data will prove more representative of the population under investigation.

The experimental procedure had been planned to proceed smoothly so that the s.'s would have confidence in E. The schedule had been timed to run for approximately fifty minutes. It was believed that a longer interval would unduly tax attention and motivation, or would be considered too great a drain upon the s.'s free time. At the initial contact, the experiment was presented as an opportunity to take part in important, campus-wide, year-long research. In order to further rapport, E. pointed out to the s.'s that he had chosen not to record the sessions surreptitiously. E. also stated that the s.'s opinions of the problems would guide future usage. E. believed that this approach would serve to convince the s.'s that they were not being whimsically manipulated. At the conclusion of a group discussion, regardless of the quality of performance, every group was told they had done a fine job.
During the twenty-four sessions, no S. behaved in such a manner as to oppose or disturb the on-going schedule. Secrecy was accepted as a necessary experimental condition. E. had carefully explained from the very beginning that he would be unable to divulge any results or explain procedures until the conclusion of the study. E. capitalized upon this to impress the S.'s with the serious purpose of the study.

The S.'s obeyed all instructions and followed orders with reference to filling out the forms, which was another verification of rapport. The instructions had been formulated briefly and clearly, and E. invited further questions. All queries were handled in a positive fashion, by phrases such as, "That is an important question," "That is a good question - every group asks me that," etc. Because questions were numerous, few errors resulted. These were noticed by E. as all forms were checked immediately after completion and returned to the S. for the correction of errors of commission and omission. The result was that each S. accurately filled out one information form, one communication form, and two leadership scales, problem solutions, and problem opinions, a total of five hundred and seventy-six responses.

A perusal of the written solutions was further testimony for rapport, as no silly solutions were offered. The S.'s tried to cope with the problems in a serious frame of reference. It was true that there were moments of levity
during the group discussions, but to E. this was added grounds for believing that rapport had been established, in that the S.'s evidently felt free to become humorous.

The fact that one hundred and three S.'s voluntarily came to two one-hour sessions speaks volumes. There were no adverse comments reported by teachers from whose classes the S.'s had been drawn. One teacher reported that several of his students approached him and demanded to be permitted to participate. From these various indications, it seems reasonable to assume that we can proceed to an analysis of the data, since they were apparently gathered under optimum conditions.

B. Solutions That Occurred Within the Time Limit

It will be recalled that the S.'s were given the problems in the form of a typewritten sheet, at the bottom of which appeared the following phrase, which was intended to focus the discussion: WHAT IS THE BEST SOLUTION? As discussion began, the S.'s were told that they would have fifteen minutes for group discussion. At thirteen minutes, a bell rang and E. announced, "Two minutes to go." If the discussion was still under way at fifteen minutes, E. announced that time had run out, and this ended the discussion.

E. believed that the typewritten phrase, which had been capitalized for emphasis, and the expressed time limit, would serve to set the task clearly - to discuss and solve the problem within the time limitation imposed.
E. therefore expected a fair number of solutions, but this did not materialize. All twelve discussions of "The Assembly-Line Problem" ran out of time. Four groups, out of the twelve discussions of "The Tennis Problem" solved the problems in the following times:

"BEFORE" SESSIONS

<table>
<thead>
<tr>
<th>Group</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>#2......10 minutes</td>
</tr>
<tr>
<td>Experimental #1</td>
<td>#1......10 minutes</td>
</tr>
<tr>
<td>Experimental #2</td>
<td>#2......12 minutes</td>
</tr>
<tr>
<td>Experimental #4</td>
<td>#4......11 minutes</td>
</tr>
</tbody>
</table>

The pattern of behavior in groups which solved the problem was similar for all four. Toward the end of the discussion, conversation lagged, pauses occurred, statements were repeated, and concurrence was evident. One or more individuals summarized the group's decision to the group. This was followed by statements of agreement by other members. The group then looked at E. and one member said, "I guess that's it. Do we have to keep on talking?" E. responded, "It's up to you. If the group feels they have solved the problem, that's O.K." The group then declared they had solved the problem, so E. ended the discussion.

The fact that no groups solved "The Assembly-Line Problem" suggested that this problem was found more difficult than "The Tennis Problem." It was true that the for-
mer problem involved more interpersonal relationships and also called for simple arithmetic in helping to understand one of the production problems involved. However, because the problem was more complex, E. tried to take this into account in several ways. E. permitted the S.'s to make notes and to keep the problem before them during the discussion. Moreover, after the S.'s had initially read the problem, E. used a diagram of the work situation to explain further details. Before group discussion began, the S.'s were encouraged to question E. pertaining to any aspect of the problem that remained unclear. The pre-discussion period lasted generally from ten to fifteen minutes. E. delayed discussion until all S.'s had signalled they were ready to begin. In contrast, "The Tennis Problem" was read and almost immediately grasped, without any explanation by E.

C. The S.'s Opinions Concerning the Problems

E. questioned the S.'s concerning their opinions of the discussion problems for several reasons. In the first place, it might assist report. If the S.'s believed that their opinions were valuable to E., they might feel more motivated to cooperate in other aspects of the experiment. The procedure might also prove a valuable clue to differences in motivation which might affect the discussion. For example, if discussion on any one problem was different from discussion on the others, the cause might be found in the
sentiments of the S.'s with relation to the problem. It would also provide a means of discovering whether there were important differences between the opinions of the experimental and the control groups. Finally, it was hoped that the result could be applied to a recurrent problem of case-discussion teachers - the never-ending search for problems of sufficient interest to sustain classroom discussion.

Therefore, after the S.'s had written what they considered the best solution, they were asked briefly to state their opinions concerning the cases. One hundred and forty-four opinions were gathered in this manner. These will now be considered.

The opinions were tabulated and categorized according to the problem preferred and the reasons given. The most striking outcome was the concomitance of opinion shown by both experimental and control S.'s. In the matter of the problem preferred when choosing between "The Coach's Dilemma" and "The Tennis Problem," the preference of both groups was similar. "The Coach's Dilemma" was preferred by 69% of the control S.'s and by 64% of the experimental S.'s.

Among the minority who favored "The Tennis Problem," such a variety of reasons was given that no particular trend was discernible. However, some generalities did emerge from an analysis of the opinions regarding "The Coach's Dilemma." Both experimental and control groups
agreed that it was important since it involved a recurrent problem of group interpersonal relationships. Furthermore, it was described as "serious," "important," and "complex" by both groups.

Because only "The Assembly-Line Problem" was discussed at the "after" session, the S.'s were asked to state whether or not they found the problem interesting and to explain their answer. There was a fair degree of accord between both groups. Eighty-two percent of the experimental S.'s judged it "interesting," as did 92% of the control S.'s. There was a tendency toward concurrence between both groups that this problem was interesting because it involved a plausible, contemporary industrial problem that had far-reaching social implications. It was frequently designated as "complex" and "real."

However, in terms of one of the important learnings that can be gained from case-discussion courses, there was one interesting difference between the groups. The criticism that the problem was not presented in sufficient detail was voiced by 11% of the control group and by 30% of the experimental group. These S.'s evidently desired a more clearly structured problem.

The writer would argue that one of the goals of case-discussion courses is to reconcile the student to the fact that he may have to make decisions on the basis of meagre information. Faced with such a condition, the individual must make the best decision in view of the limitations of the situation, with the realization that the decision may
not prove successful. Many case-discussion teachers share this viewpoint.

On the basis of this assumption, the case-discussion S.'s have apparently been less inclined to accept the limitations. However, an alternative interpretation presents itself. It is possible, that because of their case-discussion training, these S.'s better understand the complexities of interpersonal relationships, and realize the need for additional information. Thus their objection may be considered a reasonable one. By inference from this argument, it would follow that case-discussion S.'s are superior, in that they are more discriminating in recognizing the important shortcomings. Unfortunately the data did not permit us to resolve the issue in favor of either alternative. Moreover, the number of S.'s involved was so small that any interpretation would prove extremely tenuous.

It is interesting to note that the experimental and control S.'s were alike in their preference for "The Coach's Dilemma" and their judgment of "The Assembly-Line Problem" as "interesting." It demonstrates that the members of the experimental and control groups who had been matched on the variables already mentioned, have also held similar views relating to their interest in the problems. It may be that similar sentiments, with respect to the cases chosen for discussion, can be considered as another criterion on which these S.'s had been matched, though this was unknown to E. This criterion is probably related to
some of the other variables on which the S.'s had been matched. However, the nature of the relationship is not known.

D. Charting Dimensions of Groups

In keeping with the contemporary trend in social psychology, there has been an attempt to systematize the various dimensions of groups into a conceptual scheme. A survey of the literature reveals that, in general, the different dimensions of groups have been classified into three broad categories as follows:

1. Characteristics of the members who make up the group, such as sex, age, etc.

2. Characteristics of the process of group interaction, such as group climate, group structure, etc.

3. Characteristics of group products, such as solutions, decisions, manufactured object, etc.

For example, French presents the following trichotomy as determinants of group productivity: (1) process variables, such as amount of contribution, (2) structure variables, such as cohesiveness of the group, and (3) cognitive structure variables, such as perception of the abilities of other members (47, p. 53). Another schema, that of Cattell, will be more fully elaborated because it is more closely related to the three broad categories previously mentioned.

Cattell describes what he call three "panels" or as-
pects of group characteristics, namely population, structural, and syntality characteristics, and explains them as follows:

1. **Population** - characteristics of the population used in the study, for example, intelligence, race, age, etc.

2. **Structural** - characteristics of the internal structure and the interaction of the group, for example, leadership, interaction, etc.

3. **Syntal** - characteristics that emerge due to the performance of the group as a whole, for example, any group product" (38, p. 18). (In Cattell's terminology, "syntality" is descriptive of the group, in the same manner as personality is descriptive of the individual.)

These examples by French and Cattell illustrate the tendency toward a more systematic approach to the problem of group dimensions. Returning to our three broad categories, the population characteristics have already been discussed under the section relating to the matching criteria. Our next goal is to describe the evaluation of some group interaction variables, and then we shall proceed to an evaluation of some group productivity variables.
CHAPTER VI.

THE INDEX OF PARTICIPATION

Equality of Participation

In order to study equality of participation, the writer used a measure called the "Index of Participation." It was developed by Findley. Concerning evenness of participation as one criteria of an effective discussion, Findley states:

"Other things being equal, a discussion in which all members participate frequently is considered good. On the other hand if one or two members of the group monopolize the discussion so that the others get little opportunity to speak, the discussion is considered poor" (46, p. 47).

The "Index of Participation" was developed in order to measure group discussion at the Air University. Findley reports that the measure appears to be a valid index as well as a useful descriptive device, and cites supporting evidence of a qualitative nature. In one instance, the "Index of Participation" ranked the instructors in the same general order as did ratings by their supervisors. In another instance, the order of excellence of discussion and the indices of participation had a positive correlation (46, p. 50).

---

*The formula appears in the Appendix, page 171.*
Findley has suggested the following scale of interpretative values:

- Above 90 - Excellent, balanced discussion.
- 80-89 - Above average, well balanced discussion.
- 70-79 - Average, satisfactory discussion.
- 60-69 - Below average, poorly balanced discussion.
- Below 60 - Very poor, unbalanced discussion.

The Problem of Choosing Appropriate Units

The writer felt that an important problem in using the "Index of Participation" would be the choice of an appropriate unit of participation. As Guetzkow (52) has pointed out, the size of the unit to be used varies with the experimenter's purpose. For example, the unit may be elastic, such as that used by Bales, who describes his unit as:

"The unit to be scored is the smallest discriminable segment of verbal or nonverbal behavior" (6, p. 37).

However, further reading reveals that this unit may be as small as a single word, or as large as an entire sentence. The unit may be very small, as was used in the study of speech behaviors. By the use of vowel and consonant sounds, Irwin and Chen (55) were able to discriminate patterns of infant speech which they were able to relate to developmental levels. Larger units have been used. For example, Heyns, in his analysis of typescripts of conferences, used a more global approach. His behavior unit was the conference as a whole, and this was rated in terms of the extent to which the conference had functioned to solve the problem (52). French also used a wholistic aspect of
of behavior which was termed "interdependence of groups" as his unit, in his study of organized and unorganized groups under experimentally created conditions of fear and frustration (48).

A survey of the literature concerning categorization of data from discussion groups similar to those employed in the present study revealed that Bovard (24, p. 535), Festinger (44, p. 276), and Steinzor (78, p. 109) had used the entire statement made by an individual as the unit.

Findley's computations were likewise based on the entire statement made by an individual, as the unit of measurement. However, he tentatively advanced a further refinement, stating that a colleague had weighted the statements according to length of time, and that this gave a slightly higher index than frequency alone (46, p. 50). The writer tested for differences in Index, due to differences in unit by computing Indices of Participation on sheer frequency as well as on weighted frequency. A comparison of weighting procedures follows:

<table>
<thead>
<tr>
<th>Findley</th>
<th>Oulund</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td>Weight</td>
</tr>
<tr>
<td>Time</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>less than 30 sec.</td>
<td>24-48 words</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>30 - 60 seconds</td>
<td>over 48 words</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>over 60 seconds</td>
<td></td>
</tr>
</tbody>
</table>
The "Index of Participation" for two groups was computed by both procedures, the results appear below:

<table>
<thead>
<tr>
<th></th>
<th>&quot;Index of Participation&quot; (frequency)</th>
<th>&quot;Index of Participation&quot; (weighted freq.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C#3 (&quot;before&quot;)</td>
<td>92.</td>
<td>92.</td>
</tr>
<tr>
<td>E#3 (&quot;before&quot;)</td>
<td>51.</td>
<td>52.</td>
</tr>
</tbody>
</table>

Because of the high correlation between both methods of computation, E. chose to compute this measure by frequency alone - the number of statements made by each individual, for the twenty-four groups.*

In a study by Wetherell (22, pp. 473-4), participation was measured in terms of word frequency as well as by means of the number of participations. When each of these measures was correlated with final examination scores, the resulting correlations were very similar, as the following figures indicate:

Correlations between number of words spoken by individual students and their final examination scores: .30 to -.04

Correlations between number of comments by individual students and their final examination scores: .26 to -.02

Peterman has experimented on the problem of intercorrela-

---

*The ranking of the "Indices of Participation" appears in the Appendix, p. 172.
tions between various measures of participation. He reported a high intercorrelation between the total number of words used, the total amount of time talked, and the total number of responses made by a participant (74).

Hypothesis and Results

Hypothesis: It was predicted that the experimental groups would have a higher "Index of Participation," indicating a more shared participation.

For this purpose, the twelve "after" sessions were ranked from the highest index to the lowest. The test applied was the Tau B. for untied dichotomous variables (60, p. 34, p. 44). The result was not statistically significant (P > .05). Our hypothesis was disproved, the groups did not differ in participation as measured by Findley's "Index of Participation."

Examination of these indices revealed an interesting trend, in the direction of a more shared participation by the "after" groups as a whole, as compared to the "before" groups. These figures were statistically analyzed by means of the "t" test for correlated groups. The result was not significant (P > .05).

Since this trend to a more shared participation occurred at the second meeting, it may be indicative of a lessening of social restraint. We may assume that those individuals who were less communicative at the first
meeting felt more free to participate at the second meeting. Becoming accustomed to the situation functioned to lessen social restraint, and a more shared participation resulted. This trend is in keeping with Doob's comment concerning French's experiment on fear and frustration. In the following quotation, the term "unorganized" was defined as meaning those groups whose members had not been previously acquainted with each other. This was true of all S.'s in the present experiment. Doob remarks:

"If these "unorganized" groups had returned to the laboratory a second time, their newly established habits would undoubtedly have been reactivated. Their initial embarrassment would have been less, and they would have utilized again those methods of cooperating that had proved satisfactory during the first session. It would seem that crises or novel situations are likely to create the need for reciprocal actions among the members of a group and that this need produces the tradition of the future" (43, p. 227).

According to the Index of Participation, differences between groups are negligible. It may be that no changes have taken place, or that changes have taken place which Findley's Index has not proved sensitive enough to indicate.
CHAPTER VII.

EXTENT AND AGREEMENT
CONCERNING PERCEIVED COMMUNICATION CHANNELS

The Communication Data Form*

The writer's preliminary attempt to chart communication channels was by the analysis of transcripts. This proved vague and unsatisfactory. It was very difficult to judge whether a statement was in reference to the immediately preceding statement or a statement made previously. Many statements were ambiguous concerning their intended destination, it was difficult to be sure whether the statement was intended to answer a single individual, or was directed toward several persons, or toward the group as a whole. After preliminary efforts proved fruitless, the attempt was abandoned. It was then decided to ask the S.'s concerning their communication behavior.

A "Communication Data" Form was devised by the writer. It was administered at the "after" sessions only, since its administration at the "before" sessions might have influenced the S.'s behavior at the "after" sessions. In the upper box of the "Communication Data" Form, the S. was asked to state those with whom he specifically tried to communicate verbally, and whether or not he believed

*A copy of the "Communication Data" Form appears in the Appendix, p. 170.*
that he had been successful. In the lower box the S, was asked to state those whom he perceived as specifically trying to communicate with him verbally, and whether or not these attempts had been successful. These forms were analyzed in three ways: Extent of Communication Channels, Agreement on Perceived Communication Channels, and Agreement on Perceived Successes.

**Extent of Communication Channels**

The extent of communication channels for each S. was obtained by computations involving the upper box of the "Communication Data" form and these were summed to yield a group score.*

**Hypothesis and Results**

**Hypothesis:** It was predicted that because of their training, the experimental ("case-discussion" method) groups would excel the control (non-case-discussion method) groups, that the experimental groups would score higher than the controls. Presumably, one of the important skills gained from "case-discussion" method courses is that it is important to attempt to communicate with many members of the group.

Inspection of the rank order distribution of the twelve "after" sessions revealed that differences between

*The procedure and the ranking of groups appears in the Appendix, pp. 173, 174.
groups were negligible, so no statistical test was applied. The hypothesis was disproved, there was no significant difference in performance between the experimental and the control groups.

Agreement Concerning Perceived Communication Channels

Agreement concerning perceived communication channels is important, for if two individuals agree concerning the state of their communication - whatever it may be - if other conditions are favorable, such as rapport and a desire to achieve a common goal, this agreement may form the basis for a productive interaction. At the very least, it may minimize confusion due to faulty communication. It should be stressed that the issue here concerns agreement on perceived communication - regardless of whether the individuals agree that they tried to communicate, did not try to communicate, etc. A weighted score for each group was obtained by checking the upper box of one S.'s form, with the lower box of another S.'s form.

Hypothesis and Results

Hypothesis: It was predicted that because of their "case-discussion" method experience, the experimental groups would score higher than the controls. Presumably, experience in "case-discussion" method courses sharpens the awareness of the individual to the importance of communication and skill in the accurate perception of communication.

*The procedure and the ranking of groups appears in the Appendix, p. 175.*
Inspection of the rank order distribution of the twelve "after" sessions revealed that differences between groups were slight, so no statistical test was applied. The hypothesis was disproved, there was no significant difference in performance between the experimental and the control groups.

**Agreement on Perceived Communication Channels - Successes Only.**

The only difference between this measure and the preceding one was that the preceding measure was concerned with agreement regardless of the state of communication, whereas this measure concerned the degree of agreement on perceived successful communication. This variable was measured by checking the upper box on one S.'s form against the lower box on another S.'s form.*

**Hypothesis and Results**

**Hypothesis:** It was predicted that because of their "case-discussion" method training, the experimental groups would score higher than the controls.

Inspection of the rank order distribution of the twelve "after" sessions revealed that differences between groups was inconsequential, so no statistical test was applied. The hypothesis was disproved, there was no significant difference in performance between the experimental groups.

---

*The procedure and ranking of all groups appears in the Appendix, p. 176.*
and the control groups.

Conclusion

In these three measures derived from the "Communication Data" Form, the hypotheses that the experimental ("case-discussion") groups would excel the control (non-case-discussion) groups were disproved. Several explanations for these negative results yielded by the three analyses of the "Communication Data" Form may be offered. It is possible that these skills are difficult to master, and that three months is not enough time for significant changes to occur. It is also possible that the measures which were used were not sufficiently sensitive to evaluate this aspect of group interaction.
CHAPTER VIII.
GROUP AGREEMENT CONCERNING "LEADERSHIP" HIERARCHY

Introduction

In the discussion bearing upon the measure, "Agreement Concerning Perceived Communication," it was stated that if two individuals perceive the state of communication between them in a similar vein, if other conditions are propitious this agreement enhances group discussion. A related observation is that if one individual perceives the "leadership" hierarchy of the group as the other members of the group do, this also enhances group discussion. It would seem that the individual and the group were in harmony concerning informal group structure. However, if one member's perception of the "leadership" hierarchy of the group were at odds with the group's perception, this variant perception would be suggestive of an opposed frame of reference. It would probably prove detrimental to group discussion since it would lead to behavior that would be at odds with the group, tending to isolate the individual from the group.

The Rating Scale*

The "Rating Scale" was designed by E. primarily to elicit information concerning aspects which were considered important to leadership in group discussion. However,

*Appears in the Appendix, p. 169.
it was possible to use this measure to study whether or not the individual perceived the "leadership" hierarchy as did the rest of the members of his group.

Procedure

Working with the data from the "before" session the score given each member of his group by S. #1 was obtained. Then, the mean score of each other member of the group, as given by the group as a whole, was determined. Both sets of scores were then ranked, and a Spearman Rank Order Correlation Coefficient was computed, and transformed into a \( z \) score. This same procedure was followed with the data from the "after" session. This necessitated six correlations per group, a total of one-hundred and forty-four correlation coefficients. The two \( z \) scores - the "before" and "after" session scores for each individual - were then compared, and "gain" scores computed. The six "gain" scores for each member of a group were averaged in order to ascertain a "group average gain score." The twelve group "gain" scores were then ranked.*

Hypothesis and Results

Hypothesis: It was predicted that because of their "case-discussion" method training, the experimental groups would improve more in awareness of the informal structure of their groups, as indicated by their ratings of the "leadership" hierarchy.

*The correlations, \( z \) scores, and rank order appear in the Appendix, pp. 177-180.
Inspection of the rank order distribution of the "gain" scores indicated that there was little difference between groups, so no statistical test was applied. The hypothesis was refuted, there was no significant difference between the groups in this respect.

It should be emphasized that another factor enters into the perception of "leadership" hierarchy, namely, the range of "leadership" activity exercised by the members of the group. This is important, as the following example may illustrate. In some groups we may note extremes of leadership and non-leadership behavior. Under such conditions, it is not difficult to rank individuals along a continuum, because the behavior of the various members is clearly differentiated. However, in another group, leadership may be such a shared function, that members rate each other so closely together that a slight difference in position changes the ranking. It is possible to accurately determine whether this condition is operating by examining the range of leadership behavior. This point may be succinctly summarized by the statement: "A wider range of leadership behavior makes it easier to rank individuals with regard to this trait." In order to explore this factor, the range of "leadership" hierarchy was plotted for each of the "after" groups.*

*Appears in the Appendix, p 181.
The statistic applied was the Tau B. (60, p. 34, p. 44) to the rank order of the twelve "after" sessions. The result was not statistically significant (P.<.09). However, a marked trend is clearly demonstrated, that the experimental groups, as might have been predicted, tended to have a narrower range of "leadership" hierarchy than did the control groups. This suggests that the case-discussion groups might have learned to share leadership functions more equally than the control groups.

This finding makes it possible to understand the non-significant relationship between both groups in relation to the "gain" scores on the "leadership" hierarchy. As was previously suggested, when the range of activity is narrow, it is more difficult to rank individuals with reference to their behavior.

In conclusion, measuring the factor of "gain" scores as indicative of increased awareness of informal group structure led to the finding that there were no significant differences between the groups. However, further exploration of the relationship between the groups pertaining to the range of "leadership" hierarchy, served to somewhat mitigate this negative finding. Moreover, this further analysis revealed a positive finding, though it was not statistically significant, that "leadership" was perceived to be more equally shared among the experimental ("case-discussion") method groups.
CHAPTER IX.

THE WRITTEN SOLUTIONS

At the conclusion of each group discussion the members independently wrote what they considered the best solution and the reason why. This was done at both sessions by all S.'s. At the "before" session, the case used was "The Tennis Problem." At the "after" session, it was "The Assembly-Line Problem." This procedure resulted in one hundred and forty-four written problem solutions. These written solutions may be considered as productivity variables.

In order to facilitate reading of these solutions, each was typewritten on a separate page. All mistakes of grammar or punctuation were corrected so that such extraneous features would not bias the judges' decision and each solution was assigned a code number, the key to the code being known only to E. The solutions were then presented in random order to eleven judges. Each judge was presented with an envelope containing a set of problem solutions, the criteria, scoring forms, problems, and instructions.** E.

---

*The writer is indebted to the following individuals who served as judges: Graduate students: Dorothy Coddington, Kelly Wiltse, Charles Buchanan, Barry Crites, Charles Beeble, Granville Fairchild, Orville Harsh, Wendell Walker. Professors: Mary Kinnahe, H. K. L'Ecouyer, Martin Scheerer.

**This material appears in the Appendix, pp. 182-187. This includes some examples of actual solutions.
also verbally instructed each judge and remained with the judge while trial solutions were scored. This proved valuable in clarifying some nuances of the criteria.

The judges did not know what groups the solutions had come from, neither did they know which problem was used at each session. The judges were kept purposely unaware of any information which might influence their judgment. Each solution was judged independently by two different judges. Other than the initial orientation presentation to the judges, E. took no part in the judging procedure.

"The Tennis Problem"

"The Tennis Problem" was used at the "before" session. This problem originated with Max Wertheimer, one of the founders of Gestalt psychology. It was employed by him as a means of investigating problem-solving by students. To the writer, it proved a most fortunate choice in eliciting a variety of feelings and opinions concerning ways of solving problems in interpersonal relations. Wertheimer states that while the problem is simple, genuine attempts to cope with this problem have deep implications for human behavior. Wertheimer remarks:

"This was only a modest little story. Yet some of its characteristics, as for instance, the essential progress which occurred through a change of centering, are, I think, characteristics of profoundly important accomplishments in human beings in human society" (63, p. 135).

"Appears in the Appendix, p. 167. Wertheimer's title for this problem was "Two Boys Play at Badminton."
Wertheimer stated that it was his practice to tell this story to a class and wait for their reactions. The responses were interpreted according to some general principles and Wertheimer has explained these classifications in his book, *Productive Thinking* (83, pp. 124–136). This report of research provided valuable leads which were incorporated in the present study. The writer used Wertheimer's general framework for setting up different criteria in the form of categories, in terms of which the solutions were judged. For "The Tennis Problem" the writer compiled a set of three criteria, "centering," "inadequate solutions," and "means."

Rationale for the "Centering" Category

According to Wertheimer, a common error in problem-solving is the failure to properly "center" a problem. "Centering" here means focusing. Wertheimer argues that a narrow approach to "centering" inhibits adequate problem-solving because only one part of the total situation receives attention. Wertheimer's principle is that adequate problem-solving demands centering on the total situation.

Wertheimer explains centering as follows:

"Centering - the way one views the parts, the items in a situation; their meaning and role as determined in regard to a center, a core, or radix - is a most powerful factor in thinking."

"As to the concept of centering, it seems to be tacitly recognized that adequate centering, with its consequences for objectivity and justice, is of the utmost importance" (83, p. 136).
The following examples will illustrate the criteria under this category:

A. **Criteria Concerning Centering**

1. "A" (older boy) centered - a solution oriented in terms of "A's" (the older boy's) needs.
   Ex: "A should give B up as hopeless and find another partner."

2. "B" (younger boy) centered - a solution oriented in terms of "B's" (the younger boy's) needs.
   Ex: "A ought to teach B so that B would develop into a better player."

3. Game or "both" centered - a solution which involves the needs of both boys.
   Ex: "If A helps B to improve, B will provide A with more worthy competition, both boys will benefit, both will derive more satisfaction from the game."

4. "Outside of the field" solution - the problem given to the S.'s ended with the phrase "What should 'A' do?" Therefore, if some other individual was brought into the case, this would constitute an evasion of the problem as given, it would embody a violation of the stated conditions.
   Ex: "The person watching from the window should go down and settle it."

**Rationale for the "Inadequate Solutions" Category**

In his discussion of classroom research, Wertheimer pointed out that frequently the students handled the problem on a superficial level, such as labeling behavior in
stereotypes, pronouncing moral judgments, and the like. The writer considered that such behaviors represented pitfalls or errors in problem-solving and this category was set up in order to detect such errors. The different types will now be illustrated:

B. Criteria Concerning Inadequate Solutions

1. **Pigeon-holing** - involves applying a label or stereotype.
   Ex: "This is obviously a case of inferiority complex."

2. **Moralizing** - involves blaming or pronouncing moral judgments.
   Ex: "A should know better than to treat B like this."

3. **Ignoring restrictions, limitations, instructions, or adding unrealistic details.**
   Ex: "An 'Outside of the field' solution."

4. **Using a restricted frame of reference** - involves a one-sided view of the problem.
   Ex: "Either an 'A' or a 'B' centered solution."

5. "**Reinterpreting" facts** - taking a given statement and altering it so as to permit a certain kind of solution.
   Ex: "Maybe 'B' was just tired that day, maybe he'll beat 'A' next time."

Rationale for the "Means" Category

By employing this category, E. endeavored to determine the underlying means with which the S. was attempting to cope with the situation. This index of behavior is important because it tells us much about the S.'s basic view-
point toward interpersonal relations.

Wertheimer was concerned about the fundamental approach of individuals to problems in interpersonal relations. He described the typical modes of approach to "The Tennis Problem" and stated that the following were the usual types of suggestions offered by his S.'s:

"Promise the younger boy a piece of chocolate.

Start another kind of game, say chess, in which the younger is equal or even better. Or promise to play, alternately with badminton, some game in which he is by far the better.

Bring him to reason by plain scolding. He should be manly, not a sissy. Can't he take it on the chin? He has to learn to take it.

Use superior authority in bringing the younger boy to reason.

Don't bother with him, he is a sissy. It will be a lesson to him.

Offer a handicap.

Promise the younger boy that the older boy will not make full use of his superior power and skill."

(83, p. 123).

It will be noted that the categories which follow have been designed to encompass many of these suggestions which Wertheimer has reported.

C. Criteria Concerning Means

1. Bargaining - entails a "business deal."
   Ex: "A should tell B that he will help him to learn tennis if B will help him in something else."
2. **Bribing** - involves offering an inducement.  
Ex: "A should offer B some candy if he will get back on the court."

3. **Flattering** - involves using verbal cajolery.  
Ex: "A should flatter B, tell him he's pretty sharp and has plenty on the ball."

4. **Shaming** - involves using reproach.  
Ex: "A should tell B he's a sissy, that he's got to learn to take it."

5. **Mutual Means** - involves using the present and potential abilities of each in a manner beneficial to both.  
Ex: "A should coach B on B's weak points, and as B improves, he can work on A's weak points, each will be gaining through this cooperative endeavor."

6. **Withdrawal or quitting** - involves either boy giving up and walking off.  
Ex: "A should just walk off and leave B there, B is just plain hopeless."

This was the last of the three categories designed for recording the solutions to "The Tennis Problem."

It has been suggested that students who sign up for case-discussion courses constitute a special population, and that this is somewhat accountable for differences which may occur. This point of view argues that the student who takes such courses, seeks them out because he enjoys talking exchanging ideas, working on problems in interpersonal relations, etc. If this is true, one would expect to find that groups drawn from such a selective population would evince
some differences at the very beginning. In the present study, the results indicate this was not so, moreover, voluntary choice was ruled out by academic regulations. One of the courses from which experimental S.'s had been drawn was a required course.

Reliability of the Judges

In the present study, each solution was judged independently by two different judges. It is important to inquire into the degree of agreement between the judges. The data concerning the judges' agreement appears on the next pages, in Tables III, IV, and V.

Table III - Written Solutions: Number of Instances in Which Both Judges Agreed

Table III was compiled according to agreement per each category by both judges. If both judges agreed that the solution fell under the same category, an entry was made in that category. The results indicated that agreement was not very high by this method of computation, except in the "means" category of "The Assembly-Line Problem," page 90.

Absolute agreement by both judges per category is a very rigorous criterion, and does not reveal the entire scope of agreement. It is possible to analyze the data in other ways. Appropriate combinations of categories have been made in Tables IV and V, pages 92 and 96.
### TABLE III

**WRITTEN SOLUTIONS**

**NUMBER OF INSTANCES IN WHICH BOTH JUDGES AGREED**

<table>
<thead>
<tr>
<th></th>
<th>A center-</th>
<th>B inadequate</th>
<th>C means</th>
<th>N pairs of judges</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;The Tennis Problem&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>16</td>
<td>12</td>
<td>21</td>
<td>36</td>
</tr>
<tr>
<td>Experimental</td>
<td>21</td>
<td>11</td>
<td>18</td>
<td>36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>&quot;The Assembly Line Problem&quot;</th>
<th>A action</th>
<th>B centering</th>
<th>C inadequate</th>
<th>D means</th>
<th>N pairs of judges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>20</td>
<td>21</td>
<td>9</td>
<td>24</td>
<td>36</td>
</tr>
<tr>
<td>Experimental</td>
<td>14</td>
<td>23</td>
<td>10</td>
<td>29</td>
<td>36</td>
</tr>
</tbody>
</table>
"The Tennis Problem" - Judges' Agreement*

Criteria A - Centering

This criterion was interpreted in terms of whether the judges had agreed and scored the solution in category three, which represented the most adequate centering, or whether the judges agreed and scored the solution as one, two, or four, all of which were examples of inadequate centering. A disagreement occurred if one judge marked a solution in category three, and the other did not. Under these conditions, agreement was 58%.

Criteria B - Inadequate Solutions

All these categories were combined under the rationale that the score represented an inadequacy, regardless of the specific kind. An entry was made if both judges had checked at least one inadequate solution or if both checked none. Under these conditions, agreement was 65%.

Criteria C - Means

This criteria did not seem fitted to any combined approach. Agreement was measured in terms of the number of pairs of judges who checked a solution in the same category. The results are identical with those in Table III. Agreement was 55%.

*The data appear in Table IV, on the next page., 92.
# TABLE IV.

JUDGES' AGREEMENT
CONCERNING WRITTEN SOLUTIONS
TO "THE TENNIS PROBLEM"

Criteria A - Centering

| # of pairs of judges who judged the solution in category 3, or in categories 1, 2, or 4 | 42 |
| # of pairs of judges who disagreed (one scored it in category 3, the other did not) | 30 |

Agreement 58%

Criteria B - Inadequate Solutions

| # of pairs of judges who agreed and checked at least one inadequate solution, or checked none | 47 |
| # of pairs of judges who disagreed (one checked at least one inadequate solution, the other checked none) | 25 |

Agreement 65%

Criteria C - Means

| # of pairs of judges who checked the same category | 39 |
| # of pairs of judges who did not | 34 |

Agreement 55%
"The Assembly-Line Problem" - Judges' Agreement

Criteria A - Concrete Action Proposals

It will be remembered that the seven categories under this heading were grouped into three general classifications as follows:

1. Keeping Joe on the same work team.
2. Transferring Joe somewhere within the company.
3. Getting rid of Joe by discharge or retirement.

The specific categories were combined under these general classifications. Agreement was in terms of the general classifications. Even though each judge might have checked a different category, if these categories fell under the same general classification, it was scored as an agreement. Under these conditions, agreement was 94%.

Criteria B - Centering

This was scored in terms of whether the judges agreed and scored a solution in category three, or in category one or two. A disagreement occurred when one judge scored it in category three, and the other did not. Under these conditions agreement was 66%.

Criteria C - Inadequate Solutions

As was done with the data of "The Tennis Problem," all these categories were combined under the rationale that the score represented an agreement concerning inadequacy, re-

*The data appear in Table V on page 96.
Regardless of the specific kind, or agreement concerning adequacy. An entry was made if both judges had checked at least one inadequate solution or if neither judge checked an inadequate solution. Under these conditions, agreement was 57%.

Criteria D - Means

As in the case of "The Tennis Problem," this criterion did not seem fitted to any combined approach. Agreement was measured in terms of the number of pairs of judges who checked a solution in the same category. Agreement was 74%.

Conclusion

Interpreting the data in terms of combined categories of broader scope yielded a higher degree of agreement than when agreement was computed on the basis of absolute agreement per category. However, a higher degree of agreement would have been desirable.

The judges were all oriented by E. in a standard manner. The criteria were explained orally and by printed instructions. This procedure was intended to create a common frame of reference for all the judges to follow. Yet there has not been a high degree of agreement. It is possible that this precaution did not avail because the judges had differing frames of reference of their own, or that the judges interpreted the standard criteria in various ways, despite the structuring by E. A more probable explanation is that the material was rather complex and ambiguous. This
rendered it difficult to evaluate and susceptible to conflicting judgments. It may be that reducing the number of judges and increasing the training period, might have produced closer agreement. However, it must be pointed out that there were certain limitations inherent in the judging process. There were two hundred and eighty-eight judgments and these were distributed among eleven judges. A smaller number of judges would have been faced with a lengthier task.

Tabulation of the Judges' Ratings

The judges' ratings were tabulated by two methods in order to guard against differential results due to the method employed. Sample calculations were made in two different ways. The first method was to record an entry if either judge had scored it under that category. The second method was to make an entry only if both judges had scored it in that category. Trial calculations demonstrated clearly that the difference between these two methods were negligible as to final results. Therefore, in further calculations, scoring was done by the first method - an entry was recorded under the category if either judge had scored it in that category. The frequencies for "The Tennis Problem" appear on page 97, Tables VI, VII, and VIII.
TABLE V.
JUDGES' AGREEMENT CONCERNING WRITTEN SOLUTIONS TO "THE ASSEMBLY-LINE PROBLEM"

Criteria A - Concrete Action Proposals

<table>
<thead>
<tr>
<th>Description</th>
<th>Number</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td># of pairs of judges who judged the solution under the same general classification, although they may have used a different category.</td>
<td>65</td>
<td>94%</td>
</tr>
<tr>
<td># of pairs of judges who disagreed (one scored it under one general classification, the other did not).</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td># of pairs of judges of whom one omitted judgment.</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

N = 72

Criteria B - Centering

<table>
<thead>
<tr>
<th>Description</th>
<th>Number</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td># of pairs of judges who agreed and judged the solution in category 3, 1, or 2.</td>
<td>47</td>
<td>66%</td>
</tr>
<tr>
<td># of pairs of judges who disagreed (one scored it in category 3, the other did not).</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

N = 72

Criteria C - Inadequate Solutions

<table>
<thead>
<tr>
<th>Description</th>
<th>Number</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td># of pairs of judges who agreed and checked at least one inadequate solution, or checked none.</td>
<td>41</td>
<td>57%</td>
</tr>
<tr>
<td># of pairs of judges who disagreed (one checked at least one inadequate solution, the other checked none).</td>
<td>31</td>
<td></td>
</tr>
</tbody>
</table>

N = 72

Criteria D - Means

<table>
<thead>
<tr>
<th>Description</th>
<th>Number</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td># of pairs of judges who checked the same category.</td>
<td>53</td>
<td>74%</td>
</tr>
<tr>
<td># of pairs of judges who did not.</td>
<td>19</td>
<td></td>
</tr>
</tbody>
</table>

N = 72
TABLE VI.
FREQUENCIES OF FORMS OF "CENTERING"
"THE TENNIS PROBLEM"
IN THE
EXPERIMENTAL AND CONTROL GROUPS

<table>
<thead>
<tr>
<th></th>
<th>Centered</th>
<th>Other</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>10</td>
<td>26</td>
<td>36</td>
</tr>
<tr>
<td>E</td>
<td>7</td>
<td>29</td>
<td>36</td>
</tr>
<tr>
<td>Total</td>
<td>17</td>
<td>45</td>
<td>72</td>
</tr>
</tbody>
</table>

TABLE VII.
FREQUENCY OF ADEQUATE AND INADEQUATE SOLUTIONS TO "THE TENNIS PROBLEM" IN THE EXPERIMENTAL AND CONTROL GROUPS

<table>
<thead>
<tr>
<th>S.'s free</th>
<th>S.'s with one or more inadequate solutions</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>from</td>
<td></td>
<td></td>
</tr>
<tr>
<td>inadequate solutions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>7</td>
<td>29</td>
</tr>
<tr>
<td>E</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>59</td>
</tr>
</tbody>
</table>

TABLE VIII.
FREQUENCY OF "MUTUAL" MEANS FOR "THE TENNIS PROBLEM" IN THE EXPERIMENTAL AND CONTROL GROUPS

<table>
<thead>
<tr>
<th>S.'s who solved by &quot;Mutual Means&quot;</th>
<th>S.'s who did not solve by &quot;Mutual Means&quot;</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>33</td>
<td>3</td>
</tr>
<tr>
<td>E</td>
<td>32</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>65</td>
<td>7</td>
</tr>
</tbody>
</table>
Hypothesis and Results

Hypothesis: Since "The Tennis Problem" was the case used at the "before" session - at which time both experimental and control S.'s were theoretically at the same level - no significant differences will be found between the groups.

The results show a very close agreement between the groups for every category. This finding was so clear that it was not necessary to compute statistical tests for significance of the differences. Inspection alone revealed a great similarity between the groups.

Conclusion

Our hypothesis of no significant differences between the groups at the beginning of the experiment has been fully substantiated. This was favorable to the experiment, since it indicated that the groups had been properly matched and that at the "before" session, the fundamental approach to this kind of problem-solving by the experimental and control groups was the same.

The Problem-Solving Behavior of "The Average College Student"

While the primary purpose of the present experiment was to explore differences between case-discussion S.'s and non-case-discussion S.'s, the data are pertinent to another problem. It reveals information concerning the
behavior of the "average college student" in solving problems of interpersonal relations. It is legitimate to combine the data of both groups for this purpose, because at the "before" session, they represented a homogenous population. Neither group had case-discussion training. It was only later that they became differentiated in this respect.

Inspection of the "centering" criteria reveals that seventeen (24%) centered in terms of the total situation. The "adequacy of solution" criteria demonstrates that thirteen (18%) proposed solutions free from inadequacies. The "mutual means" category shows that sixty-five (90%) solved the problem in this manner.

According to Wertheimer (83, pp. 124-136) the most fruitful approach to problem solving in interpersonal relations lies in centering in terms of the total situation, which only one in four was able to do. He also stressed the principle that inadequacies such as moralizing, blaming, and others measured in the present study, hinder problem solving. Only one in five solutions was free from inadequacies. Wertheimer further stated that "mutual means" was the most productive solution, for it constructively uses the present and potential resources of both individuals. Nine out of ten were of this nature.

Individually, each of these trends is clear. Collectively, these trends appear contradictory, because only 24% use the most constructive approach to the problem, only 18%
propose solutions free from inadequacies, yet 90% use the most constructive means. Judges were questioned on this matter and stated that often a solution was oriented in terms of the needs of "B" the younger boy. However, the action proposed involved the potentialities of both boys, and thus it was judged in the "mutual means" category.

This apparent contradiction may be resolved as follows. The goal of the S.'s is to solve the problem which they propose to do by the most constructive means. However, their approach is too narrow and they fall into certain pitfalls.

Unfortunately, Wertheimer gives no specific data so that a comparison with the present results cannot be made. However, from Wertheimer's general description, unexperienced S.'s tend to demonstrate a narrow approach and to show many inadequacies in their solutions as did the S.'s in the present experiment. However, a striking difference is that 90% of the S.'s proposed to solve the problem constructively, whereas Wertheimer found this a rare occurrence. This difference cannot be explained, for Wertheimer stated that he used this problem in the classroom, but did not specify on what level.

"The Assembly-Line Problem"*

This problem was used at the "after" session. It originated with N.R.F. Maier whose psychological studies

*Appears in the Appendix, p. 168.
of problem-solving behavior of men and animals are classics of experimental methodology. This problem has been utilized in studying the problem-solving ability of college students and adults employed in industry, as an important aspect of an industrial management training program.

Maier stated that in order to highlight the possible value of the expert he selected this problem for his experiment because there is one solution to it which is definitely superior to others. Maier calls this the "elegant" solution, and explains it as follows:

"The solution regarded as having the quality of inventiveness or elegance was that of periodically having the men exchange places, progressing from one position to the next in a counterclockwise direction. In this manner the fast workers would reduce work piled up in positions occupied by slower workers, and at the same time variety would be introduced into the job and would make the work less monotonous. Such progressive changes would make production dependent upon the ability of the average man rather than on that of the slowest man. This is an objective fact if we assume there are no adverse attitudes. The special condition that permits this rotation is the fact that the work is simple, requires a minimum of learning, and demands similar aptitudes. The only problem that seems necessary to put the solution into practice is the willingness of the group to adopt it. This willingness, of course, will depend somewhat on the type of Joe that is present, but this is not
necessarily important, since the poor attitude on Joe's part might be due to his inability to keep up with the group.

An added feature of insight, which conceivably might accompany the solution, is the fact that the solution applies to all the other groups. Each one is paced by the slowest worker.

The argument against discharging the slowest worker is that it merely creates another "slowest worker." If discharge is practiced it will either lead to insecurity or a protection of the slow workers" (67, pp. 307-308).

Criteria

The rationale for the "centering," "inadequate solutions," and "means" criteria used in "The Assembly-Line Problem" were the same as those used for "The Tennis Problem," so they will not be repeated here.** There were minor differences in the categories under these criteria and these will be discussed at the appropriate place.

The main difference between the criteria for the two problems was the addition of the category "Concrete Action

---

*For experimental reasons which do not concern the present study, this information was a variable manipulated by Maier. In some of his experiments, Joe was described as being liked by his group, and as having the proper attitude. In others, Joe was described as having a poor attitude and being disliked by his group.

In the present study, this information was kept constant. Joe was presented as being liked by his group.

**The rationale appears on pages 84-86.
Proposals" to "The Assembly-Line Problem." This was done because it was one of the major devices used by Maier in evaluating the problem-solving ability of his groups.

Concerning the concrete action proposals, Maier stated that he usually received a variety of suggestions. As was the case with Wertheimer's S.'s, many of these proposals dealt with the problem on a superficial level. Maier stated that the "elegant" solution was difficult to achieve, and lists the usual types of solutions as follows:

"Transfer Joe to another unit.
Give Joe a talking-to and warn him.
Transfer Joe to a non-production job.
Retire Joe.
Let the group work it out.
Give part of Joe's work to others.
Divide the men into various units according to ability.
Put Joe in an easier position.
Discharge Joe" (67, p. 308).

Maier goes on to point out that many of the solutions offered did not remain within the limitations of the problem as stated. He specifies these limitations as follows:

"(1) the foreman is in charge of the unit and cannot influence other units.
(2) Joe cannot be transferred.
(3) Joe's primary difficulty is low aptitude in finger dexterity.
(4) the positions are equal and require the same aptitude" (67, p. 308).
The criteria "Concrete Action Proposals" have been set up especially to encompass these aspects of the typical solutions. The attempt was made to create a category which would serve to differentiate each of these various action possibilities.

A. Criteria Concerning Concrete Action Proposals

In the following description, the generic headings, which are denoted with asterisks (*) were contrived in order to facilitate judging and tabulation.

*Keeping Joe in the company, on the same work team.

1. Give the less capable less to do - these solutions recognize differences in ability and arrange for some way of permitting the more capable members of the work team to help the less capable.
   Ex: "Let Joe periodically change places with a faster man on his team."

2. Training Joe to make him more dexterous - these solutions attempt to change Joe and make him more productive.
   Ex: "Give him practice in manipulative skills."

**Keeping Joe in the company, but on a different work team or job.

3. Transfer Joe to one of the other teams - these solutions involve placing Joe with one of the other teams, all do the same kind of work as does Joe's present team.
   Ex: "Transfer Joe to A, B, or D."

4. Transfer Joe to a non-production job in which he will be under the supervision of a foreman - these
solutions are important to note, because they constitute a violation of the stated conditions. The problem specifically states that this cannot be done, other foremen on non-production jobs are unwilling to accept Joe as a transfer. The judges were cautioned to watch for this. Ex: "Make one of the non-production foremen take him anyway, they have to find a place for him."

5. Transfer Joe to a non-production job in which he will not be under the supervision of a foreman – these solutions attempt to find some other place where Joe can be employed. Ex: "Make Joe a foreman, advisor, janitor, etc."

***Removing Joe from the company

6. Fire Joe.
7. Retire Joe.

B. Criteria Concerning Centering

1. Joe centered – a solution oriented in terms of Joe's needs.
   Ex: "The main point is that the company has got to take care of Joe, he's an old man."

2. Company centered – a solution oriented in terms of the company's needs.
   Ex: "Joe is the bottleneck in production, he's got to go if he can't produce."

3. "Both" centered – a solution oriented in terms of the needs of Joe and the needs of the company.
   Ex: "Joe must have plenty of know-how, make him a foreman, the advancement would be a recognition for his faithful service and production would probably increase."
C. Criteria Concerning Inadequate Solutions

Because these criteria are identical to those used in "The Tennis Problem" they will not be repeated here. They appear on page 86 of this thesis.

D. Criteria Concerning Means

(In this problem the company must take some kind of action.)

1. Verbal appeal
   Ex: "Give Joe a pep talk."

2. Threat or pressure
   Ex: "Give Joe a warning to produce or else!"

3. Withdrawal - refusing to deal with the problem.
   Ex: "Let's wait awhile and see what will happen."

4. Break off the relationship
   Ex: "Fire Joe, fire the foreman, reshuffle the group."

5. Use Joe productively - this is considered the most adequate solution. Joe is considered an asset to be used somewhere.
   Ex: "Joe must have a good background of knowledge about the company, this is too valuable to be lost, so we must find a spot where he can use this knowledge, then everyone will benefit."

Tabulation of the Judges' Ratings

The judges' ratings on "The Assembly-Line Problem" were tabulated in the same manner as those for "The Tennis Problem." Again, the tabulation was made in two ways to guard against differential results attributable to the
method used. The first method was to record an entry if either of the judges had scored it. The second method was to make an entry only if both judges had scored it. Several sample calculations made it clear that the differences in scoring yielded negligible differences in results. Therefore, in further computations, scoring was done by the first method - an entry was recorded if either judge had scored it. The data has been tabulated and appear in Tables IX, X, XI, and XII on the next two pages.

**Hypotheses and Results**

In the following section we shall examine some specific hypotheses concerning expected differences in behavior between the groups. Since "The Assembly-Line Problem" offers a wide range of possible solutions, we shall state our prediction and the results pertaining to each solution before proceeding to the next. At the conclusion, we shall attempt to integrate and summarize the results.

**Hypotheses Concerning "Concrete Action Proposals"**

There were three general classifications under this criterion. The first was "Keeping Joe in the company, on the same work team," which involved two categories.

1. "Give the less capable less to do." This solution is desirable because it minimizes the shifting of personnel, since the changes occur only within this one work group. Next to the "elegant" solution, Maier considers it the most feasible (67, p. 318).
TABLE IX.
FREQUENCIES
OF
"CONCRETE ACTION PROPOSALS"
TO
"THE ASSEMBLY-LINE PROBLEM"
IN THE
EXPERIMENTAL AND CONTROL GROUPS

<table>
<thead>
<tr>
<th>Same</th>
<th>Train</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>Fire</th>
<th>Retire</th>
<th>Misc</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>6</td>
<td>0</td>
<td>7</td>
<td>10</td>
<td>26</td>
<td>0</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>8</td>
<td>1</td>
<td>11</td>
<td>8</td>
<td>26</td>
<td>2</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Tot.</td>
<td>13</td>
<td>1</td>
<td>18</td>
<td>18</td>
<td>52</td>
<td>2</td>
<td>17</td>
<td>2</td>
</tr>
</tbody>
</table>

"The N is not equal as alternate proposals were permitted.

TABLE X.
FREQUENCIES
OF FORMS OF
"CENTERING"
"THE ASSEMBLY-LINE PROBLEM"
IN THE
EXPERIMENTAL AND CONTROL GROUPS

<table>
<thead>
<tr>
<th></th>
<th>&quot;Both&quot; Centered</th>
<th>&quot;Other&quot; Centered</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>31</td>
<td>5</td>
<td>36</td>
</tr>
<tr>
<td>E</td>
<td>28</td>
<td>8</td>
<td>36</td>
</tr>
<tr>
<td>Tot.</td>
<td>59</td>
<td>13</td>
<td>72</td>
</tr>
</tbody>
</table>
TABLE XI.
FREQUENCY OF ADEQUATE AND INADEQUATE SOLUTIONS TO "THE ASSEMBLY-LINE PROBLEM" IN THE EXPERIMENTAL AND CONTROL GROUPS

<table>
<thead>
<tr>
<th></th>
<th>S.'s free from inadequate solutions</th>
<th>S.'s with one or more inadequate solutions</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>7</td>
<td>29</td>
<td>36</td>
</tr>
<tr>
<td>E</td>
<td>7</td>
<td>29</td>
<td>36</td>
</tr>
<tr>
<td>Tot.</td>
<td>14</td>
<td>58</td>
<td>72</td>
</tr>
</tbody>
</table>

TABLE XII.
FREQUENCY OF "USE JOE PRODUCTIVELY" MEANS FOR "THE ASSEMBLY-LINE PROBLEM" IN THE EXPERIMENTAL AND CONTROL GROUPS

<table>
<thead>
<tr>
<th></th>
<th>S.'s who solved by This Means</th>
<th>S.'s who did not solve by This Means</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>19</td>
<td>17</td>
<td>36</td>
</tr>
<tr>
<td>E</td>
<td>24</td>
<td>12</td>
<td>36</td>
</tr>
<tr>
<td>Tot.</td>
<td>43</td>
<td>29</td>
<td>72</td>
</tr>
</tbody>
</table>
Hypothesis and Results

Hypothesis: The experimental group will present this solution more frequently than will the control group.

Table IX on page 108 reveals that the experimental group does excel the control group in this respect. However, application of the Chi Square Test revealed that the trend was not statistically significant (P > .05).

2. "Training Joe to make him more dexterous." This solution does not appear adequate, because the problem states that Joe is lacking in this skill and training would seem of dubious value. However, since training may be suggested in conjunction with other solutions, it is difficult to properly assess this solution. Therefore, no prediction was made. The results show that only one individual suggested this possibility.

The second general classification was, "Keeping Joe in the company, but on a different work team or job," and in-

---

<table>
<thead>
<tr>
<th></th>
<th>5</th>
<th>31</th>
<th>36</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>13</td>
<td>59</td>
<td>72</td>
</tr>
</tbody>
</table>

*The Chi Square Test for this and the following tests under "Concrete Action Proposals" were set up in the following way to circumvent the problem of independent entries: S. does not suggest S. suggests
volved three categories.

3. "Transfer Joe to one of the other teams." This answer is difficult to evaluate. It may be considered favorable, because it involves a transfer to the same kind of work, and involves a minimum disturbance of personnel. Viewed in this frame of reference, it may be construed as a means of containing the problem — of keeping it in the same department. However, from another point of view, such a solution by-passes the basic problem. Moving the slow man to another group confronts them with the problem. Because of this conflicting interpretation, no prediction was made.

The results indicate that the experimental group gave more solutions of this type, but the result was not statistically significant (P > .05), when tested by the Chi Square Test.

4. "Transfer Joe to a non-production job in which he will be under the supervision of a foreman." We have previously stressed that this solution represents a clear violation of the stated conditions, for the problem reads: "Foremen on non-production jobs are not willing to accept Joe as a transfer." Our hypothesis was equally clear.

**Hypothesis and Results**

**Hypothesis:** The experimental group will present this solution less frequently than the control group.

The results bear out our prediction, but inspection revealed that the difference was not statistically significant (P > .05). The superiority of the experimental group is slight.
5. "Transfer Joe to a non-production job in which he will not be under the supervision of a foreman." This type of solution was not crucial to the present study. While it by-passes the basic problem, it attempts to use Joe productively within the limitations of the problem. Since this was not a crucial aspect of the study, no prediction was made. The results show that both groups scored exactly the same and that a large number of S.'s chose this solution.

The third general classification was, "Removing Joe from the company," and involved two categories.

6. "Fire Joe." This constitutes a drastic proposal for reasons of morale, company prestige, etc. and probably creates more problems than it settles.

Hypothesis and Results

Hypothesis: The experimental group will not make this recommendation as frequently as will the control group.

The results indicate only two cases - both experimental - among all the S.'s who chose this solution. While the number is small, the hypothesis does not seem tenable.

7. "Retire Joe." This solution is also a drastic proposal, though not as devastating as (6). It cannot be considered an adequate solution, since other possibilities might prove more fruitful.

Hypothesis and Results

Hypothesis: The experimental group will not present this solution as frequently as will the control group.
The results refute this hypothesis, for twice as many experimental S.'s chose this solution as did control S.'s. The Chi Square Test indicated that the difference was not statistically significant (P > .05).

**Hypothesis Concerning "Inadequate Solutions"**

As has been previously outlined, this criterion was used in an attempt to measure shortcomings in the proposed solutions.

**Hypothesis and Results**

**Hypothesis:** The experimental S.'s, because of their "case-discussion" method training will show fewer inadequacies than the control S.'s.

The data clearly refuted this hypothesis, since the score for both groups was identical.

**Hypothesis Concerning "Centering"**

The category which represented adequate centering was number three, "both-centered."

**Hypothesis and Results**

**Hypothesis:** The experimental group, because of their "case-discussion" method training will excel the control group in the use of "both-centered" solutions.

A Chi Square Test demonstrated that the difference between groups was not statistically significant (P > .05) though the trend was in the predicted direction.
Hypothesis Concerning "Means"

Of the five categories under the "means" criterion, only "Use Joe Productively" represents an adequate means.

Hypothesis and Results

Hypothesis: The experimental group, because of their "case-discussion" method training, will employ this means of solution to a greater extent than the control group.

While the results yielded a difference in the direction indicated, the superiority of the experimental group was small. The Chi Square Test indicated that the difference was not statistically significant (P > .05).

Conclusion

The results in general have not supported our hypotheses with statistically significant findings. However, four trends appeared, which favor the experimental group, namely (1) Attempts to keep Joe on the same work team; (2) Fewer violations of conditions stated in the problem; (3) More adequate centering; and (4) More solutions which attempt to use Joe productively. The control group was superior to the experimental group in one area, namely, they showed less tendency to solve the problem by retirement. Concerning the number of solutions which involved inadequacies, both groups were identical.
While the majority of these trends were in the predicted direction, the findings were not statistically significant. It may be that "The Assembly-Line Problem" has proved too complicated to be dealt with adequately in a fifteen-minute leaderless group discussion. Alternative explanations could be offered, but it would seem premature to do so here. Instead we shall proceed to the judgment of the transcripts which were related to the judgments of the written solutions and then attempt to integrate and interpret the various results.
CHAPTER X
THE GROUP DISCUSSION TRANSCRIPTS

Introduction

It will be remembered that each S. was given the problem in typewritten form to be read individually and when all had read it, E. said, "You'll have fifteen minutes on this problem." "You may begin now." Ordinarily, E. took no part in the ensuing leaderless group discussion. E. entered the discussion only to correct misunderstandings, as this might give the discussion a tangential direction which would render the results atypical. For example, on several occasions someone misread the information given in the problem with regard to the group in which Joe worked. When this happened, E. waited a bit to allow another member to point this out. If this did not occur, E. interrupted and said, "Excuse me, but I think there's a misunderstanding here. Someone said Joe was in group D, but the problem stated that Joe is in group C., isn't that correct?"

When time ran out, E. terminated the discussion. If a group had apparently solved the problem to their own satisfaction within the time limit and wanted to end the discussion, E. said, "It's up to the group. If you feel you've solved the problem and don't want to continue, you don't have to." Four groups solved the problem within the time limit.
While the problem was being discussed, E. took sequential notes of each statement, including the number of the speaker as well as an identifying phrase. This notation technique proved satisfactory for only a few words were lost in the course of transcription. A finished transcript involved about eight hours of listening and preparation and averaged about ten typewritten pages.\*  

Transcripts as Scientific Behavioral Data  

It is true that a transcript is a reproduction of the actual behavior and subject to limitations inherent in any facsimile. Transcripts have an important shortcoming, in that the richness of gesture, expression, and other non-verbal behavior are lacking. However, verbal behavior may be reproduced more accurately and completely by transcription than might have been assimilated by anyone listening to the discussion while it was in progress.

A transcript has merit if it has been collected under rigorous experimental conditions for it remains a permanent scientific record of the actual behavior, can be consulted repeatedly, and lends itself to virtually unlimited analysis. A transcript has further value in that it would seem to enhance objectivity, for biased judgment due to factors such as voice and personality have been largely eliminated. Thus

\*A complete transcript appears in the Appendix, pp. 192-203.
transcripts represent scientific data that remain a permanent objective record amenable to examination at will, the analysis of which is limited only by the ingenuity of the investigator.

Probably the most satisfactory and complete method of collecting behavioral data would be to have judges observing behavior during the discussion, and rating it at the conclusion of the discussion, as well as having a typewritten transcript available for reference. Unfortunately, the former method was not feasible, so the latter method proved expedient. By means of the procedures previously described, the twenty-four group discussions were reproduced in typescripts.

The Transcript Judgment Design

From the viewpoint of psychological experimentation, the experimental design must always take into account "hidden" variables which may bias the judging procedure. One of the most cogent findings of Fechner's psychophysical research almost a century ago was the discovery that the sequence of stimulus presentation influenced judgment. For example, it has been clearly established that under many conditions the judgment of a second stimulus is influenced by the judgment of the initial stimulus. Fechner discovered this phenomenon present in the judging of weights and lights. Other experimenters have studied these effects, which have become known as "time" and "space" errors. In the study of attitudes it has also been demonstrated that an initial presentation
serves to establish a frame of reference which influences further judgment.

One of the most effective methods of mitigating the effects of these phenomena is by means of counterbalancing. This procedure involves the presentation of two different stimuli in a pattern which balances influential effects. The prototype of such a design is the "ABBA" sequence, which is a frequently used method of counterbalancing. In the present study there were three important variables which E. wished to take into account lest they bias the judgments. They were: (1) the type of group, whether experimental or control; (2) the time of discussion, whether "before" or "after;" and (3) the problem discussed, whether "The Tennis Problem" or "The Assembly-Line Problem." The latter two variables were related in that all groups discussed "The Tennis Problem" at the "before" session and "The Assembly-Line Problem" at the "after" session, but this was unknown to the judges.

As a preliminary measure, the judges were kept unaware of these three aspects. This was not deemed sufficient, for experimental rigor dictated the incorporation of counterbalancing into the transcript judgment design. Therefore, each judge was requested to judge his set of four transcripts in a prescribed order. The order had been set up by an intricate design which encompassed the forty-eight transcript judgments.*

*This design appears on the next page, p. 120.
### TABLE XIII. THE TRANSCRIPT JUDGING DESIGN

<table>
<thead>
<tr>
<th>Judge</th>
<th>Group</th>
<th>Sequence</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>I</td>
<td>Experimental-Control</td>
<td>BAAB</td>
</tr>
<tr>
<td>II</td>
<td>II</td>
<td>Experimental-Control</td>
<td>ABBA</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
<td>Experimental-Control</td>
<td>ABAB</td>
</tr>
<tr>
<td>IV</td>
<td>IV</td>
<td>Experimental-Control</td>
<td>BABA</td>
</tr>
<tr>
<td>V</td>
<td>V</td>
<td>Experimental-Control</td>
<td>BAAB</td>
</tr>
<tr>
<td>VI</td>
<td>VI</td>
<td>Experimental-Control</td>
<td>ABAB</td>
</tr>
<tr>
<td>VII</td>
<td>I</td>
<td>Control-Experimental</td>
<td>ABBA</td>
</tr>
<tr>
<td>VIII</td>
<td>II</td>
<td>Control-Experimental</td>
<td>ABAB</td>
</tr>
<tr>
<td>IX</td>
<td>III</td>
<td>Control-Experimental</td>
<td>BABA</td>
</tr>
<tr>
<td>X</td>
<td>IV</td>
<td>Control-Experimental</td>
<td>BAAB</td>
</tr>
<tr>
<td>XI</td>
<td>V</td>
<td>Control-Experimental</td>
<td>ABAB</td>
</tr>
<tr>
<td>XII</td>
<td>VI</td>
<td>Control-Experimental</td>
<td>BAAB</td>
</tr>
</tbody>
</table>

Judging Sequence:

1. **BE1**
2. **AE1**
3. **AC1**
4. **BC1**
5. **AE2**
6. **BE2**
7. **AC2**
8. **BC2**
9. **AE3**
10. **BE3**
11. **AC3**
12. **BC3**
13. **AE4**
14. **BE4**
15. **AC4**
16. **BC4**
17. **AE5**
18. **BE5**
19. **AC5**
20. **BC5**
21. **AE6**
22. **BE6**
23. **AC6**
24. **BC6**

"The judging sequence is to be interpreted as follow: The sequence for Judge I was:

1st - "before" session, Experimental Group #1
2nd - "after" session, Experimental Group #1
3rd - "after" session, Control Group #1
4th - "before" session, Control Group #1."
The transcript judgment design counterbalanced the order of presentation of experimental and control transcripts, and at the same time counterbalanced the order of presentation of transcripts from "before" and "after" sessions. As a perusal of the design indicates, the second judge followed a sequence which was also counterbalanced according to these two factors, but was counterbalanced in a different manner from that of the first judge. As a result of this complicated methodology, each set of four transcripts was read in a different, but commensurately counterbalanced sequence by each of the two judges, thus ensuring some correction for factors which it was believed might bias judgment.

As an additional precaution, each judge was told his set of four transcripts might include four, three, or two of one kind of group (experimental or control). Actually, each judge received the four transcripts from one matched pair, namely: Experimental Group #1, "before" and "after" sessions, and Control Group #1, "before" and "after" sessions. It was believed that this combination of transcripts would result in the complete set of one matched group being judged in the same frame of reference.

Each transcript was assigned a code number, the key to the code being known only to E. E. was very fortunate concerning the important problem of gaining the assistance of competent judges: Twelve members of the University of Kansas Staff voluntarily gave generously of their time so that this task could be accomplished. Six of the judges
were psychologists, six were teachers experienced in the "case-discussion" method of teaching."

The judges did not know what kind of group they were judging, neither did they know which problem was discussed at each session. The judges were given no information pertaining to the experiment that might influence their judgments.

Each judge was informed in advance concerning the length and requirements of the task. E. presented each judge with an envelope containing the four transcripts, a rating scale for each one, a copy of the criteria, and a letter of instructions." The task was organized so that each transcript was judged by two different judges who worked independently. The distribution of the transcripts was arranged so that each transcript was judged by a psychologist and by a teacher experienced in the "case-discussion" method of teaching. All these individuals were members of the University of Kansas staff. E. took no part in the judging procedure other than providing the necessary orientation.

---

"The writer is indebted to the following individuals who served as judges: Professors Louise Barker, Lawrence Bee, Carroll Clark, Hilden Gibson, Fritz Heider, Marston McCluggage, Lee Meyerson, Edward G. Nelson, William Thompson, Charles Warriner, Edward Wike, M. Erik Wright.

"This material appears in the Appendix, pp. 188-191."
While the judging process might have been managed in other ways, there were limitations imposed by the nature of the task. In the first place, the task was exacting. It required concentration over a number of hours. Moreover, it was imperative to gain the services of competent individuals, and it would have been an imposition to request their services over too long a period of time. Another problem was that a judge would need a minimum number of transcripts in order to form a frame of reference within which each transcript could be judged. These problems were resolved by assigning four transcripts to each judge. Since twelve judges cooperated, this plan permitted each transcript to be judged twice, so that a total of forty-eight judgments resulted.

Reliability of the Judges

Since each transcript was judged by a psychologist and by a "case-discussion" method teacher, it is pertinent to inquire concerning the agreement between the judges. Three variables were analyzed. The first was the ability of the judges to judge the type of group from the transcript - whether it was an experimental or a control group. The second, agreement pertaining to the emergence of a group solution and its adequacy. The last comparison was between both judges' ratings for each of the remaining seven variables, all of which were judged by rating on a continuum.

1. Agreements concerning type of group

This measure is different from the other two kinds of
reliability, since the judges had no knowledge of the type of group, and such judgment was based on the general impression gained by reading the transcript. It was considered important to determine whether there were differences in discriminatory ability between the two classes of judges, as this held implications for other aspects of the experiment. For example, if all the judges were highly accurate, this would signify that the groups had sharply distinguishing characteristics. If one group of judges proved superior, this might be attributed to the discipline represented by the judges. If neither group proved to be accurate, this would indicate that the groups being judged did not present sharply distinguishing characteristics.

No prediction was made, because if differences were large, discrimination would be enhanced. If differences were small, judgments would be more subject to error. The results of this tabulation appear in Table XIV, (p. 125). These results indicate small differences at the "before" session. The identical results of judgments pertaining to the "after" session demonstrates that there were no differences between both types of judges in their ability to judge the type of group. When the total judgments are taken into consideration, regardless of type of judge, all judges are correct slightly more than half of the time. This may be interpreted as indicating that the different types of groups did not have outstanding distinguishing characteristics, and that neither type of judge proved clearly superior.
### TABLE XIV.

**ABILITY TO JUDGE TYPE OF GROUP FROM DISCUSSION TRANSCRIPT**

<table>
<thead>
<tr>
<th></th>
<th>Correct</th>
<th>Incorrect</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Before&quot; Session</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychologists</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>&quot;Case-Discussion&quot;</td>
<td>4</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Teachers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total N &quot;Before&quot; Session</td>
<td>10</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td>&quot;After&quot; Session</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychologists</td>
<td>7</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>&quot;Case-Discussion&quot;</td>
<td>7</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Teachers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total N &quot;After&quot; Session</td>
<td>14</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td>Total N &quot;Both&quot; Sessions</td>
<td>24</td>
<td>24</td>
<td>48</td>
</tr>
</tbody>
</table>
2. **Agreements Concerning Solution and Adequacy**

The judges were asked to state whether there had been a group solution to the problem. The results have been tabulated and appear in Table XV (p. 127). Reading in lowest row, "Total N - both sessions" it is clear that agreements and disagreements are about equal. The ratio is thirteen agreements to eleven disagreement. Because there were no group solutions to "The Assembly-Line Problem" which was used at the "after" session, the criterion concerning adequacy needs no further comment.

As previously mentioned, we may conclude that either this material is difficult to judge or that the judges viewed the material in different frames of reference. It is possible that disagreement arose due to differing interpretations of the phrase "a group solution." Some judges might set a rigid criterion, such as unanimity, whereas other judges might permit a minority dissension. It is possible that agreement might have been higher had this criterion been made more explicit. Other relationships disclosed by Table XV will be pointed out later when the hypotheses and results are discussed.

3. **Agreement Concerning Continua Ratings**

Since each of the seven remaining variables were rated on a continuum, it is pertinent to inquire into the agreement between the two judges. Agreement was analyzed by computing a Spearman Rank Difference Correlation Coefficient between the ranks assigned by the two judges to the
TABLE XV.

RELIABILITY OF JUDGES
CONCERNING PRESENCE OR ABSENCE OF
GROUP SOLUTION

<table>
<thead>
<tr>
<th></th>
<th>Agreement</th>
<th>Disagreement</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Before&quot; Session</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Experimental</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Total N &quot;Before&quot; Session</td>
<td>9</td>
<td>3</td>
<td>12</td>
</tr>
</tbody>
</table>

| "After" Session  |           |              |    |
| Control          | 2         | 4            | 6  |
| Experimental     | 2         | 4            | 6  |
| Total N "After" Session | 4 | 8 | 12 |

Total N "Both" Sessions | 13 | 11 | 24

*Agreement was defined in terms of concurrence of both judges.
seven variables. Twenty-four correlation coefficients were calculated, transformed into Z scores, and averaged to obtain an average Z score for all the transcript judgments. This Z score was transformed into an average correlation coefficient by which agreement was measured. The result was an average Z score of .38, the average Spearman Rank Difference Correlation Coefficient was .36. This was an indirect check on reliability.

The results indicate that agreement was not statistically significant though the trend was toward agreement. Several explanations are plausible. Even though each judge was oriented to the judgment procedure in a standard manner, it may be that the psychologists and the "case-discussion" method experienced teachers had different frames of reference in terms of which they rated the transcripts. Though the correlation was low, it was positive, and permits us to analyze the data further.

In conclusion, the reliability of the judges concerning the emergence of group solutions was low, as was their reliability on the continua ratings. As was suggested in the case of the written solutions, either this type of material is difficult to judge, or it was judged in different frames of reference.

The Tabulation of the Data

The continuum along which judgments were recorded measured exactly fifteen millimetres, so the judges' mark could be read off on a scale from zero to fifteen. Since
each transcript was judged independently by two judges, these two scores were averaged to obtain a mean score for each variable on which the transcript had been judged. The outcome will be discussed in the section concerning the hypotheses and results.

The Criteria

Two of the criteria have been briefly discussed in the preceding section. They will be discussed further in this section, as well as the remainder of the criteria, which were rated on a continuum. All the criteria were established to measure some important aspects of group discussion. The underlying assumption was that each criterion was important to group discussion. These criteria represent recognized ways of evaluating group discussion and have been used by many investigators for this purpose. The explanations of the criteria are the same as those given to the judges.

1. **Insight** - An understanding of the basic issues involved in a problem, such as the feelings, needs, and motivations of the individuals involved, the economic demands of the situation, and the implications of these factors for the total situation seems to be a prerequisite for adequate problem-solving.

2. **Group solution and adequacy** - It was considered important to ask the judge whether or not there had been a group solution. The judge was also asked to state the solution and rate its adequacy. An adequate solution was
defined as one which was not couched in terms of a one-sided viewpoint, such as a solution in terms of the needs of one individual, or any limited aspect, but one that considered the total situation.

3. **continuity** - It was believed that continuity was a desirable characteristic of group discussion, since order or continuity suggests that the group is proceeding with some direction, that a goal has been cognized, whereas, a haphazard, disjointed, or aimless series of statements indicates that direction is lacking.

4. **atmosphere** - It was believed that a warm, friendly and permissive atmosphere, one in which individuals tend to respond spontaneously, enhances group discussion, whereas, a neutral, strained or negative atmosphere does not.

5. **task-centered** - It was believed that a group which directed its energies toward the task with which the group was confronted, whose immediate purpose was the facilitation of the problem solution, was functioning more adequately than a group which showed little interest in the task.

6. **group centered** - It was believed that a group which directed its energies toward the functioning of the group as a group, whose immediate purpose was the solution of within-group problems, such as the maintenance, strengthening, and regulation of the group, including also the facilitation of communication, summarizing, reconciling differences, etc. was functioning more adequately than a group that was not. It should be mentioned that this variable
need not conflict with task-centeredness. An adequately functioning group probably would not neglect either aspect, since they are complementary.

7. **individual centered** - It was believed that in a group in which energy was directed primarily toward the satisfaction of the participants' individual needs, whose immediate purpose was the reaching of individual goals which were neither task nor group relevant, was not functioning adequately. In the present context "individual" goals connotes goals that are individual in the sense that the satisfaction aimed at by the participant cannot be shared by the other members of the group. This variable is in sharp conflict with the two previous variables. Task-centered and group-centered behavior are relevant to the group interaction process and group problem-solving process, whereas individual-centered behavior is not. It should be emphasized that this is the only variable in which a low score represents more adequate functioning than a high score.

8. **successful communication** - Communication was defined as successful when, in the eyes of the judges, individuals showed by their immediate or subsequent responses that they had interpreted the contributions of others as they were intended.

9. **judgments concerning the type of group** - While this does not constitute a variable it was believed important to determine whether the judges were able to correctly designate
the type of group. This aspect of the study has already been discussed under the section concerning the judges' agreement on pages 123 and 124.

HYPOTHESES AND RESULTS

Of the nine criteria which were rated by the judges, seven involved rating on a continuum and two involved written reports. Because these two must be analyzed in a different manner, they will be dealt with first.

Solutions and Adequacy

Concerning criterion two, the judge was asked to state whether a group solution had emerged, and if this occurred, to state it briefly and to rate its adequacy.

Hypothesis: The experimental group will present more group solutions and more adequate ones.

The results appear in Table XVI (p. 133). The data indicate that there were small differences at the "before" session. This is favorable to the experiment since it lends empirical evidence to support the assumption that both groups were about equal at the "before" session.

However, concerning the "after" session, the results do not support the hypothesis, for neither group achieved a solution. These results may be interpreted as indicating that "The Assembly-Line Problem" proved too difficult for the S.'s, or that the S.'s had not progressed in problem-solving skills to the point where they were able to achieve a group solution.
TABLE XVI.

FREQUENCY OF GROUP SOLUTIONS*

<table>
<thead>
<tr>
<th>Session Type</th>
<th>Control</th>
<th>Experimental</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Before&quot;</td>
<td>3</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Total N &quot;Before&quot; Session</td>
<td>5</td>
<td>7</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Session Type</th>
<th>Control</th>
<th>Experimental</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;After&quot;</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Total N &quot;After&quot; Session</td>
<td>0</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Total N &quot;Both&quot; Sessions</td>
<td>5</td>
<td>19</td>
<td>24</td>
</tr>
</tbody>
</table>

*A group was scored as having a solution only if both judges agreed.*
It will be remembered that the judges were asked to briefly state the group solution. It is interesting to note that in all cases where both judges agreed that a group solution had emerged, the two judges wrote identical descriptive statements. This finding gives some positive indication of reliability of the judges. Since the ratings of adequacy of solution concerned the five solutions which occurred at the "before" session, these will not be evaluated. The number of solutions is small and our interest was not in analyzing solutions which occurred at the "before" session.

This finding is related to the reliability of the judges concerning group solution. Inspection of Table XV (p. 127) reveals a consistent pattern which sheds light upon other aspects of the experiment. A comparison of experimental and control groups according to the agreements and disagreements of the judges, for either session, reveals that differences between the groups in performance is small. For example, at the "before" session, five pairs of judges of control group transcripts agreed that these groups had achieved a group solution, and that four pairs of judges of experimental group transcripts agreed that these groups had done likewise. Inspection of the "after" session reveals that these same ratings total two for each group.

However, when we compare the total "before" session scores with the total "after" session scores, a clear reversal is apparent. The row, "Total N 'before' sessions" indicates that agreements are three times as frequent as
are disagreements. However, the row, "Total N 'after' sessions" indicates that agreements occur half as frequently as disagreements. This reversal demands interpretation.

Since the differences between groups were small, it would seem that the groups were alike in that at the "before" session, there was more agreement, and at the "after" session, there was less agreement. Thus the differences apparently may be attributed to the session. If we accept the proposition that more agreements than disagreements signifies that the data lent itself more readily to judgment, we may interpret this reversal as indicating that the data from the "before" session were easier to judge.

Relative ease of judgment may indicate something about the nature of the problem, in that "The Tennis Problem" was less complex. It will be remembered that four groups solved this problem at the "before" session and this must have been obvious to the judge who read these transcripts. This would account for a higher degree of agreement at the "before" session. No groups solved the "after" session problem.

Relative ease of judgment may also reflect a less complex interaction process. It may be that it is easier to agree when a group either solves a problem or doesn't, but that it is more difficult to agree when the data are ambiguous. It seems that disagreement of the judges due to
the complexity of the interaction process or disagreement of judges due to the complexity of the problem are inter-related. A more difficult problem often results in a more complex interaction.

Type of Group

The second criterion which required a written judgment rather than a rating on a continuum was criterion nine. The judge was asked to state his opinion regarding whether the transcript was from a "case-discussion" method group or a non-case-discussion method group. The results which appear on page 125, Table XIV, indicate that concerning the post-test, neither group of judges excelled the other, since the scores for both groups were identical. Taking all judges as a whole, their correct judgments totalled thirteen, incorrect judgments were eleven. These differences are not statistically significant. These results seem to reflect upon the impressions which the judges formed concerning differences between the groups. It may be inferred from the post-test data that evidently the differences were not great or accuracy of judgment would have been higher. This finding is clear, and must be related to other aspects of the study in order to verify this inference.

In conclusion, the results of the data elicited by these two criteria which involved a written judgment, indicated that differences between groups were small. In the next section, the data from the remaining seven criteria, which were rated on a continuum, will be analyzed and then all of
criteria on which the transcripts were judged will be summarized and integrated.

Continua ratings

Concerning the remaining criteria which were judged by ratings on a continuum, a blanket hypothesis was made as follows:

Hypothesis and Results

Hypothesis: The experimental group, because of their "case-discussion" method of training, will excel the control group in the following group interaction-process and problem-solving skills: The experimental group will show more insight into the nature of the problem and their solutions will prove more adequate. The experimental group will demonstrate a higher degree of continuity in the development of the discussion, and be more successful in communication. Their interactions will be more task-centered and group-centered, but less individual-centered, and they will function in a more permissive and friendly atmosphere.

The data appear in Table XVII (p. 139). The data were analyzed by a "t" test and the computation procedure was
as follows: The gain score for experimental group #1 was calculated by comparing their rating on variable #1 at the "before" session with their rating at the "after" session. The same procedure was followed for control group #1. Then the gain scores for both groups were compared, resulting in a difference score between this matched pair of groups. In this manner, difference scores between each matched pair of groups was calculated for each variable. The results appear in Table XVII (p. 139), row five, "t" test scores.

The results are consistent and in the predicted direction, but only the difference on variable #4, "atmosphere" was statistically significant (P.<.03). The results constitute the most striking finding of the present study. Inspection of Table XVII (p. 139) shows that the experimental groups have gained more than the control groups. There is a range of gains, and not all are large. However, without exception, the experimental group excels the control group on each of the seven variables.

While the results are not statistically significant they will bear some interpretation. Inspection of the largest differences indicates that the experimental group functions in a more friendly, permissive atmosphere, and that they are definitely not individually centered. Concerning the remaining differences which are not so large, we can state that the experimental group shows a tendency to maintain a continuity of discussion, shows some evidence of successful communication, and is somewhat task-centered. These findings
TABLE XVII.

TABULATION OF THE TRANSCRIPT RATINGS

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insight</td>
<td>Continuity</td>
<td>Atmosphere</td>
<td>Task Centered</td>
<td>Individual Centered</td>
<td>Communication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.</td>
<td>E.</td>
<td>C.</td>
<td>E.</td>
<td>C.</td>
<td>E.</td>
<td>C.</td>
<td>E.</td>
</tr>
<tr>
<td>10.18</td>
<td>3.18</td>
<td>9.32</td>
<td>6.63</td>
<td>10.90</td>
<td>8.69</td>
<td>11.23</td>
<td>9.81</td>
</tr>
<tr>
<td>2.72</td>
<td>7.32</td>
<td>5.83</td>
<td>5.75</td>
<td>6.90</td>
<td>9.99</td>
<td>9.22</td>
<td>6.63</td>
</tr>
<tr>
<td>7.69</td>
<td>6.63</td>
<td>8.04</td>
<td>6.94</td>
<td>9.37</td>
<td>11.25</td>
<td>10.58</td>
<td>6.35</td>
</tr>
<tr>
<td>6.72</td>
<td>8.33</td>
<td>8.74</td>
<td>8.74</td>
<td>8.74</td>
<td>8.74</td>
<td>8.74</td>
<td>8.74</td>
</tr>
<tr>
<td>-2.49</td>
<td>-1.37</td>
<td>-1.23</td>
<td>-1.12</td>
<td>.68</td>
<td>.03</td>
<td>.77</td>
<td>.97</td>
</tr>
<tr>
<td>-1.17</td>
<td>-2.01</td>
<td>-1.66</td>
<td>-1.66</td>
<td>-1.66</td>
<td>-1.66</td>
<td>-1.66</td>
<td>-1.66</td>
</tr>
<tr>
<td>.92</td>
<td>1.39</td>
<td>2.08</td>
<td>.74</td>
<td>1.86</td>
<td>-1.84</td>
<td>1.18</td>
<td>1.18</td>
</tr>
</tbody>
</table>

5. "t" test scores:

| .70 | .63 | 2.53 | .35 | .92 | 1.16 | .47 |
| (P.<.26) | (P.<.28) | (P.<.03) | (P.<.37) | (P.<.20) | (P.<.15) | (P.<.33) |
| not sig- | not sig- | sig- | not sig- | not sig- | not sig- | not sig- |

*It will be noted that in every case, the E. group has excelled the C. group.

**All of the "t" tests are one-tail tests. A "t" of 2.02 is required for significance at the .05 level.
all concern group interaction process variables. The superiority of the experimental group pertaining to insight, which is the only variable more directly related to productivity, was small.

However, because the performance of the experimental group excelled the control group on all these seven variables, the possibility must be raised that these variables are interrelated and that therefore the same thing was being measured in seven different ways. It may be that these seven variables represent some underlying global trait, such as social or verbal facility, but this is difficult to determine. It does not seem likely that performance on one variable permits prediction of performance on other variables. For example, a group could devote so much time to group-centered interaction, that task-centered interaction would suffer, resulting in a low correlation. A group could show a high degree of individual-centered behavior in a highly permissive atmosphere, whereas under conditions believed optimum to group problem-solving, atmosphere would be highly permissive, but individual-centered behavior would be low. Insight need not necessarily correlate with atmosphere, task-centeredness, group-centeredness, and successful communication, for it is possible to have a high degree of behavior of the latter types, and little insight into the problem, as when a group discusses a problem on a superficial level.
The General Pattern of Both Sessions

Inspection of Table XVII on page 139 reveals the curious fact that at the "before" session, the control group excelled the experimental group on each of the seven variables. Several explanations are possible. Prior to the "before" session, the experimental group had experienced several meetings of their "case-discussion" method classes. It may be that they perceived the experimental situation as related to their class work in some way, and that they were less relaxed, and were trying too hard. Another possibility is that these few class meetings were so utterly different from the traditional classroom routine, that the experimental group S.'s were frustrated and unable to cope with the situation. In the writer's experience, this is a frequent occurrence at the beginning of the term in a "case-discussion" method class. Students seem uneasy and reluctant to participate and interchange ideas, because the role of the teacher is unstructured.

Another possibility is that students who seek out "case-discussion" method classes represent a special population of students who are lacking in discussion skills or who enjoyed discussion and that this may affect group behavior. As was pointed out previously, this possibility cannot be fully tested because a minority of the S.'s were from "case-discussion" method classes which were required courses.

Conclusion

There were no significant differences between judges
who were psychologists and judges who were "case-discussion" method experienced teachers concerning ability to judge type of group from transcripts. The judges as a whole were correct slightly more than half the time. Judges' agreement concerning whether a group solution had emerged was also low. An average Spearman-Rank Difference Correlation Coefficient was obtained for the judges' ratings of the seven variables which were rated on a continuum. This was an indirect measure of agreement, and yielded a value of .36, which was low, and probably represents a minimum of reliability.

It was predicted that the experimental group would excel the control group in the matter of group solutions and adequacy of solution. However, since no group solutions occurred at the "after" session, there was no difference between groups concerning problem-solving. While the experimental group excelled the control group on insight, the margin was small. Since insight is directly related to problem-solving, these last two findings indicate that the experimental group had not progressed in the acquisition of problem-solving skills. Several findings from other aspects of the experiment, coupled with the lack of problem-solving by either group just cited, indicate that "The Assembly-Line Problem" which was used at the "after" session, was more complex, and more difficult to solve by a fifteen-minute discussion. The finding that the judges were more able to agree concerning group solutions to "The Tennis Problem,"
which was used at the "before" session, seems to further support this contention.

The results of the judges' ratings of the group discussion transcripts may be considered as a trend in the predicted direction concerning group interaction process skills. The experimental groups did excel the control groups on insight, continuity, atmosphere, task-centeredness, group-centeredness, and demonstrated less individual-centered behavior, and more successful communication.

These results may be interpreted as providing some indication that the experimental group has made progress in the achievement of group interaction skills, and shows a trend toward improvement in the problem-solving skill of insight. This finding is important, because group interaction skills are generally acknowledged as prerequisite to adequate group problem-solving. Our results permit this progress to be attributed to the "case-discussion" method of instruction which the experimental group S.'s have experienced for three months. However, the gains have been modest and should be considered as trends in the predicted direction, as they do not represent statistically significant results.
CHAPTER XI.
DISCUSSION

Introduction

This chapter will comprise a summary and interpretation of the results which issued from the present study. Indications of adequacy of tools, techniques, and experimental design will be reported, as well as some interesting incidental findings. The relevance of the findings to curriculum improvement and further research will be drawn.

Adequacy of Experimental Tools and Techniques

The experimental procedure proved adequate in that every group was presented with a standard situation and all groups proceeded through both meetings according to plan. The problems sustained leaderless group discussion. By the use of recording equipment and the notation technique devised by E., the group discussions were reproduced in the form of typewritten transcripts with nearly perfect accuracy. The experimental procedure resulted in a complete set of data for each S. and a record of each group discussion.

Adequacy of Experimental Design

The experimental design included the use of matched S.'s for control and experimental groups, and the same group was measured before and after a three month training period which was experienced by only the experimental group. Empirical evidence supports the adequacy of the
matching procedure. At the "before" session the solutions written by the experimental and control S.'s were virtually identical. Moreover, the S.'s opinions and preferences concerning the problems were also very similar. According to the transcript judges, the experimental and control groups produced about the same number of group solutions at the "before" session. These judges were unable to distinguish the type of group from reading the "before" session transcripts. These findings reflected favorably upon the experimental design, since apparently both groups behaved in much the same manner at the "before" session, when differences were not predicted. There was one difference between groups which has already been pointed out, that the control group excelled the experimental group on the transcript criteria at the "before" session. However, the reason for this remains unknown.

**Incidental Findings**

Often an experiment uncovers incidental information which should be reported. Three such findings came to light in the present study. These aspects pertained to the reliability of the judges, the adequacy of the problems, and the problem-solving behavior of "the average college student".

The reliability of the judges was not high concerning the written solutions when reliability was determined in terms of the most rigorous criterion - absolute
agreement. However, when reliability was assessed according to dichotomies that were meaningful in that they involved adequate vs. inadequate characteristics of solutions, reliability was enhanced. Judged in terms of this criterion, reliability on four criteria were 57%, 66%, 74%, and 94%. It was concluded that the material was probably difficult to judge and that agreement might have been increased if the judges' training period had been increased. Unfortunately, certain limitations did not permit this.

The reliability of the judges in the matter of the transcript ratings was not high. The average Spearman Rank Difference Correlation Coefficient was .36. This low, positive correlation represents a minimum of reliability. This was an indirect measure, because it was not possible to measure reliability directly. It was only possible to determine what the lower level of the reliability would be.

Neither the judges who were psychologists nor the judges who were "case-discussion" method experienced teachers excelled in ability to judge type of group from reading the transcripts. The judges proved reliable in that five pairs of judges who agreed that a group solution had been reached by group discussion, in each case, each pair of judges wrote identical descriptive statements of the solution.

The adequacy of the problems is worthy of comment, since teachers and experimenters are constantly engaged in the quest for problems that provoke and sustain discussion.
While there were differences among the S.'s concerning preferences and opinions, all three problems seem useful media for leaderless group discussion. They would probably prove fruitful for a classroom discussion in which the teacher took part. They may be considered appropriate in relation to course goals in that the problems are not narrow, but involve far-reaching social implications. Since there is evidence that "The Assembly-Line Problem" is the most complex of the three, this problem probably should be distributed and clarified in advance of discussion. Since the essentials of the other two problems can be grasped almost immediately, advance distribution does not seem necessary.

The groups were carefully matched before experimentation began, and experimental results verified the adequacy of the matching procedure, for the groups tended to have similar preferences and opinions concerning the problems. They also wrote remarkably similar solutions to "The Tennis Problem". At this stage of the experiment, neither group had had experience in "case-discussion" method courses. Analysis of the results indicated that while the goal of the S.'s might have been to solve the problem by the most constructive means, their approach was too narrow and they fell into certain pitfalls.
Major Findings

In this section the major findings from the present study will be analyzed and integrated into a meaningful pattern. The first discovery of differences between experimental and control groups came from analysis of the "leadership" rating scale. It was found that there was no significant difference between the groups pertaining to the perception of the individuals. vs. the groups' perception of the "leadership" hierarchy. However, further analysis indicated that this may have been due to the fact that the range of "leadership" ratings for the experimental groups was much more narrow than that of the control groups. It was determined that this difference was almost statistically significant, (P. < .09). This was interpreted as suggesting that the experimental groups had excelled the control groups in that the experimental groups tended to share leadership functions. This outcome was consistent with other predictions that were made.

An analysis of the written solutions to "The Assembly-Line Problem" provided four trends in the predicted direction, one contrary trend, and one inconclusive finding. None of these was statistically significant. The first trend involved attempts to keep Joe on the same work team. This may be interpreted as evincing a desire to contain the problem - to solve it with the minimum shifting or disturbance of personnel and interpersonal relationships. The second trend showed that the experimental group produced
more solutions that attempted to use Joe productively. This may be interpreted as demonstrating a desire to use the potential of the individual to bring satisfaction to the worker and to the company. This inference is supported by the third trend, in which the experimental group excelled the control group by centering the problem in the most adequate manner. This may be interpreted as evidence that the experimental S.'s tended to take into account the total situation, rather than approach the problem from a restricted viewpoint. It will be remembered that Wertheimer emphasized the importance of centering by stating that adequate problem-solving demanded adequate centering. The fourth trend in which the experimental group demonstrated superiority was the finding that the experimental group produced fewer solutions that violated conditions which were stated in the problem. This was interpreted as meaning that the experimental S.'s were more willing to accept the limitations of the situation. It is generally agreed that this is a more realistic approach to problem-solving, than an approach which distorts the problem situation so that it may fit a particular solution. The latter approach may be considered an evasion of the task.

This last finding permits the resolution of alternate interpretations which were advanced concerning the S.'s opinions of the problems. It will be remembered that the experimental S.'s voiced the criticism that relevant information was lacking more often than did the control S.'s.
The alternative interpretation was proposed that either they were unwilling to accept the limitations of the situation, which was considered an unfruitful viewpoint, or that they keenly sensed that important information was lacking, which was considered a reasonable criticism. The fourth trend supports the latter interpretation. By their behavior in the task of writing solutions, the experimental group demonstrated that they willing to accept the limitations of the situation, and their criticisms may be interpreted as meaning that they were more keenly aware of the lack of certain information than were the control group.

In one respect the groups showed no differences. Both achieved identical scores concerning inadequacies which were detected by the judges in their solutions. The one contrary trend stands out clearly. The experimental groups produced a greater number of solutions which involved retirement. This was considered a drastic method of solving the problem since other alternatives were possible. It should be stated that no S. proposed retirement as the only solution, retirement was presented as an alternative. Therefore, it may have been considered as a last resort if other proposals failed. In this context, retirement is preferable to discharge, for retirement would benefit the worker by providing an income and would benefit the company by freeing them to hire a capable replacement. However, it would seem that the psychological
repercussions - the impact on the worker and the effect of this action on company-worker relationships, might outweigh these benefits.

The conclusion drawn from analysis of the written solutions was that the margins of superiority of the experimental group were not large. This may be considered a reflection of the difficulty of the problem. Perhaps we should not expect an individual to solve adequately a problem which six individuals acting as a group have proved incapable of solving after fifteen minutes of discussion. This aspect of the study may be summarized by stating that the experimental group excelled the control group on four measures, was excelled by the control group on one measure, and paralleled their performance on another one.

In the matter of group discussion transcripts, there were no differences between groups on the variable that pertained to the emergence of a solution from the group discussion, and the adequacy of such a solution - not one group achieved a solution, all ran out of time at the "after" session. On each of the remaining variables, the "t" tests and inspection of the relative mean gain differences indicated a consistent trend. The experimental group excelled the control group on each of the following seven variables: insight, continuity, atmosphere, task-centeredness, group-centeredness, individual-centeredness, and success of communication. Since it is generally agreed that these variables are important to group discussion
processes, this performance is important. The score for
the variable, "atmosphere" was statistically significant,
\( P < .03 \).

It should be mentioned that one of these variables, "insight" is more directly related to problem-solving than are
the others. The margin of excellence on "insight" was low,
which seems consistent with the other low findings related
to problem-solving. The fact that the experimental group
functioned more adequately in group discussion processes
suggests that they have made more progress in this direc-
tion than in the direction of group problem-solving. This
may be a reflection upon their training or upon the diff-
iculty of the problem.

The fact that the superiority of the experimental group
was small may be related to two other findings. Of the var-
ious measures used to investigate differences in perfor-
mance between groups, two other measures indicated that
there were no differences in performance, the Communica-
tion Data Form and the Index of Participation. At the time
of analysis, it was suggested that either the groups did
not differ, that these measures had proved too crude to de-
tect subtle differences, or that these measures were not
valid.

The transcript judgment analysis suggests that the
overall performance of the experimental group surpasses
that of the control group, that the former group has made
more progress in group discussion skills. This is import-
ant, for it is generally agreed that these skills are pre-
requisite to group problem-solving skills.

Recommendations for Curriculum Change

The fruitfulness of a study requires that it yields im-
plications which extend beyond the experimental situation.
The writer trusts that the present study will prove heur-
istic, and believes that the results warrant the suggest-
on of certain practical applications and further exper-
imental possibilities.

In the matter of curriculum improvement, the data tend
to give some support to the contention of other studies
that "case-discussion" courses provide valuable experi-
ences that are not usually gained in the typical lecture
method course. Concerning "case-discussion" courses of the
Human Relations type, it seems that it might prove profit-
able to extend the training period, alter the type of train-
ing, or effect some combination of these possibilities.

It often happens that training in "case-discussion"
classes cannot afford all members equal opportunities to
learn group interaction process and group problem-solving
skills because large enrollments curtail individual part-
ticipation. Our results indicate that students need to be-
come accustomed to vis-a-vis relationships apart from the
classroom situation, so that a transfer of learning will
be effected to areas more within the scope of daily inter-
personal relationships in college, in work, and in social
situations. Unless this happens, students may be content
to permit the teacher to assume responsibility for furthering group interaction processes. Our results seem to indicate that a promising beginning has been made, but there is no room for complacency; further progress is essential.

Perhaps more participation, and experience in small groups could be managed by increasing the training period and instituting changes in methodology. Under the guidance of the teacher, a permissive atmosphere for classroom discussion could be established at the beginning of the course. Later, the classroom group could be encouraged to institute its own observer-feedback-evaluation procedure, and these functions could be shared by the members. Concomitantly, students could be provided with experience in functioning as members of small groups, so that each individual would have a greater opportunity to function as a responsible group member. The institution of small, leaderless groups would tend to maximize individual responsibility. It is hoped that these suggestions do not appear too unrealistic, at least some combination of these possibilities might prove feasible. Finally, if some of these suggestions were put into practice, further experimentation would become imperative in order to test the effectiveness of such procedures.

Recommendations for Further Investigation

As a preliminary to future experimentation, some way of increasing the reliability of the judges should be devised.
If transcripts are used, it might be possible to employ a smaller number of judges so that conflicting frames of reference might be fewer. However, the work load may then become onerous. Perhaps reliability of transcript judgment may best be accomplished by means of a general orientation meeting attended by all judges so that differences could be reconciled, and the criteria could be further clarified. Unfortunately, these procedures were not feasible in the present study.

It might be possible to eliminate the use of transcriptions. Judges who worked as observers could be trained on other groups before experimentation was begun. If these observers rated the actual discussion behavior, the laborious process of transcribing the discussions could be eliminated. However, it would be advisable to have this important record of group behavior, for a group discussion transcript yields rich information.

Twenty minutes of the experimental period could be used to better advantage by eliminating the Communication Data Form and the "Leadership" Scale. The discussion period could then be extended twenty minutes. It may be that more time was needed to solve these problems through group discussion. If a group lacked the necessary skills, one would predict that extra time was of little value, but if the group possessed the skills but lacked the time, more solutions could be predicted.
If some of the curriculum changes which were suggested could be instituted, such innovations could function as variables that would permit an extension of the experimental design. For example, the present design could be ramified to include S.'s with two semesters of "case-discussion" method experience, S.'s with small group experience, S.'s with large group observer-feedback-evaluation experience, etc. These would serve to evaluate the suggestions offered on the basis of the results of the present study.

Conclusion

These gains made by the experimental group have been modest and they have been, with few exceptions, in the predicted direction, except for the data on communication and equality of participation. Statistically these gains were not significant, but this must be considered in the context of the situation in which they were accomplished.

The experimental situation represented a novel experience that was a wide departure from classroom routine as practiced in "case-discussion" method courses from which the experimental S.'s had been drawn. Moreover, there is evidence that "The Assembly-Line Problem" which was the crucial test, was a very complex problem. Finally, one must take into account the brevity of the training period - only three hours per week for three months - and the fact that this type of training was diametrically opposed to traditional methods.
When these conditions are taken into consideration, the writer believes it would be unrealistic to expect sweeping advances. In the writer's opinion, an appreciation of these aspects leads one to conclude that these modest gains represent an achievement which may be attributed to the "case-discussion" method of teaching.
CHAPTER XII.

SUMMARY

The purpose of the present study was to investigate the effectiveness of the "case-discussion" method of teaching by evaluating the extent to which certain goals had been achieved by this method of instruction. This study was conducted at the University of Kansas during the academic year of 1952-53.

The experimental design involved the use of experimental and control groups which were matched according to certain criteria, including friendship, experience in "case-discussion" courses, college standing, and grade point average. Twelve groups met twice during the same semester, the personnel of each group remained constant, and there was a three month interval between meetings. During this interval, the experimental group S.'s were having experience in the "case-discussion" method of teaching, but the control S.'s were not.

The experimental procedure at the "before" session began with a "warmup" case in order to accustom the S.'s to discussion in the presence of a microphone. This was followed by a fifteen minute leaderless group discussion of another case that was recorded with the knowledge of the S.'s. At the conclusion of the discussion, the S.'s individually wrote what they considered the "best" solution and why. Then they wrote brief opinions concerning the problems, and filled out a "leadership" rating scale.
The procedure at the "after" session followed the same general pattern, with two exceptions. A new, more difficult problem was used, and a "Communication Data Form" was filled out by the S.'s. Rapport was excellent, and a complete set of data was obtained from each S. The recordings of the twenty-four group discussions were transcribed into typewritten form.

Analysis of the "leadership" rating scale revealed that for the experimental groups, "leadership" tended to be a shared function, and this difference was almost statistically significant, (P.< .09). The written solutions were analyzed according to certain criteria by graduate students and faculty members who served as judges. This resulted in four trends favorable to the experimental group, one inconclusive finding, and one contrary trend. Concerning the favorable trends, the first trend was interpreted as a tendency to solve the problem with a minimum disturbance of personnel. The second indicated an attempt to solve the problem by using the potential of the individual in a manner that would bring satisfaction to the worker and the company, which was considered the most adequate solution. Another trend indicated that the experimental S.'s were attempting to diagnose the problem in the context of the total situation, rather than in a restricted manner. Moreover, the experimental S.'s seemed more willing to accept the limitations imposed by the problem, but their criticisms indicated that they were more
keenly aware of the lack of relevant information than were the control S.'s.

The inconclusive finding was that both groups achieved identical scores concerning inadequacies present in their proposed solutions. The contrary trend indicated that the experimental group proposed more solutions which involved retirement, which was considered a drastic action since better alternatives were available.

The conclusion drawn from the analysis of the written solutions was that the margins of superiority of the experimental group were not large. There was evidence which suggested that the problem used as the crucial test was complex, and the relatively small margin of the experimental group was considered in part, a reflection of the difficulty of the problem.

The group discussion transcripts were judged by twelve members of the University of Kansas Staff. Standard criteria were set up in terms of variables generally agreed upon as being important characteristics of group interaction processes. At the post session there was no difference between groups on the variable that pertained to the emergence of a solution from the group discussion and the adequacy of such a solution, for not one group achieved a solution.

However, the experimental group excelled the control group on each of the other seven variables. They were: insight, continuity, atmosphere, task-centeredness,
group-centeredness, individual-centeredness, and success of communication. These gains were modest but they were consistent. Since it is generally agreed that the achievement of group interaction skills are prerequisite to the mastery of group problem-solving skills, this progress may be interpreted as representing a promising beginning.

Statistically only one of these gains was significant, the difference on variable #4, atmosphere, \( (P. < .03) \). However, in view of the conditions under which they were achieved, they seem noteworthy. The experimental situation was a wide departure from classroom routine and the task of solving the "after" session problem was a rigorous one. Finally, one must take into account the brevity of the training period - only three hours per week for three months - and the fact that this type of training was diametrically opposed to the methods by which these S.'s had been trained throughout their schooling. In view of these considerations, it was suggested that these gains represented an achievement that could very likely be attributed to the "case-discussion" method of teaching.

* * *
APPENDIX

1. Confidential Data Form..........................164
2. The Information Form, Coding System...........165
3. The Coach's Dilemma............................166
4. The Tennis Problem..............................167
5. The Assembly-Line Problem......................168
6. Leadership Rating Scale.........................169
7. Communication Data Form.........................170
8. Table XVIII, Index of Participation
   Formula........................................171
9. Table XIX, Index of Participation..............172
10. Table XX, Summary Chart for Communication
     Data...........................................173
11. Table XXI, Summary Chart for Communication
     Data: Subjects - A
         Group as a Whole - B....................174
12. Table XXII, Agreement in Perception of
     Communication Channels......................175
13. Table XXIII, Agreement in Perception of
     Communication Channels (Successes Only).....176
14. Table XXIV, Correlation Coefficients and
    Z Transformations, A, B, C, D................177-180
15. Table XXV, Range on Leadership Scale.........181
16. Criteria for Judging Written Solutions,
    Tennis Problem..............................182
17. Judge's Scoring Form - Tennis Problem........183
18. Samples of Solutions to "The Tennis Problem..184
19. Criteria for Judging the Written Solutions,
    Assembly-Line Problem.......................185 & 186
### APPENDIX - CONTINUED

<table>
<thead>
<tr>
<th>Sample/Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>20. Samples of Solutions to the Assembly-Line Problem</td>
<td>187</td>
</tr>
<tr>
<td>21. Letter to the Judges</td>
<td>188</td>
</tr>
<tr>
<td>22. Criteria for Judging the Transcripts</td>
<td>189-190</td>
</tr>
<tr>
<td>23. Transcript Rating Scale</td>
<td>191</td>
</tr>
<tr>
<td>24. Transcript</td>
<td>192-203</td>
</tr>
</tbody>
</table>
CONFIDENTIAL DATA SHEET

Name ___________________________ School ___________________________
(last) (first) (initial) (ex: Bus., Eng., Coll.)
Lawrence Address ___________________________ Class ___________________________
_________________________________________ (soph) (jr) (sr)
Phone # ___________________________ Teacher ___________________________
_________________________________________ (this class)
Marr. or Single ___________________________ Name & No. ___________________________
______________________________ (this class)
Age ___________________________
______________________________

Below enter your FREE afternoon hours. Ex: 1-2; 3-4; 1-5, etc (when not in class, lab, etc.)

Monday ___________________________
Tuesday ___________________________
Wednesday ___________________________
Thursday ___________________________
Friday ___________________________

Concerning whether you have had or are now taking the following courses, please write "yes," "no," or "now."

Business Administration Practices I or II (B.A.P.) ___________________________
Fundamentals of Speech ___________________________
Human Relations ___________________________
Names of Sociology Courses ___________________________
Experience in Debate ___________________________
Drama ___________________________
Pub. Speak. ___________________________
THE INFORMATION FORM CODING SYSTEM

The information forms were classified according to control or experimental groups and arranged according to courses by a simple, but effective numerical coding system devised by E. In order to eliminate confusing experimental and control S.'s control classes were marked with red numbers, experimental classes with blue, and every class or section of a course was given a different number. The experimental classes were numbered in blue from 1-10. Each control class or section of a course was given a different number and the information forms from one class were all given the same number. The control classes were handled likewise. Furthermore, the first digit was the key which designated which school these control classes were from. For example, the series 10-19 indicated classes from the School of Business, the series 20-29 indicated classes from the School of Journalism, etc. In arranging a group, this system proved most helpful in that it ensured that the S.'s had been drawn from different classes, thus minimizing friendship factors.
THE COACH'S DILEMMA

The old basketball coach at school was talking to the new coach. "You'll have trouble with John, he's too cocky. He is just fifteen years old and the star player on the 'Juniors'. He's been a member of the team and has a reputation for insisting on special privileges, like coming into the gym when it is closed, or borrowing extra equipment. And does he complain when he doesn't get what he wants! His parents are foreigners and he lives in a very poor neighborhood. Oddly enough, his friends come from much better neighborhoods. Because of his ability, the team tolerates him, but they don't really like him. I heard them say they wish he'd leave. Well, good luck, I have to go now."

When the new coach appeared, the entire team began testing the limits by cutting up. They wanted to see just how much they could get away with. The coach said nothing, but gave the offenders a disapproving look. After John had received several dirty looks, he deliberately stalled the ball. When the coach walked over, John swore at him. Just then, the bell rang, ending the period, everyone had to clear out of the gym immediately. WHAT SHOULD THE COACH DO? WHY?
The two boys were playing tennis. I could hear as well as see them from my window, although they did not see me. One boy was twelve, the other ten years old. They played several sets. The younger boy was by far poorer, he was beaten in all the games. I heard some of their conversation. The loser - let us call him B - became more and more unhappy. He had no chance. A often served him so cleverly that he could not possibly return the ball. The situation grew worse. Finally, B threw down his racket, sat on a tree trunk, and said, "I won't play any more." A tried to persuade him to continue. No answer from B. A sat down beside him. Both looked unhappy.

What should A do?

Why?
THE ASSEMBLY LINE PROBLEM

Visualize a sub-assembly situation in which seven men, working in a circle, assemble a part of a car (carburetor or instrument panel, for example). The article enters the circle at one point, and each person adds his pieces and pushes the unit to the next worker who adds his elements. When the unit leaves the circle, it is a completed part product.

Suppose there are four such parasol sub-assembly stations, each one supervised by a foreman. Suppose further that Station A assembles 85 units per day; Station B, 80 per day; Station C, 60 units per day and Station D, 50 units. It is a fact that Station D previously assembled 60 units. The foreman was dissatisfied with the production and reprimanded the group. Following the reprimand production fell to 50 units per day. SEE DIAGRAM.

The assembly work is simple and requires a minimum of training for each step. The aptitude requirement is primarily good finger dexterity. The materials for each assembly position are located in bins which are kept supplied by material handlers. Thus each worker has his essential material at his elbow. The job has been analyzed by time-and-motion experts so that the positions are of equal difficulty. Pay is based on hourly rates.

The total factory production is dependent upon receiving the required number of assembled units from these four stations. The production is now so low that the factory production as a whole had to slow down. The desired quota is 300 parts per shift for the four stations combined. ANY QUESTIONS?

We are concerned with Station C, producing at the rate of 60 units. The work piles up at the position of Joe Brown. The unit must pass through him (position V), he always has several piled up waiting for him. Foremen
on non-production jobs are not willing to accept Joe as a transfer. Joe is a man of 60 with 30 years of service in the company. Emphasis on improving production has brought his deficiencies to light. Joe is a nice congenial fellow. He is liked by the group of workers, but is slow and unhandy. However, he tries. Joe has to work for a living. WHAT WOULD BE THE BEST SOLUTION?
"LEADERSHIP" RATING SCALE

Number ______________  Name ____________________________
(last)  (first)

Rating Scale: A reminder - your name and all data
gathered in this experiment will be kept strictly con-
fidential. Your judgment is of vital importance to
this experiment.

On the following rating scale, please rate each in-
dividual, including yourself.

Mark your judgment by placing a vertical line for
each individual, including yourself, and writing the in-
dividual's number on top of the line. You may place your
vertical lines anywhere along the horizontal line. When
completed, you should have six vertical lines along the
horizontal line, and an identifying number on top of each
line. For example:

```
  3 1 5 6 2 4
```

* * * * *

Think in general terms of what you would consider
"leadership ability" in working through this problem.
Please rank each individual, including yourself. Mark
with a vertical line the one you consider "best," on
the right as far up as you judge him, then place each in-
dividual in order, behind him, where you think they belong.

least

most
COMMUNICATION DATA

Communication, as used here, concerns the transmission of ideas, information, feelings, etc. Please note below with an "X":

A. Those with whom you specifically tried to communicate verbally.

B. Whether or not you successfully communicated. If you did not try to communicate, leave the second column blank.

<table>
<thead>
<tr>
<th>Member Number</th>
<th>I tried to communicate</th>
<th>I was successful</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group as a whole</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A. Now please note those whom you believe specifically tried to communicate verbally with you.

B. Whether or not their attempts were successful. If they did not try to communicate, leave the second column blank.

<table>
<thead>
<tr>
<th>Member Number</th>
<th>Tried to communicate with me</th>
<th>Was successful</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE XVIII.

INDEX OF PARTICIPATION FORMULA*

\[
I.P. = 100 \left( 1 - \frac{2[N \bar{x}^2 - (\sum x)^2]}{(N - 2)(\sum x)^2} \right)^2
\]

Example:

<table>
<thead>
<tr>
<th>Individual</th>
<th>Frequency (x)</th>
<th>(x^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>23</td>
<td>529</td>
</tr>
<tr>
<td>B</td>
<td>17</td>
<td>289</td>
</tr>
<tr>
<td>C</td>
<td>12</td>
<td>144</td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>E</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>F</td>
<td>16</td>
<td>256</td>
</tr>
<tr>
<td>G</td>
<td>9</td>
<td>81</td>
</tr>
<tr>
<td>H</td>
<td>12</td>
<td>144</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td><strong>104</strong></td>
<td><strong>1568</strong></td>
</tr>
</tbody>
</table>

\[
I.P. = 100 \left( 1 - \frac{2[8(1568) - (104)^2]}{6(104)^2} \right)^2
\]

\[
= 100 \left( 1 - \frac{2(128564 - 10816)}{6(10816)} \right)^2 = 100 \left( 1 - \frac{3456}{64896} \right)^2
\]

\[
= 100(1 - .053)^2 = 100(.947)^2 = 90.
\]

### TABLE XIX. INDEX OF PARTICIPATION

<table>
<thead>
<tr>
<th>Before Session</th>
<th>After Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td>Score</td>
</tr>
<tr>
<td>1</td>
<td>92.</td>
</tr>
<tr>
<td>2</td>
<td>76.9</td>
</tr>
<tr>
<td>3</td>
<td>75.8</td>
</tr>
<tr>
<td>4</td>
<td>73.4</td>
</tr>
<tr>
<td>5</td>
<td>68.5</td>
</tr>
<tr>
<td>6</td>
<td>66.49</td>
</tr>
<tr>
<td>7</td>
<td>66.09</td>
</tr>
<tr>
<td>8</td>
<td>55.5</td>
</tr>
<tr>
<td>9</td>
<td>55.05</td>
</tr>
<tr>
<td>10</td>
<td>52.4</td>
</tr>
<tr>
<td>11</td>
<td>51.</td>
</tr>
<tr>
<td>12</td>
<td>37.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Both Sessions</th>
<th>Group #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td>Score</td>
</tr>
<tr>
<td>1</td>
<td>92.</td>
</tr>
<tr>
<td>2</td>
<td>91.3</td>
</tr>
<tr>
<td>3</td>
<td>87.6</td>
</tr>
<tr>
<td>4</td>
<td>83.7</td>
</tr>
<tr>
<td>5</td>
<td>80.1</td>
</tr>
<tr>
<td>6</td>
<td>78.4</td>
</tr>
<tr>
<td>7</td>
<td>76.9</td>
</tr>
<tr>
<td>8.5</td>
<td>75.8</td>
</tr>
<tr>
<td>8.5</td>
<td>75.8</td>
</tr>
<tr>
<td>10</td>
<td>73.4</td>
</tr>
<tr>
<td>11.5</td>
<td>68.5</td>
</tr>
<tr>
<td>11.5</td>
<td>68.6</td>
</tr>
<tr>
<td>13</td>
<td>68.49</td>
</tr>
<tr>
<td>14</td>
<td>67.7</td>
</tr>
<tr>
<td>15</td>
<td>66.09</td>
</tr>
<tr>
<td>16</td>
<td>60.9</td>
</tr>
<tr>
<td>17</td>
<td>60.68</td>
</tr>
<tr>
<td>18</td>
<td>55.8</td>
</tr>
<tr>
<td>19</td>
<td>55.5</td>
</tr>
<tr>
<td>20</td>
<td>55.05</td>
</tr>
<tr>
<td>21</td>
<td>53.8</td>
</tr>
<tr>
<td>22</td>
<td>52.4</td>
</tr>
<tr>
<td>23</td>
<td>51.</td>
</tr>
<tr>
<td>24</td>
<td>37.2</td>
</tr>
</tbody>
</table>

**Key**

- B = Before Session
- A = After Session
- C = Control Group
- E = Experimental Group
- # = # of Group
TABLE XX.

SUMMARY CHART FOR COMMUNICATION DATA

GROUP A & G #1

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0</td>
<td>00</td>
<td>/\</td>
<td>/\</td>
<td>-\</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-0</td>
<td>00</td>
<td>00</td>
<td>/\</td>
<td>/\</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>00</td>
<td>00</td>
<td>00</td>
<td>/\</td>
<td>00</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>/\</td>
<td>00</td>
<td>00</td>
<td>00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>/\</td>
<td>/\</td>
<td>04</td>
<td>00</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>/\</td>
<td>04</td>
<td>04</td>
<td>00</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Weighted Score**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Weight</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>0-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0/</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>--</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>/-</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>//</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Weighted Score = 26

**Symbols:**

- 00 = Neither S. tried to communicate.
- 0- = One S. tried, and was not successful.
- 0/ = One S. tried and was successful.
- -- = Both S.'s tried, neither was successful.
- /- = Both S.'s tried, one was successful.
- // = Both S.'s tried, both were successful.

This chart should be read as follows:

We note that the cell which appears at column 1, row 2, contains the symbols: (-0). The symbol on the left (-), indicates that S.#1 tried to communicate with S.#2 and was not successful. The symbol on the right (0) indicates that S.#2 did not try to communicate with S.#1. Thus the symbols (-0) mean "one S. tried, and was not successful."

The totals in the Weighted Score diagram indicate the number of times each of the six possible communication combinations (represented by the six combinations of symbols), occurred.

The bottom row "G" indicates the number of individuals who tried to communicate with the group as a whole and whether or not they were successful.
# TABLE XXI

## A. SUMMARY CHART FOR COMMUNICATION DATA
**(SUBJECTS)**

<table>
<thead>
<tr>
<th>Rank</th>
<th>Score</th>
<th>Group #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>43</td>
<td>A E #6</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td>A C #4</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>A E #2</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>A E #6</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>A C #2</td>
</tr>
<tr>
<td>7</td>
<td>33</td>
<td>A E #3</td>
</tr>
<tr>
<td>7</td>
<td>33</td>
<td>A E #1</td>
</tr>
<tr>
<td>7</td>
<td>33</td>
<td>A C #3</td>
</tr>
<tr>
<td>9</td>
<td>31</td>
<td>A E #4</td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>A C #5</td>
</tr>
<tr>
<td>11</td>
<td>27</td>
<td>A C #6</td>
</tr>
<tr>
<td>12</td>
<td>26</td>
<td>A C #1</td>
</tr>
</tbody>
</table>

## B. SUMMARY CHART FOR COMMUNICATION DATA
**(GROUP AS A WHOLE)**

<table>
<thead>
<tr>
<th>Rank</th>
<th>Score</th>
<th>Group #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>12</td>
<td>A E #6</td>
</tr>
<tr>
<td>1.5</td>
<td>12</td>
<td>A C #5</td>
</tr>
<tr>
<td>4.5</td>
<td>11</td>
<td>A C #6</td>
</tr>
<tr>
<td>4.5</td>
<td>11</td>
<td>A E #4</td>
</tr>
<tr>
<td>4.5</td>
<td>11</td>
<td>A C #3</td>
</tr>
<tr>
<td>4.5</td>
<td>11</td>
<td>A E #2</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>A E #1</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>A E #3</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>A E #5</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>A C #4</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>A C #1</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>A C #2</td>
</tr>
</tbody>
</table>

### Key

- **A** = After Session
- **C** = Control Group
- **E** = Experimental Group
- **#** = # of Group
**TABLE XXII.**

**AGREEMENT IN PERCEPTION OF COMMUNICATION CHANNELS**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>****</td>
<td>6</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>*</td>
<td></td>
<td>-</td>
<td>****</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>*</td>
<td></td>
<td>*</td>
<td>-</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>*</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>***</td>
<td></td>
<td></td>
<td>-</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Totals 17 13

For Example: Notice the star in the cell at Row 1, Column 6. This indicates that whatever S.#1 has stated concerning S.#6; S.#6 has stated concerning S.#1.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Score</th>
<th>Group #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23</td>
<td>A C #2</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>A C #6</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>A E #2</td>
</tr>
<tr>
<td>4.5</td>
<td>17</td>
<td>A E #6</td>
</tr>
<tr>
<td>4.5</td>
<td>17</td>
<td>A C #5</td>
</tr>
<tr>
<td>6.5</td>
<td>16</td>
<td>A E #1</td>
</tr>
<tr>
<td>6.5</td>
<td>16</td>
<td>A C #4</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>A E #3</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>A E #4</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>A E #5</td>
</tr>
<tr>
<td>11.5</td>
<td>14</td>
<td>A C #1</td>
</tr>
<tr>
<td>11.5</td>
<td>14</td>
<td>A C #3</td>
</tr>
</tbody>
</table>

**Key**

A = After Session  
C = Control Group  
E = Experimental Group  
# = # of Group
### Table XXIII.

**Agreement in Perception of Communication Channels (Successes Only)**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>2</strong></td>
</tr>
<tr>
<td>2</td>
<td><strong>2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>2</strong></td>
</tr>
<tr>
<td>3</td>
<td><strong>1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>1</strong></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>0</strong></td>
</tr>
<tr>
<td>5</td>
<td><strong>4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>4</strong></td>
</tr>
<tr>
<td>6</td>
<td><strong>4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>4</strong></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>13</strong></td>
</tr>
</tbody>
</table>

For Example: Notice the star in the cell at Row 1, Column 6. This indicates that S.#1 stated that he tried and was successful in communicating with S.#6, and that S.#6 stated that he perceived S.#1 trying to communicate with him, and that S.#1 had been successful.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Score</th>
<th>Group #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>AC #2</td>
</tr>
<tr>
<td>2.5</td>
<td>13</td>
<td>AC #3</td>
</tr>
<tr>
<td>2.5</td>
<td>13</td>
<td>AE #5</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>AE #6</td>
</tr>
<tr>
<td>5.5</td>
<td>10</td>
<td>AC #5</td>
</tr>
<tr>
<td>5.5</td>
<td>10</td>
<td>AE #2</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>AC #6</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>AE #3</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>AE #4</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>AE #1</td>
</tr>
<tr>
<td>11.5</td>
<td>5</td>
<td>AC #1</td>
</tr>
<tr>
<td>11.5</td>
<td>5</td>
<td>AC #3</td>
</tr>
</tbody>
</table>

**Key**

A = After Session  
C = Control Group  
E = Experimental Group  
# = # of Group
TABLE XXIV-A
CORRELATION COEFFICIENTS AND TRANSFORMATIONS

<table>
<thead>
<tr>
<th>Group G #4</th>
<th>Group G #4</th>
<th>Group G #5</th>
<th>Group G #5</th>
<th>Group G #6</th>
<th>Group G #6</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. #</td>
<td>Before</td>
<td>After</td>
<td>D</td>
<td>S. #</td>
<td>Before</td>
</tr>
<tr>
<td>1</td>
<td>.37</td>
<td>.20</td>
<td>- .17</td>
<td>1</td>
<td>.888</td>
</tr>
<tr>
<td>2</td>
<td>.56</td>
<td>.77</td>
<td>- .21</td>
<td>2</td>
<td>.632</td>
</tr>
<tr>
<td>3</td>
<td>.20</td>
<td>.43</td>
<td>- .23</td>
<td>3</td>
<td>.802</td>
</tr>
<tr>
<td>4</td>
<td>.94</td>
<td>.76</td>
<td>- .18</td>
<td>4</td>
<td>1.738</td>
</tr>
<tr>
<td>5</td>
<td>.77</td>
<td>.66</td>
<td>- .11</td>
<td>5</td>
<td>1.020</td>
</tr>
<tr>
<td>6</td>
<td>.49</td>
<td>.56</td>
<td>- .07</td>
<td>6</td>
<td>.538</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Average &amp; Score</td>
<td></td>
</tr>
<tr>
<td>Group G #5</td>
<td>Group G #5</td>
<td>Group G #5</td>
<td>Group G #5</td>
<td>Group G #6</td>
<td>Group G #6</td>
</tr>
<tr>
<td>S. #</td>
<td>Before</td>
<td>After</td>
<td>D</td>
<td>S. #</td>
<td>Before</td>
</tr>
<tr>
<td>1</td>
<td>.45</td>
<td>.59</td>
<td>- .16</td>
<td>1</td>
<td>.888</td>
</tr>
<tr>
<td>2</td>
<td>.71</td>
<td>.66</td>
<td>- .05</td>
<td>2</td>
<td>.887</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>.66</td>
<td>- .34</td>
<td>3</td>
<td>3.000</td>
</tr>
<tr>
<td>4</td>
<td>.83</td>
<td>.93</td>
<td>- .10</td>
<td>4</td>
<td>1.188</td>
</tr>
<tr>
<td>5</td>
<td>.94</td>
<td>.71</td>
<td>- .23</td>
<td>5</td>
<td>1.738</td>
</tr>
<tr>
<td>6</td>
<td>.83</td>
<td>.45</td>
<td>- .40</td>
<td>6</td>
<td>1.188</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Average &amp; Score</td>
<td></td>
</tr>
<tr>
<td>S. #</td>
<td>Before</td>
<td>After</td>
<td>D</td>
<td>S. #</td>
<td>Before</td>
</tr>
<tr>
<td>1</td>
<td>.94</td>
<td>.71</td>
<td>- .23</td>
<td>1</td>
<td>1.738</td>
</tr>
<tr>
<td>2</td>
<td>.54</td>
<td>.60</td>
<td>- .06</td>
<td>2</td>
<td>.604</td>
</tr>
<tr>
<td>3</td>
<td>.77</td>
<td>.71</td>
<td>- .06</td>
<td>3</td>
<td>1.020</td>
</tr>
<tr>
<td>4</td>
<td>.94</td>
<td>.83</td>
<td>- .11</td>
<td>4</td>
<td>1.738</td>
</tr>
<tr>
<td>5</td>
<td>.94</td>
<td>.60</td>
<td>- .34</td>
<td>5</td>
<td>1.738</td>
</tr>
<tr>
<td>6</td>
<td>.89</td>
<td>.67</td>
<td>- .22</td>
<td>6</td>
<td>1.421</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Average &amp; Score</td>
<td></td>
</tr>
</tbody>
</table>
### TABLE XXIV-B

**Correlation Coefficients and Z Transformations**

<table>
<thead>
<tr>
<th>Group C #1</th>
<th>Group C #1</th>
<th>Group C #2</th>
<th>Group C #2</th>
<th>Group C #3</th>
<th>Group C #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>S.#</td>
<td>Before</td>
<td>After</td>
<td>D</td>
<td>S.#</td>
<td>Before</td>
</tr>
<tr>
<td>1</td>
<td>.60</td>
<td>.54</td>
<td>-.06</td>
<td>1</td>
<td>.693</td>
</tr>
<tr>
<td>2</td>
<td>.60</td>
<td>1.00</td>
<td>.40</td>
<td>2</td>
<td>.693</td>
</tr>
<tr>
<td>3</td>
<td>.77</td>
<td>.77</td>
<td>---</td>
<td>3</td>
<td>1.020</td>
</tr>
<tr>
<td>4</td>
<td>.76</td>
<td>.83</td>
<td>.07</td>
<td>4</td>
<td>.996</td>
</tr>
<tr>
<td>5</td>
<td>.93</td>
<td>.77</td>
<td>-.16</td>
<td>5</td>
<td>1.658</td>
</tr>
<tr>
<td>6</td>
<td>-.11</td>
<td>.54</td>
<td>.65</td>
<td>6</td>
<td>.110</td>
</tr>
</tbody>
</table>

Average Z Score: .377

Average Z Score: .525

Average Z Score: .317

Average Z Score: .661
TABLE XXIV-G

CORRELATION COEFFICIENTS AND Z TRANSFORMATIONS

<table>
<thead>
<tr>
<th>Groups</th>
<th>Correlations</th>
<th>z Transformations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group E #4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. #</td>
<td>Before</td>
<td>After</td>
</tr>
<tr>
<td>1</td>
<td>- .04</td>
<td>.11</td>
</tr>
<tr>
<td>2</td>
<td>.45</td>
<td>.14</td>
</tr>
<tr>
<td>3</td>
<td>.66</td>
<td>.89</td>
</tr>
<tr>
<td>4</td>
<td>.09</td>
<td>.89</td>
</tr>
<tr>
<td>5</td>
<td>.49</td>
<td>.89</td>
</tr>
<tr>
<td>6</td>
<td>- .04</td>
<td>.94</td>
</tr>
<tr>
<td>Average &amp; Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group E #5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. #</td>
<td>Before</td>
<td>After</td>
</tr>
<tr>
<td>1</td>
<td>.43</td>
<td>.66</td>
</tr>
<tr>
<td>2</td>
<td>.63</td>
<td>.94</td>
</tr>
<tr>
<td>3</td>
<td>.44</td>
<td>.89</td>
</tr>
<tr>
<td>4</td>
<td>.66</td>
<td>.83</td>
</tr>
<tr>
<td>5</td>
<td>.93</td>
<td>.37</td>
</tr>
<tr>
<td>6</td>
<td>.97</td>
<td>.89</td>
</tr>
<tr>
<td>Average &amp; Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group E #6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. #</td>
<td>Before</td>
<td>After</td>
</tr>
<tr>
<td>1</td>
<td>.89</td>
<td>.26</td>
</tr>
<tr>
<td>2</td>
<td>.89</td>
<td>.66</td>
</tr>
<tr>
<td>3</td>
<td>.63</td>
<td>.20</td>
</tr>
<tr>
<td>4</td>
<td>.77</td>
<td>.20</td>
</tr>
<tr>
<td>5</td>
<td>.89</td>
<td>.26</td>
</tr>
<tr>
<td>6</td>
<td>.60</td>
<td>.67</td>
</tr>
<tr>
<td>Average &amp; Score</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Average & Score
TABLE XXIV-D
CORRELATION COEFFICIENTS AND % TRANSFORMATIONS

Correlations

<table>
<thead>
<tr>
<th>Group E #1</th>
<th></th>
<th></th>
<th>Group E #2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S.#</td>
<td>Before</td>
<td>After</td>
<td>D</td>
<td>S.#</td>
<td>Before</td>
</tr>
<tr>
<td>1</td>
<td>.89</td>
<td>-.53</td>
<td>-.142</td>
<td>1</td>
<td>1.42</td>
</tr>
<tr>
<td>2</td>
<td>.83</td>
<td>-.14</td>
<td>-.97</td>
<td>2</td>
<td>1.18</td>
</tr>
<tr>
<td>3</td>
<td>.37</td>
<td>-.77</td>
<td>-.114</td>
<td>3</td>
<td>.388</td>
</tr>
<tr>
<td>4</td>
<td>.09</td>
<td>-.14</td>
<td>-.23</td>
<td>4</td>
<td>.090</td>
</tr>
<tr>
<td>5</td>
<td>.54</td>
<td>.67</td>
<td>.13</td>
<td>5</td>
<td>.604</td>
</tr>
<tr>
<td>6</td>
<td>.09</td>
<td>.53</td>
<td>.44</td>
<td>6</td>
<td>.090</td>
</tr>
</tbody>
</table>

Average % Score = -.509

% Transformations

<table>
<thead>
<tr>
<th>Group E #1</th>
<th></th>
<th></th>
<th>Group E #2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S.#</td>
<td>Before</td>
<td>After</td>
<td>D</td>
<td>S.#</td>
<td>Before</td>
</tr>
<tr>
<td>1</td>
<td>.16</td>
<td>.59</td>
<td>.43</td>
<td>1</td>
<td>.161</td>
</tr>
<tr>
<td>2</td>
<td>.48</td>
<td>-.66</td>
<td>-.114</td>
<td>2</td>
<td>.522</td>
</tr>
<tr>
<td>3</td>
<td>-.60</td>
<td>-.20</td>
<td>.40</td>
<td>3</td>
<td>-.693</td>
</tr>
<tr>
<td>4</td>
<td>.77</td>
<td>-.66</td>
<td>-.43</td>
<td>4</td>
<td>1.020</td>
</tr>
<tr>
<td>5</td>
<td>.66</td>
<td>-.53</td>
<td>-.19</td>
<td>5</td>
<td>.792</td>
</tr>
<tr>
<td>6</td>
<td>.37</td>
<td>-.20</td>
<td>.57</td>
<td>6</td>
<td>.588</td>
</tr>
</tbody>
</table>

Average % Score = .020

<table>
<thead>
<tr>
<th>Group E #3</th>
<th></th>
<th></th>
<th>Group E #3</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S.#</td>
<td>Before</td>
<td>After</td>
<td>D</td>
<td>S.#</td>
<td>Before</td>
</tr>
<tr>
<td>1</td>
<td>.88</td>
<td>.77</td>
<td>-.11</td>
<td>1</td>
<td>1.375</td>
</tr>
<tr>
<td>2</td>
<td>.37</td>
<td>.20</td>
<td>-.17</td>
<td>2</td>
<td>.388</td>
</tr>
<tr>
<td>3</td>
<td>.77</td>
<td>.04</td>
<td>-.81</td>
<td>3</td>
<td>1.020</td>
</tr>
<tr>
<td>4</td>
<td>.89</td>
<td>.43</td>
<td>-.46</td>
<td>4</td>
<td>1.42</td>
</tr>
<tr>
<td>5</td>
<td>.57</td>
<td>.60</td>
<td>.03</td>
<td>5</td>
<td>.647</td>
</tr>
<tr>
<td>6</td>
<td>.83</td>
<td>.31</td>
<td>.52</td>
<td>6</td>
<td>1.186</td>
</tr>
</tbody>
</table>

Average % Score = -.550
TABLE XXV.

RANGE ON LEADERSHIP SCALE
AFTER SESSION

<table>
<thead>
<tr>
<th>Rank</th>
<th>Score</th>
<th>Group #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>80</td>
<td>A C #6</td>
</tr>
<tr>
<td>2</td>
<td>75</td>
<td>A E #5</td>
</tr>
<tr>
<td>3</td>
<td>68</td>
<td>A C #5</td>
</tr>
<tr>
<td>4</td>
<td>67</td>
<td>A C #2</td>
</tr>
<tr>
<td>5</td>
<td>65</td>
<td>A C #3</td>
</tr>
<tr>
<td>6</td>
<td>61</td>
<td>A C #4</td>
</tr>
<tr>
<td>7</td>
<td>48</td>
<td>A E #3</td>
</tr>
<tr>
<td>8</td>
<td>44</td>
<td>A E #4</td>
</tr>
<tr>
<td>9</td>
<td>36</td>
<td>A C #1</td>
</tr>
<tr>
<td>10</td>
<td>35</td>
<td>A E #6</td>
</tr>
<tr>
<td>11</td>
<td>32</td>
<td>A E #1</td>
</tr>
<tr>
<td>12</td>
<td>22</td>
<td>A E #2</td>
</tr>
</tbody>
</table>

**Key**

A = After Session
G = Control Group
E = Experimental Group
# = # of Group
CRITERIA FOR JUDGING WRITTEN SOLUTIONS - TENNIS PROBLEM

* Each solution will be rated according to categories A, B, and C.

** On the "Judge's Rating Form" space is allowed for comments, where you may note any unusual aspect, doubts as to category, or any feature of the solution not anticipated in the categories.

A: Criteria Concerning Centering:
1. "A" (older boy) centered: - a solution oriented in terms of "A"'s the older boy's needs.
2. "B" (younger boy) centered: - a solution oriented in terms of "B"'s, the younger boy's needs.
3. Game or Both "A" and "B" centered: - a solution which involves the needs of both boys.
4. "Outside of the Field" solutions: - The problem set up for the subject to solve is specifically, "What should "A" do?" Therefore, if someone else is brought into the case, this would seem an evasion of the problem, a sort of "magical" solution which goes beyond the limitations of the solution.
EX: "The person watching from the window should settle it."

*** Overlapping is not permitted. If a single solution contains elements of an "A" or "B" centered solution and a "Game" centered solution, score this as a "Game" centered solution.
If any element of a solution is an "Outside of the Field" aspect, the solution is marked in this category.

B: Criteria Concerning Inadequate Solutions:
* Overlapping is permitted, a solution may be scored in as many of the following categories as appropriate.

1. Pigeon-holing - involves applying a label or stereotype.
   EX: "This is obviously a case of inferiority complex."
2. Moralinizing - involves blaming, or moral judgment.
   EX: "A should know better than to treat B like this."
3. Ignoring restrictions, limitations, instructions, or adding unrealistic details:
   EX: An "Outside of the Field" solution.
   EX: Either an "A" or "B" centered solution.
5. "Reinterpreting" facts - Taking a given statement of conditions and altering it so as to permit a certain kind of solution.

C: Criteria Concerning Means:
* Overlapping NOT permitted - Judge in terms of the main theme.

1. Bargaining - entails a "Business deal" - "I do this, if you do that."
2. Bribing - offering an inducement, such as candy.
3. Flattering
4. Shaming
5. Mutual Means - using the present and potential abilities of each in a manner beneficial to both.
6. Withdrawal or quitting - walking away.
A: Criteria Concerning Centering:
(No overlapping)
1. "A" (older boy) centered——
2. "B" (younger boy) centered——
3. Game or both centered——
4. "Outside of the Field"——

B: Criteria Concerning Inadequate Solutions:
(Overlapping O.K.)
1. Pigeon-holing
2. Moralizing
3. Ignoring restrictions, limitations, etc.
4. Restricted frame of reference
5. "Reinterpreting" facts

C: Criteria Concerning Means:
(No Overlapping
1. Bargaining
2. Bribing
3. Flattering
4. Shaming
5. Mutual Means
6. Withdrawal, quitting
S. #1, Control Group #1:

A, very evidently the superior player, should take time to drill the younger player in the fundamentals of the game. This would improve both players' game, as I believe both are very much interested in the sport and seem to indicate a desire to become more proficient. A appears to be overeager and should slow down to enable B to catch up with him. This solution would give each boy a partner. Evidently A needs a partner as indicated by his strong insistence that B continue the game. B also needs one, else why would he continue as far as he did?

S. #6, Experimental Group #1:

I think that A should try and regain B's friendship and restore his confidence by explaining that there is a difference in ages and maybe get B back out on the court for a little practice and not play a game, but just bat the ball back and forth, to improve both of them. Maybe A should let up a little bit so that B won't feel so bad.

A has already made B feel bad by beating him so severely so now he has to make him feel good so that they will still be able to play and no hard feelings will be had by B.
CRITERIA FOR JUDGING THE "ASSEMBLY-LINE" PROBLEM

(code #) (date) (Judge's signature)

A - Categories in terms of Concrete Action Proposals.
(each solution probably fits at least one of the following numbered categories in this group. If it fits more than one, check each category that is appropriate, if it fits none, write a brief descriptive statement)

* Keeping Joe in the company, on the same work team.

1. Give the less capable less to do. Solutions which recognize differences in ability and arrange for some way of permitting the more capable member of the team to help the less capable.
   EX: Let Joe periodically change places with a faster man on his team.

2. Training Joe to make him more dexterous. EX: Give him practice in manipulative skills.

** Keeping Joe in the company, but on a different work team or job.

3. Transfer Joe to one of the other teams. EX: Transfer Joe to Station A, B, or D.
   (All of these do the same kind of work as does Joe's present team.)

4. Transfer Joe to a non-production job, in which he will be under the supervision of a foreman. (Please be especially careful in watching for this one, it is important to catch it, for the problem specifically states that this cannot be done, other foremen on non-production jobs are unwilling to accept Joe as a transfer.)

5. Transfer Joe to a non-production job, in which he will NOT be under the supervision of a foreman. EX: Make Joe a foreman, advisor, janitor, storekeeper, etc.
   NOTE: If more than one of these alternatives is proposed, make a check for each one and write next to each check the alternatives proposed.

** Removing Joe from the company.

6. Fire Joe

7. Retire Joe

For the entire above group, you may have as many checks as are appropriate.

* * * * * * *
B - Categories Concerning Centering.
(Only one of the following categories is to be checked)

1. Joe centered - solutions concerned only with the needs of Joe. ................. Joe
2. Company centered - solutions concerned only with the needs of the company. Company
3. "Both" centered - solutions concerned with the needs of Joe and the needs of the company. ..................... "Both"

*Overlapping is NOT permitted.

NOTE: If a single solution involves aspects of a "Joe" centered, or a "company" centered solution AND a "both" centered solution, mark this a "both" centered solution. If any other aspect is "centered" please explain.

C - Categories Concerning Inadequate Solutions.
(These involve what might be termed "errors" or "pitfalls")
(As many of the following categories as appear appropriate may be marked. It is possible that none will be applicable.)

1. Pigeon-holing - involves applying a label or stereotype. .........................
2. Moralizing - involves blaming or moral judgment ........................................
3. Ignoring restrictions, limitations, instructions, or adding unrealistic details. (If appropriate, please check and explain)

Explanations ________________________________________________________________

EX: One example which will automatically be noted here is the solution—"Transfer Joe to a non-production job in which he will be under the supervision of a foreman." This violates the conditions of the problem. This is just one obvious example, there may be others.

4. Using a restricted frame of reference. Involves a one-sided view of the problem. ................................................

EX: a "Joe" or "company" centered solution, or any other that involves a one-sided view, whatever it may be.

5. "Reinterpreting" facts - taking a given statement of conditions and altering it so as to permit a certain kind of solution. ................................................

EX: The problem states Joe lacks dexterity. Therefore if someone says, "Maybe Joe is better than the report shows," or "Maybe Joe purposely slowed down because he knew they were checking up on him" - and the solution makes use of this "distortion."

D - Categories Concerning the Means the Company will use to solve the problem. (Overlapping is NOT permitted. Judge in terms of the main theme.)

1. Verbal appeal - Give Joe, foreman, or group, a pep talk __________________
2. Threat or pressure - Tell Joe, foreman, or group to produce or else. ............
3. Withdrawal - refusing to deal with the problem, no action taken .................
4. Break off the relationship - fire or retire Joe, foreman, reshuffle group .......
5. Use Joe productively - in such a solution, Joe is considered as an asset to be used somewhere. It may be suggested that Joe be retrained, reassessed, tested, etc. before so doing, or it may be suggested that he be transferred without prior assessment. An attempt is made to use Joe's potential.
SAMPLES OF SOLUTIONS TO THE
"ASSEMBLY LINE PROBLEM"

S.#6, Experimental Group #3:

I think there are several possibilities to be considered in solving this problem. Maybe Joe should be considered as a foreman for either section C or D. He is congenial, has thirty years of service, he tries hard, but is slow and unhandy because he does not have the finger dexterity necessary to do this work efficiently. However, other foremen are not willing to accept Joe on transfer. Maybe we should see if he has the capabilities for a foreman. If he is energetic he might be useful in this capacity, in view of his other qualifications. However, what would be done with the foreman he replaced? Maybe somewhere up the chain there is some position that could be filled by that foreman. I think this might be considered, too. That other positions in the assembly circle be viewed to see if he could handle those. If he would not make a good foreman he might prove more efficient in another position than he is now.

S.#2, Control Group #1:

The best solution to me would seem to be in definitely removing Joe Brown from his position as number 5 man in Station C and placing him in a non-productive job, regardless of individual foremen's prejudices. If this failed to help the problem, Joe could perhaps be given a personnel job, due to his agreeable personality in the eyes of the workers, generally.
Dear ________________:

The enclosed transcripts are of a leaderless group discussion by six students. The recordings have been faithfully reproduced. If a word, phrase, or sentence does not seem to "fit," this is not due to any typographical error or error in transcribing. Laughter and pauses have been noted. The latter have been timed when they exceeded five seconds.

You will probably find it more meaningful to first read the form, "Criteria for Judging the Transcripts," then the form, "Transcript Rating Scale." If you then read the problems - there are two, and there are two transcripts per problem - it will give you a better understanding of these discussions.

It will be greatly appreciated if you will follow the sequence outlined below for the transcripts, as this has been set up according to an experimental design. The transcripts have been numbered in red for your convenience.

1st. code #__________
2nd. code #__________
3rd. code #__________
4th. code #__________

It is permissible to re-read and re-judge, should you so desire.

In these judgments, the judge is requested to take the viewpoint of an objective observer. The judge may imagine that he has been listening to the discussion, though he has not been able to see the behavioral reactions of the members.

When judging has been concluded, please place all material in this large envelope and have your Secretary telephone me that they are available. The envelope may be left __________________. Should any questions arise concerning the judging, be sure to telephone me immediately. Since these transcripts must be judged before the ____________, anything you can do to expedite the process will be appreciated. Your generous cooperation will be one of the most pleasant remembrances of my years at K. U.

Sincerely

Len Ostlund
CRITERIA FOR JUDGING THE TRANSCRIPTS *

1. **What was the degree of insight into the nature of the problem?**
   
   To what extent did the group understand the basic issues? Did they see the problem at a superficial level, or did they understand the implications of the total situation in terms of personal dynamics, such as feelings, needs, motivations, or the economic demands of the situation, etc.?

2. **Concerning the solution.**
   
   a. Was there a group solution? (From the discussion as present, was there a group solution?)
   
   b. If "yes" briefly state the group solution.
   
   c. If "yes" how adequate was the solution?

   Was the solution couched in terms of a one-sided viewpoint, such as in terms of the needs of one individual, or any limited aspect rather than the total situation?

3. **What was the degree of continuity in the development of the discussion?**

   Was there an order, continuity, or flow to the discussion so that the group proceeded with some direction, or was it a haphazard, disjointed, aimless series of statements?

4. **To what degree was the atmosphere one of warmth, friendliness, and permissiveness?**

   Group atmospheres may run the gamut from a warm, friendly, and permissive atmosphere in which the individuals tend to respond spontaneously; through a reserved, neutral or formal atmosphere; to a strained, negative, or hostile atmosphere.

5. **To what degree was the group "task-centered"?**

   A "task-centered" group directs its energies toward the task with which the group is confronted. Their immediate purpose is the facilitation of the problem solution.

* Please record all responses on the "Transcript Rating Scale."
Criteria for Judging the Transcripts.

6. To what degree was the group "group-centered"?

A "group-centered" group directs its energies toward the functioning of the group as a group. Their immediate purpose is the solution of within-group problems such as the maintenance, strengthening, and regulation of the group. For example, facilitating communication, summarizing, reconciling differences, etc.

7. To what degree was the group "individual-centered"?

An "individual-centered" group directs its energies toward the satisfaction of the participant's individual needs. Their immediate purpose is the reaching of individual goals which are neither task nor group relevant. The goals are individual in the sense that the satisfaction aimed at by the participant cannot be shared by the other members of the group.

8. To what degree was communication successful?

Communication is successful when the individual shows by his words that he has interpreted the communication as it was intended. Communication may be considered successful when members express the fact that they have understood and/or show this by their subsequent statements.

9. Some of these transcripts are of groups of students who have had case-discussion courses, others have not. This set of four (4) transcripts may involve 4, 3, or 2 of one kind of group. In your judgment state whether each transcript was a case-discussion experienced group or a non-case-discussion experienced group.
TRANSCRIPT RATING SCALE

Code # ______ Date__________________________ Judge__________________________

(Please mark with a vertical line anywhere along the continuum.)

1. What was the degree of insight into the nature of the problem?
Low 1 2 3 4 5 High

2. Concerning the solution.
   a. Was there a group solution? Yes No.
   b. If "yes", briefly state the solution:

   c. If "yes", how adequate was the solution?
Low 1 2 3 4 5 High

3. What was the degree of continuity in the development of the discussion?
Low 1 2 3 4 5 High

4. To what degree was the atmosphere one of warmth, friendliness, and permissiveness?
Low 1 2 3 4 5 High

5. To what degree was the group "task-centered"?
Low 1 2 3 4 5 High

6. To what degree was the group "group-centered"?
Low 1 2 3 4 5 High

7. To what degree was the group "individual-centered"?
Low 1 2 3 4 5 High

8. To what degree was communication successful?
Low 1 2 3 4 5 High

9. Was this a case-discussion group?__________________________

* * * * * * *
EXPERIMENTAL GROUP #5 - AFTER SESSION
(CODE # A B #5)

DISCUSSION OF "THE ASSEMBLY-LINE PROBLEM"

Experimenter: "We'll have fifteen minutes on this problem. You may begin now."

5. Well, one thing you might do is to find out what makes A, B so high in their production, and by that, perhaps, determine what's lacking in group C, at the same time, D.

1. Well, just looking here on the surface, it looks like, ah, it says, the work piles up at the position of Joe Brown and then, ah, that's for group C, and in D, ah, I don't know how much you can take from this, but--it says the foreman was dissatisfied and reprimanded them. I imagine he ate them out pretty good. It said production even fell lower. So, they didn't take too kindly to that.

2. Here's a point that you could use. Due to the fact that it said that, ah, this Joe, I believe, well, he was well-liked, he was a nice fellow, and since that, ah, D is pretty low anyway, due to the fact, possibly, that the foreman has reprimanded this group, is to transfer ah, Joe from the C group to the D group, whereas it might raise the D group somewhat and then get--add--somebody else to the group where he was replaced and it might speed up production there. All along the line.

1. Ah, you think then that by combining this congeniality of Joe, maybe moving him into a more administrative position, say, in D here, to smooth down the ruffled feelings of the people there, and at the same time move him out of C, where he is a bottleneck?

2. Yeh.

6. Well, there is also another thing to look at. He is 60 years old and it says in here
6. (continued) that ah, finger dexterity is a very important factor. Well, when you get to be 60 years old you don't have quite the dexterity that you used to.

4. Might...

2. True.

6. (continued) and it is possible that he can't do it any more.

4. It might be. It might help, too. He's C's third man down here. Maybe if they moved him up to first man he wouldn't have as much to lift. It wouldn't be quite as heavy because it wouldn't have so many parts. I wouldn't know if that would have any--of course, it says that each job takes about the same amount of time.

6. Well.

5. If he can't do the job for group C, I don't (2 min.) see how he can do much better in group D.

6. I don't either.

2. The only thing is, it might improve the relationships of the others in that group who've been reprimanded.

5. It might do that.

2. That's what I was referring to.

5. It might help the group as a whole even though his own production wouldn't...(Interruption by No. 2)

2. Wouldn't be any better.

6. But even if it did help the group as a whole it would speed up theirs and there he would be, to slow it back down again.

2. But maybe he wouldn't slow down as much accordingly--that could-- (unclear)
Ah, this idea that I mentioned a little bit before about putting him in an administrative position, ah. I don't know what the foremen do on this job, but I would assume that they're not in this, ah, where they set up the station. In that way he wouldn't be slowing anything down. You could shift him from that position with D.

5 and 3 talk at once...(unclear)

5..........(to No. 3) Go ahead.

3..........As long as group D is, is producing at the rate of 50 units per shift, anyway, ah, and he's been the bottleneck in group C, and (3 min.) they're still being—maintaining—production rate of approximately 60 units per day. Surely his pace will be more like that of the D group so that perhaps he will fit in there just as well, and maybe overall production will be raised a few points by his transfer to the slower group.

6..........You'd raise it back to 60 possibly, but that still leaves us with a shortage...(Interruption by No. 5)

5..........(unclear)

6..........(continued) deficiency.

3..........Well, that might—that might free group C, however, to increase production.

2..........I believe...

6..........Possibly.

2..........(continued) I believe that it—that No. 1 mentioned that—put him in as, say, sort of a foreman of D and since he is liked by the men that it would help bring up production, yet he wouldn't be doing the actual assembling—the assembling himself, so he wouldn't be interfering with production, yet he would have a job, and at the same time help improve production in that group.

6..........If the foreman does that work.

2..........Yes if?
6. We don't know for sure.

2. That's right. We'd have to base it on that assumption.

5. Well, to formulate that solution you have to get the problem, and it looks to me like Joe is certainly the problem here, and I take a capitalistic view of it. If he's not producing, we have to do something with him, ah, to get the production up. It looks like the production is your problem. You've got to have 300 units, and even though he's a nice guy, and all that, if he can't produce, why he's got to move, got to go.

6. Ah, surely, if it's any company that has mass production like they, must have some sort of a pension plan.

5. Yeh.

1. I believe he probably has some years to go yet.

6. Thirty years? They usually....

4. They usually can...(unclear)...

6. They usually...(unclear)... all talk at once

3. Well, retirement's--I think we all think retirement is in order if he is at retirement age and has a pension coming to him, or, or... adequate compensation, at all.

6. But I would think 30 years...(Interuption by No. 5)

5. (unclear)

6. (continued) might possibly be retirement...

(unclear) (Interuption by No. 5)

5. He must have earned it or they wouldn't have kept him that long.

6. Yes.
5. (continued) Must have been a good employee. Maybe he just can't get the hang of the mass production methods.

6. Mm, hm.

5. (continued) on this particular job.

6. Well, chances are he is just slowing down and that's just the thing.

3. Yes.

6. (continued) I mean, those little things. I (5 min.) know if it was a carburetor like I think they mentioned in here, why it would be—I don't know. I'm pretty rough on that myself.

5. Well, you and I might have a rough time putting a carburetor together..(Interruption by No. 2)

2. It takes the right kind of fingers. With hands like we've got...(unclear)

6. Well, sure...(talk )

5. Sure......... (together)

4. It might be. We could move him to one of these positions where he supplied the...(Interruption by No. 5)

5. Well, it says non-production doesn't want him... (Interruption by No. 3)

3. You'd, you'd take him out of the group then, actually. He'd be out of this--he'd be out of this ring.

4. Yeh.

3. This certain position.

6. Well...

4. (continued) It does say they don't want him.

3. He'd be supplying...
Foremen on non-production job are not willing to accept the transfer.

They don't want him anywhere.

That's the best....

Pretty good idea, because you can't just let him go, I mean, it looks to me like you owe Joe something...(Interruption by No. 2)

He's been there for 30 years. You owe him something, definitely.

Now, surely, if--it says that they are not willing to accept a transfer, but it sure, it would be possible somehow, to work a transfer where he wouldn't have to have such finger dexterity if that is possibly the cause of trouble. I mean, if there is something which is a little bit larger and doesn't take quite that, and still it's a job, and due to the 30 years I doubt very much if they would cut him in pay.

Yes.

(continued) At least they would be fair to him.

I can't really see why they wouldn't want him if he's congenial and...(unclear) (Interruption by No. 5)

That's what I couldn't--I couldn't understand why non-production wouldn't want him on their job.

'Cause I should think that would be the thing they'd work for, these men that can be--get along with the others, see....

It seems that they're basing the fact that they don't want him on the fact that he has been a bottleneck in this one thing, where he might not be somewhere else.

Such things as personnel work or something.

That's right. Personnel management would come in there, or placement, er.....

(Pause, 16 seconds)
3. Well, we'd find Joe willing to do anything if we could think of something for him to do. He'll be happy to do it, 'cause....

1. He tries.

5. He has to work for a living, and....

3. It seems to be a little bit uncertain about him here, but he was a nice guy to get along with.

6. Well, that foreman's job, it's usually as it says--it's a supervisory position, and if he's been there for 30 years, he knows what he's doing. I mean, he knows what's required,

(7 min.)

1. Yes.

6. (continued) and I think he would make a good supervisor.

1. Yeah, I think they could get the most out of him by keeping him in, in this administrative position like you said, 6, and still he'd be fair enough that he could smooth the ruffled feathers on our little birds in number--station D.

6. If, ah, he's a friendly type of person, he's very congenial, and besides that, he's been there for 30 years, and he should know very definitely exactly what's wanted. He knows the business. He knows what, ah...(pause) ah, if something should happen, he probably would know what to do. (pause)

3. The problem with moving him to an administrative position is whether or not he can efficiently supplant the men that are already there. They might be serving very well.

5. This probably wouldn't do any good in Joe's particular case because if a man can't do it, he can't do it, but on the other shifts in order to get your production up, which seems quite a problem here--ah, pay incentive plan might be worked out. It looks like 75 units per sub-assembly station is what they'd need to make the 300...
6. That's right.

5. (continued) and the two groups are exceeding that, and two are pretty low.

(Pause, 7 seconds)

4. If you could get the one that was previously doing 60 up to 85 it would take care of Joe's there with 60.

6. Well, if you could get them both up to 65 it would take care of the situation.

5. Yeh. Of course, then you would have two horses pulling part of the load of the other two and...

1. Yeh.

5. (continued) and it's a bad thing there.

1. Yeh.

5. (continued) but, ah, it would be better than not having the production...

6. Mm, hm.

5. (continued) required 300 units, I'll grant you that.

6. One up to 65 and the other one up to 70, maybe.

5. It says that pay is based on hourly rate. I don't know whether they've ever considered the incentive plan or not. They sometimes can back-

(9 min.) fire, too.

(Pause, 14 seconds)

5. Reading this over again I see that production is so low that the factory production as a whole had to slow down.

6. Mm, hm.

5. Why that's... ( Interruption by No. 2)

2. So it means something has to be done.
5...........Yeh. It has to be done no matter how drastic it is. They would probably come out ahead of it if they could retire Joe, if that's the trouble. Of course, that doesn't take care of group D...

2...........No.

5...........(continued) but if they'd retire Joe they might come out ahead, 'cause they must have other people idle if they can't supply enough of these parts that are going to a major assembly division some place else.

(Pause, 8 seconds)

2...........It would look like, ah, since Joe's a problem and also D's a problem if you could combine them together. In other words if you could make use of Joe in the group D, such as a supervisory position, if he's able to handle it, it would be best because it would speed up D and it would solve the bottleneck in C, too. (10 min.)

5...........You might have an idea there...(pause) if that would do.....

2...........If that would do it.

5...........Mm, hm.

1...........I'd have to agree there, with that, of course. That's been my sentiment from the beginning of this discussion.

2...........I think, that's, in general, been all our sentiments.

5...........Mm, hm. You'd at least localize your problem.

2...........Yeh.

5...........(continued) in one group that way.

2...........Yeh, I know.

5...........(continued) if nothing else, and C would probably move up--at least they're blaming it on Joe. Of course if C stayed the same then you would know it was something else....(pause)
1. This calls for some definite policy or action. We might as well, since this is the most obvious, try it first.

2. Mm, hm.

3. I think a simple transfer to another station—D, namely—will probably result in exposing a few more facts.

4. Well it could be done with a minimum of difficulty just ...(Interruption by No. 3)

5. Why of course. (Nos. 5 and 3 talk at once)

6. (continued) ...be a matter of going to work someplace...(Interruption by No. 3)

3. You could put him in exactly in the same relative position.

5. (continued) else, same job and everything.

4. I wonder if you could put him in, maybe A, where they are making 85, and cut that down and bring up the other one. Of course that would be hard on morale.

3. Well, I doubt—I doubt....

2. (unclear)...

3. (continued) I imagine that he'd knock any group that he was with down to 60 units.

2. You'd still have your D bottleneck at the same time,

5. Yeh.

2. (continued) and if we could try to bring all the problems together at one point we'd be better off than having several little things that we...(Interruption by No. 5)

5. Even if we had three groups that were pretty efficient, it would be better than just having two groups.

(Laughter)

2. Yes.... (several talking at once)

Somebody... two that work....
5. (continued) Four would be ideal, but if you can't have that why then you should try and get three....(pause)....They say perfection is a matter of relativity, well...(laughter) I don't know. (pause)

3. I wish we could get Joe somewhere--somewhere that he's....

6. I just--I just have a feeling that he's just simply unable to do any better, physically.

Somebody...Nm, hm.

1, 2, 5, & 6 talk at once)1...I think he's doing as well as he can.

2. Yeh, and retirement is the only thing.

5. I'm inclined to agree with you.

6. Uh, huh, because, ah, it says that the job has been analyzed by time and motion experts so that the positions are of equal difficulty. Now, everybody else is doing a fine job, but it bottles up there and, and that gives me the idea that this is pretty well in Joe. Probably Joe is just not physically able to do any better for one reason or another, and chances are it's that finger dexterity with age.

4. Yes.

3. Would you--would you recommend then that we, that we'd think the best thing to do is to transfer him to D, then see what happens?

1. Yeh, try that first.

3. And, and, ah, failing in that you'd have to try something else, ah, even going as far as retirement?

5. I'd agree with that and I would also say I'm afraid that he's doomed to failure in group D, too. I agree with No. 6 here. I think that Joe is just, just unable physically to--to perform work.
Physically.

If it were a matter of education why they could certainly teach him, but after 30 years, the man must know more about it than I do.

Sure.

Yes.

Well, one thing that might, ah, add or detract from that is that it states that each one of these sub-assembly stations are supervised by foremen. Does the foreman enter into the work or is it purely oversees it?

(unclear)

That might make a difference, too, cause if he enters into the work—-I would agree with you that definitely, it looks to me like he won't be able to do it, but if he merely oversees, with that 30 years of experience, it would seem to me that he would, maybe, do very well at that. But that's the question.

Well, I have a suggestion. Here we have—so far we've looked at the facts and it looks to me like we've decided he won't be able to work. Why don't we look through and see if we can't find a position where he would be doing supervisory work. He knows what it is, he knows all the facts about it. Let's look through the organization and see if there can't be a, something like that found.

If the foreman in group D doesn't actually do work and I think that's a good place for him.

Yes, of course. He can fit—he can fit with a 50 unit production rate, of course. He could fit in a 50 unit production...(Interruption by Nos. 1 and 5)

But the idea is to keep it there, yeh...
5. If he's non-production why he might... (Interruption by No. 2)

2. Might even drop it further.

6. I would suggest though, that he be possibly put into a strictly superintendent type of thing...(Interruption by No. 5)

5. Parts man or supply....

6. (continued) He knows his business. We know that, and we also know that he's slowing down, so consequently, the two of them--it looks to me--would fit together into a good supervisor of some kind.

4. He's well liked.

6. (continued) He's well liked, he knows his business... (Interruption by No. 5)

5. I don't see why...

6. he's too slow for anything else.

5. (continued) he wouldn't make a good supervisor.

2. I think so too. If there's one to be found.

6. Otherwise if we have a pension plan, oh, it (16 min.) looks like that's the best.

* * * * *

Experimenter: (ending the discussion)
"It looks like time ran out on us.
That was a fine discussion."

* * * *
BIBLIOGRAPHY


