A SUPERVISORY HUMAN RELATIONS TRAINING EVALUATION PROJECT: MEASURING ATTITUDE CHANGES IN SUPERVISORS AND IN THEIR SUPERIORS AND SUBORDINATES

by

Scott Robert Purtell

A Thesis Submitted to the Faculty of the

DEPARTMENT OF MANAGEMENT

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 7 8

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Scott R. Turbell

APPROVAL BY THESIS COMMITTEE

This thesis has been approved on the date(s) shown below:

Edwin B. Flippo — June 6, 1978

EDWIN B. FLIPPO, Thesis Director
Professor of Management — Date

Committee:

John W. Dickson — June 6, 1978

David A. Tansik — June 6, 1978

Robert M. Hollmann — June 7, 1978

ACKNOWLEDGMENTS

I would like to express my deep appreciation to Dr. Edwin B. Flippo, my thesis director, for his help and guidance throughout the duration of this project. I am also indebted to Drs. Robert W. Hollmann and John Dickson, members of the thesis committee, for their invaluable contributions to my research.

Many thanks are also extended to the City of Tucson and the Division of Training for making this study possible. Special thanks are extended to Conrad Cormier, Helen Cigrand, Gail Rothrock, Deborah Tanner, and Phil Carlin for their ideas, constructive criticisms, patience, understanding and support. I would also like to thank all of those City of Tucson employees who gave me their time and cooperation so that the data for this project could be collected.

A final expression of gratitude and sincere thanks goes to a very special friend, Ronald W. Decker, for reasons he knows well. However, the ultimate expression of gratitude goes to my wife, Gail. Her faith and understanding during my undergraduate and graduate years at The University of Arizona has meant more to me than I could ever express here.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

LIST OF TABLES

ABSTRACT

The objective of this thesis is to evaluate whether a supervisory human relations training program is able to change attitudes, and hopefully, behavior. Attitudes are generalized responses and can be used to predict behavior in a given particular situation. More specifically, the purpose is to measure any changes in the attitudes of a group of first-line supervisors as a direct result of having participated in the training program. Moreover, the thesis will also measure the attitudes of the superiors and subordinates in relation to their perceptions of their supervisors' behavior as a result of the training program.

The ultimate test of this training evaluation project, or any other such project, is whether or not any positive changes in attitude are reflected in future changes in organizational behavior. However, analysis of the data collected in this study indicates that no significant differences were observed between the post-training scores and the pre-training scores for any of the groups. Neither changes in attitudes or changes in behavior were observed as a direct result of the training program. The specific results obtained from this study do not preclude its contributions to the overall field of training evaluation research.

CHAPTER I

INTRODUCTION

State of the Art

This study deals with the evaluation of a supervisory human relations training program. The concept of executive development has undergone a management "revolution" over the past 30 to 35 years. A great increase in the number and variety of development programs has been born out of this revolution. Both public enterprise and private industry are fully involved in executive development programs. However, the art and practice of measuring the effectiveness of managerial training falls far short of the desirable. Management development programs in general have proved to be somewhat of a disappointment.

According to Robert J. House (1965, p. 16):

> Management development is explained as a process of change resulting from imparting information plus selected reinforcements which are unique to the managerial environment. These reinforcements influence the degree to which learning takes place and the degree to which learning can be translated into improvement in attitude and job performance.

Consideration of Dr. House's statement provides a basis for the logical assumption that a properly designed training program can be an extremely effective vehicle through which

management can realize its goals. However, in order for training to function as an effective tool for management, it (training programs) must continually undergo systematic evaluation in some form or other. A much more detailed consideration of the state-of-the-art in measuring the effectiveness of managerial training is presented in the following chapter.

Objectives and Approach

The basic objective of this study is to propose and test one particular research model that hopefully will improve the process of evaluating managerial training. Many managerial training programs concentrate upon affecting attitudes, which in turn should impact behavior. Consequently, any comprehensive evaluation model should attempt to measure both trainee attitude changes and trainee behavior changes during a subsequent period. The model proposed in this study will incorporate both types of measurements.

The general approach used in testing the model of training evaluation is through a case study. The setting for this study involves a supervisory human relations training program developed by the Training Division of the City of Tucson. The training, which is entitled the "Leadership Program," attempts to teach first-line supervisory personnel techniques of human relations and supervisory management.

Accompanying this training process is an attempt to influence trainee attitudes in the direction of more "considerate" behavior toward subordinates along with a greater willingness to elicit subordinate participation in managerial decisions. Attention is also given to affecting trainee attitudes toward goal attainment, interpersonal relations, and role-structuring.

Attitudes are important, but according to Blumenfeld and Holland (1971, p. 638), "The purpose of training evaluation is to determine if desired changes in behavior (emphasis supplied) did occur." An effective training evaluation model should therefore incorporate multiple measurements at different stages of a program. In this case study, the impact of the supervisory human relations training program will be measured in two ways. First, the management trainees' attitudes will be evaluated both before and after the training. Second, assessments provided by the management trainees' superiors and subordinates are obtained to improve the accuracy of conclusions concerning training effectiveness. Thus, changes in attitude will be identified through measurement of the trainees. Changes in behavior will be identified through reports of superiors and subordinates.

Importance of the Study

This study is important in terms of its contribution to the development of training evaluation research. Continued and meaningful research is a necessity in the area of the measurement and evaluation of management training programs.

The study has value because it attempts to evaluate a management training program in a governmental setting. From this evaluation comes further information and knowledge on the state of training evaluation, the value of training, and the overall efficiency and effectiveness of a given training program. From this study opinions can be formulated, assumptions can be made, and conclusions can be drawn as to the needs for training, the further organizational uses of training, the effects of training, the results of training, and the effectiveness of training.

Finally, in a survey of the results of 400 experimental studies concerned with management development, Robert J. House (1965) found that in many instances the programs had little or no measurable effect on business performance or managerial behavior. Improved models for evaluating training impact could well provide firms with clues concerning how these training programs could be designed to produce more of their intended values.

Limitations of the Study

Several limitations are present as a result of selecting the particular research setting for this study. These limitations may or may not have had a significant effect on the outcome of this project. The major limiting factor is that this is a single case study and the ability to generalize its findings is restricted.

Three additional limiting factors are associated specifically with the research setting itself. First, there is the climate and atmosphere associated with the work environment of the test groups. The City of Tucson governmental organization (hereafter referred to as the City of Tucson) is presently characterized by much tension and conflict. This atmosphere is created largely from factors such as impending work force reductions, budget cuts, and council elections. Politics plays a major role in forming and shaping the state of the environment within the City. Second, is the sample size of the test groups. The experimental group of supervisors is composed of only 17 test subjects, while the control group is made up of only 13 individuals. Likewise, only 14 test subjects comprise the experimental group of superiors, with the control group having a sample size of eleven. The final additional limiting factor is the experimental mortality associated with one of the test groups. Between the pre-measurement phase

and post-measurement phase of the study, the experimental group of supervisors lost three members. These three losses occurred as the result of accidents and illness.

Outline of the Study

This study is divided into five major sections, each section being one chapter. Following these introductory remarks, Chapter II deals with the Leadership Program as a case example and some extensive background information on training evaluation. The training program is detailed as to the personnel involved, goals and objectives, structure, and methods of instruction. A survey of the literature concerning training evaluation is also provided.

Chapter III presents the design of the study, the hypotheses, the test subjects, and the instrumentation utilized. Chapter IV contains the results of the study, in conjunction with a detailed consideration of the statistical methodology employed, and concludes with a brief analysis of the research findings. The final chapter consists of a summary, specific recommendations, and a discussion of the study's implications for future research on training evaluation.

CHAPTER II

BACKGROUND AND SETTING FOR THE STUDY

Review of the Literature

Many research projects have been conducted in the area of training evaluation. For the most part, the results of much of this research have been vastly inconsistent. A number of these studies have focused on supervisory human relations training programs using both the Leadership Opinion Questionnaire (LOQ) and the Supervisory Behavior Description (SBD), developed by Dr. Edwin A. Fleishman, and are therefore relevant to this study. Both the LOQ and SBD are designed to measure two significant dimensions of supervisory leadership, Initiating Structure and Consideration. One of the many uses of these instruments is in training evaluation and, more specifically, in evaluating the impact of management development programs. Chapter III contains a section on instrumentation which provides a complete description of both the LOQ and SBD.

Although numerous training evaluation projects have been conducted under a multitude of different research settings, the overall results have been somewhat disappointing. Furthermore, many of these studies could not retain their

claims of significance under close critical analysis. According to Bunker and Cohen (1977, p. 526), " . . . this trend toward continued avoidance of the evaluation issue comes at a time when measurement of training impact would appear to be increasingly more important. The demand for additional and more complex training programs is likely to continue to escalate."

With this general overview of the state of training evaluation research in mind, a number of specific studies will be examined. These studies will be categorized and presented in terms of their major research findings--positive, negative, or mixed.

Positive Results

In an early project, Raymond A. Katzell (1948) studied a supervisory human relations training program involving 60 management individuals of the Illinois Central Railroad. The eight-week program was divided into three units: the Supervisor and His Job, the Supervisor and Human Nature, and the Supervisor and Leadership. The objective of the training program was to increase effectiveness of supervision and the ability to deal with peers and superiors. Both experimental and control groups were utilized. The File-Remmers How Supervise questionnaire was administered to both groups prior to training and at the completion of training. The results indicated a significant change in

the post-training scores as compared with the pre-training scores for the experimental group.

In 1954 Francis J. DiVesta evaluated a training program involving 94 medical administrative supervisors in a military school. Utilizing both experimental and control groups on a pre-training and post-training measurement basis, the LOQ and 13 other instruments were used in an attempt to assess the effects of the training program. With respect to the LOQ, the results showed an increase in Consideration scores and a decrease in Structure scores for the group exposed to training.

A few years later, the Bell Telephone Company was the object of a training evaluation project (Stroud, 1959). At Bell, 122 supervisors and their superiors were measured. The purpose of the study was to assess the value of the training program by comparing expected behavior with actual behavior. Using the SBD, data were collected on present behavior and on behavior that had occurred one year prior to training. The results indicated that the SBD was an effective measure of performance changes in areas covered by the training program.

A study principally utilizing the LOQ (Biggs, Huneryager, and Delaney, 1966) measured the effects of a human relations training program. The test subjects were 32 potential supervisors attending the Creighton University

Conference. Both pre-training and post-training measures were recorded. The results showed an increase in Consideration scores and a decrease in Structure scores at the post-training measurement.

A study concerned with the relationships between the leader behavior of industrial supervisors and the behavior of their work group members (Fleishman and Harris, 1962) produced one significant outcome; the results documented the fact that there are a myriad of interactive effects between different combinations of Consideration and Structure. This study examined 57 production foremen and their work groups in a motor truck manufacturing plant using the SBD as the measurement tool.

One final study producing positive results (Hand and Slocum, 1972) evaluated a managerial human relations training program in terms of the program's ability to change attitudes. Forty-two line and staff managers, and their subordinates, in a specialty steel plant were studied using both the LOQ and SBD. Both experimental and control groups were measured prior to training and 18 months after training. The results indicated that changes in attitudes were accompanied by performance changes for the experimental group. In other words, what was learned in training was transferred to actual on-the-job performance. This particular study was a follow-up of previous work (Hand and

Slocum, 1970) wherein the indicated results were exactly the opposite of those just reported. The 1970 project is contained in the following section dealing with studies resulting in negative findings.

Negative Results

While a number of training evaluation projects have produced positive and significant results, many more studies have not. Much of the training evaluation research conducted over the past 25 years has caused the rejection of many research hypotheses due to the generation of negative end results. The most comprehensive study confirming the overall unhealthy state-of-the-art of training evaluation was conducted by Robert J. House in the early 1960s at the University of Michigan.

In the most extensive research yet conducted into the state of training evaluation, Dr. House (1965, p. 15) " . . . analyzed the results of 400 experimental studies concerned with management development and found much disappointment and disillusionment." This project covered both private industry and public enterprise on a nationwide scale and even included some work in England. A vast number of different research settings were studied, ranging from light industry to heavy industry and from laboratory settings to actual field research. Many university development programs and various other conferences and seminars

were included in the research along with several government and civil service programs.

Many different types of training programs, using different instruments designed to measure and evaluate management development, were studied. These programs focused on human relations, managerial performance, measuring leadership, managerial effectiveness, and leadership status. Many of the training programs were also concerned with measuring attitude changes, cohesiveness, the selection process (for the training program), and overall management development and supervisory training.

According to Dr. House's analysis of these 400 training evaluation studies, "In many cases, these programs have had little or no demonstrable or measurable effect on business performance or manager behavior. Several management development programs have actually caused problems and undesirable behavior (1965, p. 15)." For example, many of these studies resulted in increased conflict and tension between the supervisor and his superior, less overall satisfaction with the company, and even job terminations as a direct result of the training program.

In a study utilizing both the LOQ and SBD, Fleishman (1953) investigated the relationships between how a foreman leads his work group and the attitudes and behavior of those above him in the organization. He studied 100 foremen from

17 International Harvester manufacturing plants on a pre- and post-training basis. The results indicated that Consideration scores increased and Structure scores decreased as a result of the training. However, subsequent behavior exhibited on the job was somewhat lower in Consideration and higher in Structure. More importantly, it was found that the type of superior under which the foreman operated affected the (foreman's) leadership of his work group more than did their having participated in the training program. In other words, a superior high in Consideration would tend to have foremen high in Consideration who, in turn, would be perceived high in Consideration by the work group. This same basic relationship was also found for the dimension of Structure.

In a follow-up study (Fleishman, Harris, and Burtt, 1955) the same International Harvester motor truck manufacturing plants were used. In this extension of the 1953 research, 122 foremen were studied along with 60 of their superiors and 394 of their subordinates. The foremen involved in this two-week leadership training program were measured prior to training, at the conclusion of training, and one year after training using the LOQ. The superiors and subordinates were likewise measured using the SBD. The results showed a slight increase in Consideration scores and reconfirmed the independence of these two scales.

However, the training program did not make the foremen more homogeneous with respect to their leadership attitudes.

Mathew B. Miles conducted another training evaluation project (1965) in the area of human relations. He used pre-training, post-training, and post-post training (eight months) measures to detect any changes in 34 school principals as a result of a two-week human relations training program. Both experimental and control groups were measured using the SBD and several other perceived-change on the part of the program participants as measured by the SBD.

Hand and Slocum (1970) conducted a field experiment in a specialty steel plant concerning a management development program designed to improve human relations practices. Forty-two managers from the same hierarchical level participated in one and one-half hour training sessions per week for 28 consecutive weeks. The research design employed both experimental and control groups measured on a pre-, post-, and post-post training (three months) basis. The independent variable under study was the longitudinal effect of the training program, while the dependent variables were knowledge, attitudes, and behavior. The instruments used to evaluate the program were Kirkpatrick's "Inventory of Human Relations," Berger's "Acceptance of Self and Others

Scale," Likert's "Profile of Organizational Characteristics," and the LOQ and SBD.

The results of this experiment revealed the training to be ineffective. No significant changes of any kind were detected in attitudes or behavior at the post-post measurement taken three months after conclusion of the training. According to Hand and Slocum (1970, p. 407), "It may be conjectured that the 'environment' of the organization under study was perceived by the participants to be unable to support their change in attitude; thus, this created strong pressure to return to the previous attitudinal state."

In a final study that produced negative end results, Elkins (1977) evaluated a basic management training program given to 110 government supervisors and managers in Contra Costa County, California. This one-week program covered communication, leadership styles, organizational climate, motivation, performance evaluation, and objectives setting. Both experimental and control groups were used in an attempt to determine the impact of the training on the participants' behavior after they returned to work. The results indicated that, as a result of the training program, the participants were less confident of their ability to perform their jobs at the post-training evaluation, were somewhat more knowledgeable regarding basic management, and exhibited

no significant changes in goal accomplishment back on the job.

Mixed Results

Some training evaluation research studies have resulted in simultaneously positive and negative findings. Other studies have concluded that neither positive or negative results occurred; that, in effect, there was no measurable change of any kind detected as a result of training. Finally, some studies have produced inconclusive findings resulting in an inability to draw definite conclusions about the impact of the training program. Two such studies will now be examined.

Theodore J. Carron (1964) reports on the evaluation of a human relations training program for supervisors in terms of changed attitudes. He studied 23 male supervisors in research, development, and engineering departments of a chemical company. Utilizing the LOQ, both experimental and control groups were measured prior to training, at the conclusion of training, six months later, and 17 months later. The results from Carron's study were mixed, both positive and negative. Positive changes in the supervisors' attitudes did occur over the training period. However, these changes were not permanent as measured at the two post-post evaluations.

In another study, Bunker and Cohen (1977) evaluated a development program at a large independent telephone company. The program involved training 131 male supervisors in an intense one-week program. Bunker and Cohen employed both experimental and control groups in a seven-week evaluation period consisting of pre-training measurements and post-training measurements. In the end it became " . . . apparent that different conclusions could have been reached relative to the effectiveness of the training program . . . (p. 537)."

A summary of the training evaluation research presented thus far is contained in Table 1.

What do we know about the state-of-the-art of training evaluation? Most important is the realization and acceptance of the fact that it is still an art and not an exact science suited to the making of specific statements. However, "In the past two decades, personnel practitioners have learned a lot about training, learning, and management development (Elkins, 1977, p. 305)." From this increase in training knowledge must come the realization on the part of business that it can no longer afford to forgo or conserve effort and expense in the area of training evaluation.

There is some evidence that suggests business is finally understanding the importance of evaluating management development programs. A Bureau of National Affairs

Table 1. Summary table of training evaluation research.

Researchers	Location and Subject	Sample Size	Instruments Used	Major Findings
Katzell (1948)	Ill. RR, managers	60	How Supervise	positive changes in effectiveness of supervision
DiVesta (1954)	military school, medical administrative supervisors	94	LOQ and 13 other instruments	increase in Consideration, decrease in Structure
Stroud (1959)	Bell Telephone, supervisors and superiors	122	SBD	SBD effective measure of performance changes
Biggs, Huneryager, Delaney (1966)	Creighton University, potential supervisors	32	LOQ	increase in Consideration, decrease in Structure
Fleishman, Harris (1962)	truck manufacturing plant, supervisors	57	SBD	interactive effects between different combinations of C and S
Hand, Slocum (1972)	steel plant, managers and subordinates	42	LOQ, SBD	changes in attitudes accompanied by changes in performance

Table 1--Continued

Researchers	Location and Subjects	Sample Size	Instruments Used	Major Findings
House (1965)	many different	400	LOQ, SBD, and many others	state of training evaluation is disappointing and disillusioning
Fleishman (1953)	truck manufacturing plants, foremen	100	LOQ, SBD	actual job performance lower in C and higher in S, superior atmosphere more relevant than training program
Fleishman, Harris, Burtt (1955)	truck manufacturing plant, foreman, superiors, subordinates	122, 60, 394	LOQ, SBD	slight increase in Consideration, no improvement in leadership attitudes
Hand, Slocum (1970)	steel plant, managers	42	Inventory of Human Relations, Acceptance of Self and Others Scale, Profile of Organizational Characteristics, LOQ, and SBD	training ineffective, no change in attitudes or behavior, environment unable to support training concepts

Table 1--Continued

Researchers	Location and Subjects	Sample Size	Instruments Used	Major Findings
Elkins (1977)	Contra Costa County, CA, government supervisors and managers	110	various perceived-change tools	reduced confidence in ability to perform jobs, no changes in goal accomplishment
Carron (1964)	chemical company, technical supervisors	23	LOQ	positive changes in attitudes but did not last over time
Bunker, Cohen (1977)	telephone company, male supervisors	131		inconclusive

study (Nixon, 1973, pp. 13-14) surveyed over 300 business organizations. These organizations represented an extensive cross-section of types, sizes, and locations of various business firms. The survey reported that 70 percent of the companies were making a concerted and serious effort to evaluate their development programs. In another study by Catalanello and Kirkpatrick (1968), 154 business firms were surveyed. These companies were, again, involved in a wide spectrum of different business activities. The results indicated that 78 percent of the firms were attempting to measure the impact of training, with over half attempting an evaluation of attitudes and/or behavior.

From knowledge of the state-of-the-art, we can ask what is wrong with training evaluation research? First of all, much of the research in this area is based on incorrect assumptions, unclear thinking, faulty reasoning, and improperly structured research designs and analysis. The number of really good training evaluation studies is pathetically small in comparison to the total number of research studies undertaken. From Dr. House (1967, p. 31) comes the firm assertion that, "In too many cases, management development has failed to produce measurable results." One final point remains to be mentioned. In his 1965 study of 400 training evaluation projects, Dr. House labeled the results as disappointing and disillusioning and went on to

say that, "Only a few rigorous evaluations of management development have been reported . . . (1967, p. 16)." It is most disturbing to find these results confirmed twelve years later by Bunker and Cohen (1977). In their research, Bunker and Cohen found training evaluation to be ". . . one of the most under-researched and neglected areas of industrial/organizational psychology (p. 525)."

Personnel

The training program, which is officially entitled the "Leadership Program," is conducted by the Division of Training of the City of Tucson. Conrad "Connie" R. Cormier heads the Division of Training and administers the Leadership Program. Mr. Cormier has been with the City of Tucson for eight years and currently holds the position of Employee Development Officer. His education background includes a BA in Government in 1953 and a 1969 MA in Political Science, both from the University of New Hampshire. He has also done graduate work at Northern Arizona University, The University of Arizona, the University of Southern California, and the University of Utah. Mr. Cormier is a Facilitator for Management Instruction and a member of the faculty at Pima Community College in Tucson.

From 1953 to 1967 Mr. Cormier was employed in various management positions with the U.S. Department of Defense

in the intelligence field. Mr. Cormier has international experience in his field, having served in Puerto Rico and Turkey as well as Hawaii and Washington, D.C. He is a current member and past president of the American Society for Training and Development (ASTD).

Helen R. Cigrand is the Assistant Employee Development Officer for the City of Tucson. Mrs. Cigrand has been with the City for five years and has been associated with the Leadership Program since 1975. She received her BA in 1969 from Loras College in Dubuque, Iowa, and is currently working toward a MBA from The University of Arizona. Mrs. Cigrand is also a faculty member at Pima Community College.

Prior to coming to Tucson, Mrs. Cigrand was a personnel assistant at a large, midwestern utility firm. Her areas of competency include benefits, classification, labor relations, procurement, selection, and wage and salary administration. Mrs. Cigrand is also a member of ASTD.

Two other individuals are principally involved in the Leadership Program: Phyllis Siken and Glenn Trowbridge. Mrs. Siken holds a master's degree in Counseling and Guidance from The University of Arizona in 1977. She is a management consultant and president of her own consulting firm, the Human Relations Institute. Mrs. Siken has been associated with the Leadership Program through a contractual arrangement with the City of Tucson since 1975.

Glenn Trowbridge is the Assistant Personnel Director for the City of Tucson. His educational background includes a 1969 BA in psychology and an MA in 1974, both from San Diego State College. He is the former Assistant City Manager for National City, California. Mr. Trowbridge has been associated with the Leadership Program for one year.

Goals and Objectives

The goals and objectives of the Leadership Program are to improve the supervisory management techniques of first-line supervisory personnel of the City of Tucson. These goals and objectives will be accomplished by improving the participants' knowledge and understanding of supervisory and human relations techniques. Therefore, the training program is designed to " . . . update the experienced leader and to provide needed knowledge and skills to the new leader," according to City Manager Joel Valdez (1977, p. 1).

The ultimate end result of these goals and objectives is to evolve a City structure built on trust, openness, problem-solving, and an awareness of the myriad of interpersonal reactions and interactions taking place within the organization. Specifically, the announced goals of the Leadership Program, which were revised and updated by

Conrad Cormier and Helen Cigrand in the Spring, 1977, are the following:

1. Have an understanding and knowledge of supervisory techniques and styles.
2. Know the process of recruitment and selection of employees, classification and compensation of positions, and grievance handling.
3. Produce a productivity plan in groups and receive a course grade for the plan.
4. Be aware of City benefits, OSHA requirements, job enlargement and enrichment methods, and employee development.
5. Be able to perform problem-solving and conduct performance appraisals.
6. Gain an awareness of human relations techniques and motivation applications.
7. Be aware of the necessity of integrating the interests and needs of individual employees with the goals and objectives of the City.

Structure and Content

The Leadership Program was started in 1971, the result of a directive issued by the then-Director of Personnel of the City of Tucson. The 1971 program consisted of 15 hours of instruction. Each year the program was reviewed, modified, and expanded by Conrad Cormier. In 1975 the program consisted of 30 hours of training, in 1976 the program was 40 hours, and in 1977 the training program was expanded into its present 70-hour format.

The Leadership Program is offered each fall and each spring. Three classes are offered each "semester."

The class meets twice a week for approximately eleven weeks and each class is two and one-half hours long. In the Fall of 1977 Session I began on 16 August and ended on 1 November, Session II began on 6 September and ended on 22 November, and Session III began on 27 September and ended on 15 December. The Session III class served as the experimental group for this project.

Currently the Leadership Program is composed of four training modules. Module 1, entitled "Supervisory Techniques," contains the following topics of study: (1) Managerial Functions, (2) First-Line Supervisor's Role, (3) Leadership Styles, (4) Recruitment and Selection, (5) Classification and Compensation, (6) Employee Benefits, (7) Equal Employment Opportunity (EEO), (8) Loss Control, (9) Team Building, (10) Productivity, (11) Problem-Solving, and (12) Performance Appraisal. From the topics contained in Module 1, the following can be related directly to the dimension of Consideration: Managerial Functions, First-Line Supervisor's Role, Leadership Styles, Team Building, Problem-Solving, and Performance Appraisal.

"Interpersonal Communications in the Work Setting" is the title of Module 2, which contains these topics of study: (1) How We're Programmed, (2) How Perceptions of Self and Others Influence Our Work, (3) Interpersonal Relationships, (4) Barriers to Effective Organizational

Functioning, (5) Styles of Supervision, and (6) Problem-Solving and Time Management. Examination of the content of Module 2 reveals that the topics of How Perceptions of Self and Others Influence Our Work, Interpersonal Relationships, Barriers to Effective Organizational Functioning, and Styles of Supervision all relate to the area of human relations and considerate behavior.

Module 3 is entitled "Human Relations" and covers the following topics: (1) Communications, (2) Attitudes, (3) Motivation and Human Needs, (4) Human Relations and Leadership, (5) Group Dynamics, (6) Employees with Problems, (7) You and People You Interact With, and (8) Workshop in Human Relations Applications. In this module all of the topics bear a direct relationship to the dimension of Consideration.

The final module, Module 4, is entitled "Employee-Employer Relations." This unit covers the following study topics: (1) On Discipline, (2) Gripes, Grievances, and the Grievance Procedure, (3) Employee Relations in Tucson, and (4) Working with the Union. The only study topic in this module with direct human relations applications is the one on Employee Relations in Tucson.

Methods of Instruction

Module 1 consists of 22 hours of instruction given by Helen Cigrand. Mrs. Cigrand utilizes a combination of

structured lectures, including use of guest speakers, slides, and films with unstructured role-playing exercises and experiential group activities. Module 2 is composed of 18 hours of instruction by Phyllis Siken. Mrs. Siken employs a completely unstructured approach consisting of role-playing exercises, group experiential activities, and discussions.

Conrad Cormier teaches Module 3, which is also 18 hours in length. Mr. Cormier also employs unstructured methods of instruction. Much emphasis is placed on role-playing and group exercises. Some discussion and limited lectures are used. Finally, Module 4 is a 12-hour unit taught by Glenn Trowbridge. Mr. Trowbridge utilizes a completely structured approach of instruction, relying heavily on lectures with some class discussions.

CHAPTER III

RESEARCH DESIGN

Hypotheses

Based on the theory of human relations training and learning theory, the following specific research hypotheses were postulated for this study:

Hypothesis 1

It is hypothesized that, as a result of the Leadership Program, the Consideration score as measured by Fleishman's LOQ will increase more and the Structure score will decrease more for the group that received the training than for the group that did not receive the training.

Hypothesis 2a

It is hypothesized that, as a result of the Leadership Program, the Consideration score as measured by Fleishman's SBD will increase and the Structure score will decrease for the group that received the training, in the opinion of their superiors.

Hypothesis 2b

It is hypothesized that, as a result of the Leadership Program, the Consideration score as measured by the

SBD will increase and the Structure score will decrease for the group that received the training, in the opinion of their subordinates.

Design of the Study

The basic foundation of the experimental design of this research project is presented in Table 2.

Table 2. Experimental design

Group	No.	Measurement		
		Pre	Post	Post-Post (1 mo.)
Supervisors	17	LOQ	LOQ	---
Control	13	LOQ	LOQ	---
Control	10	---	LOQ	---
Control	26	---	LOQ	---
Superiors	14	SBD	---	SBD
Control	11	SBD	---	SBD
Subordinates	51	SBD	---	SBD
Control	39	SBD	---	SBD

The LOQ-measurement phase of the experimental design is adapted from the Solomon Four-Group Design (Solomon, 1949; Campbell and Stanley, 1966). The Solomon Four-Group Design is presented below:

B_1	X	B_2
B_3		B_4
	X	B_5
		B_6

B_1 is the experimental group of supervisors prior to participating in the Leadership Program (X). B_2 is the experimental group upon completion of the training program. B_3 is the control group measured at the same time as B_1. B_4 is the control group measured at the same time as B_2, but without having had any training. B_5 is a group of supervisors similar in character to B_1 measured by the LOQ only upon completion of the training program. B_6 is likewise a group of supervisors similar in character to B_1 measured at the same time as B_5, but without having participated in the Leadership Program.

The SBD-measurement phase of the experimental design is known as The Pretest-Posttest Control Group Design (Campbell and Stanley, 1966). This design is presented below:

A_1 X A_2

A_3 A_4

(and)

C_1 X C_2

C_3 C_4

A_1 represents the experimental group of superiors measured before their supervisors' training (X), while A_2 represents the same group upon completion of the supervisors' training. A_3 represents the control group of superiors measured at the same time as A_1. A_4 is the control group measured at the same time as A_2, but without having had their supervisors participate in the training program.

C_1 represents the experimental group of subordinates measured prior to the training of their supervisors (X), and measured again at the end of the program (C_2). C_3 is the control group of subordinates measured before their supervisors' training, and at the end of training (C_4) without the benefit of supervisor participation.

The independent variable being studied in this research design is exposure to a supervisory human relations training program. The dependent variables associated with the research design are the Consideration and Structure scores as measured by the LOQ and SBD.

Both the Pretest-Posttest Control Group Design and Solomon Four-Group Design are sound research designs because they are able to control various sources of invalidity. Both of these experimental designs are able to control the effects of the following sources of internal invalidity: (1) history, (2) maturation, (3) testing, (4) instrumentation, (5) statistical regression, (6) selection, (7) experimental mortality, and (8) selection-maturation interaction (Campbell and Stanley, 1966, p. 5). Both designs are somewhat less able to control external sources of invalidity such as the interaction effect of testing, the interaction effect of selection, and the reactive effects of experimental arrangements (Campbell and Stanley, 1966, pp. 5-6).

Test Subjects

The experimental test subjects are a group of 17 first-line supervisory personnel from various departments of the City of Tucson. The departments represented are: Fire, Sanitation, Human and Community Development, Water, Transportation, Library, Administration, Planning, Budget and Research, Fleet Services, Computer Services, Operations, Purchasing, City Clerk, City Courts, Personnel, and Finance-Accounting. The composition of each class is arranged so that the various City departments are represented in order to bring into focus, for training purposes, the degree of

interrelationships between these different departments. In many instances the test subjects have volunteered for the Leadership Program while others have been asked to attend by their immediate superiors. The test subjects themselves differ as to age, sex, education, race, seniority, and other factors. The basis for similarity among the group of test subjects is the organizational level from which they have come and the basic functional duties they perform.

The control group is comprised of 13 first-line supervisors similar to the experimental group in both departmental representation and personal characteristics. This group of supervisors has gone the longest period without any kind of human relations training, and therefore will enter the next session of the Leadership Program in January, 1978.

The group of supervisors known as B_5 is comprised of the participants of Session II of the Leadership Program (Fall, 1977). Similarly, the participants of Session II in the Spring, 1978, comprised the B_6 group of supervisors. Again, both of these groups possess very similar characteristics to the experimental and control test subjects discussed in the preceeding two paragraphs.

The superiors of each of the experimental test subjects make up the group of 14 experimental "superiors." At the time of this project, three supervisors did not have an

immediate superior; hence, a group of 14 instead of 17. The control group of superiors is made up of the superior of each member of the control group of supervisors. Again, for various reasons, two of the control group supervisors did not have an immediate superior at the time of this study.

From a list of the subordinates supplied by each experimental supervisor, a random selection of three individuals per supervisor was made in order to make up the experimental group of 51 subordinates. Similarly, a random selection of three individuals was made from a list supplied by the control supervisors, in order to make up the 39 control subordinates.

Instruments

LOQ

The LOQ, developed by Dr. Edwin A. Fleishman, was used to measure the reaction of the supervisors to the Leadership Program in terms of their attitudes on the Consideration and Structure scales. These two factors are defined by Pfeiffer, Heslin, and Jones (1976, p. 296):

> Consideration. Reflects the extent to which an individual is likely to have job relationships with his subordinates characterized by mutual trust, respect for their ideas, consideration of their feelings, and a certain warmth between himself and them. A low score indicates the individual is likely to be more impersonal in the relations with group members.

> Structure. Reflects the extent to which an individual is likely to define and structure his own role and those of his subordinates toward goal attainment. A high score on this dimension characterizes individuals who play a very active role in directing group activities through planning, communicating information, scheduling, criticizing, and trying out new ideas. A low score characterizes individuals who are likely to be relatively inactive in giving direction in these ways.

The LOQ is a self-administering instrument. It is necessary, for scoring purposes, that all 40 items be completed. If an item does not apply to the subject's situation, it should be answered as if it did apply. All respondents must be allowed to complete the questionnaire, which usually takes 15 to 20 minutes.

For scoring purposes, the alternatives to each item are weighted 4, 3, 2, 1, or 0. Within the test booklet itself is a self-scoring format with complete instructions. The maximum possible score on each of the two scales, Consideration and Structure, is 80.

In terms of the reliability of the LOQ, the range of both internal consistency and test-retest reliabilities is from .62 to .89 on the Consideration dimension, and .69 to .88 on the Structure dimension. The scores obtained on each scale have been proven to be independent of one another. Furthermore, the LOQ is designed to maximize construct validity (Fleishman, 1969, pp. 1-5).

The LOQ was originally developed out of the Ohio State University studies on leadership in the 1950s. The

most significant finding of this research project was the identification and isolation of the factors "Consideration" and "Initiating Structure" as the primary dimensions of leadership behavior in the organization. Over the past 25 years the LOQ has been used in a wide variety of business organizations for a number of different reasons. The LOQ has been proven effective if used for selection purposes, as a training aid, in training evaluation, and in training design (Fleishman, 1969, pp. 8-10). Notable management scientists who participated in the original research, and the research conducted over the past 25 years, were E. A. Fleishman, E. F. Harris, H. E. Burtt, R. M. Stogdill, A. E. Coons, A. W. Halpin, J. K. Hemphill, and B. J. Winer.

SBD

The SBD, also developed by Dr. Fleishman and a product of the Ohio State University Leadership Studies, was used to measure changes in the supervisors as recorded by their superiors and subordinates. The SBD measures the attitudes of the superiors and subordinates toward the behavior of the supervisors on the same scales of Consideration and Structure as the LOQ.

The administration of the SBD is exactly the same as for the LOQ. The only difference is that the SBD contains 48 items, instead of the 40 contained in the LOQ.

For scoring purposes, the SBD is again very similar to the LOQ. The difference here is that a separate scoring key is provided.

In terms of the reliability of the SBD, the range of internal consistency reliabilities and test-retest reliabilities is from .56 to .98 for Consideration, and .46 to .81 for Structure. As with the LOQ, the scales of the SBD have been proven stable, independent of one another, and the instrument is designed to maximize construct validity (Fleishman, 1972, pp. 2-7).

The SBD has also been utilized in a wide variety of organizations over the past 25 years. Through this research it has been shown that the SBD is an effective tool when used for appraisal purposes, counseling purposes, as an instructional aid in the evaluation of organizational climate, and in the evaluation of the impact of specific supervisory development training programs (Fleishman, 1972, p. 10).

CHAPTER IV

RESEARCH FINDINGS

Results

The raw data collected in accordance with the experimental design (see Table 2) represents over 300 interviews with 181 test subjects. The data are presented by group: Supervisors, Superiors, and Subordinates, and by time of measurement: pre-training, post-training, and post-post training (one month later). Data are presented for each scale, Consideration (C) and Structure (S), in terms of the group mean and standard deviation. These data are presented in Table 3.

A more complete description of this data, including the specific individual scores for each group, is contained in Appendix A. On the basis of this data, the statistical methodology to be applied will include analysis of variance, hypothesis testing, and nonparametric median tests.

Statistical Methodology

According to Campbell and Stanley (1966, p. 25) no single statistical tool exists which can adequately deal with all of the data provided by the Solomon Four-Group Design. Consequently, the best available statistical

Table 3. Research data.

	N	Mean C	Mean S	Standard Deviation C	Standard Deviation S
Supervisors					
B_1 Exp-pre	17	54.6	48.1	6.3	6.5
B_2 Exp-post	14	53.4	46.7	6.7	4.1
B_3 Con-pre	13	54.7	42.5	7.7	10.9
B_4 Con-post	13	53.5	39.7	6.4	10.0
B_5 Con-post	10	56.2	46.3	7.6	5.6
B_6 Con-post	26	53.1	46.7	6.4	7.3
Superiors					
A_1 Exp-pre	14	76.1	43.3	15.8	7.6
A_2 Exp-post	14	75.4	41.4	9.4	7.6
A_3 Con-pre	11	77.8	44.5	18.2	11.5
A_4 Con-post	11	78.3	41.6	15.6	9.9
Subordinates					
C_1 Exp-pre	51	69.9	36.7	23.3	9.6
C_2 Exp-post	51	69.2	38.5	20.6	8.7
C_3 Con-pre	39	85.5	41.4	15.1	9.6
C_4 Con-post	39	80.6	41.6	9.5	7.4

procedure is a 2 X 2 analysis of variance design, shown below:

	No X	X
Pretest	B_4	B_2
No Pretest	B_6	B_5

The null hypothesis (H_o) to be tested is that there is no significant difference in the mean scores from the pre-training measurement to the post-training measurement as recorded by the LOQ. The alternate hypothesis (H_a) to be tested against the null is that a change in mean scores did in fact occur and that this change occurred as postulated in Research Hypothesis 1. These tests will be conducted at the .05 level of significance.

The results of the analysis of variance performed on the Consideration scores is presented below:

Sources of Variation	Sums of Squares	Deg. of Freedom	Mean Squares	F
Training	19.254702	1	19.254702	.3948062
Pretest	30.754878	1	30.754878	.6306105
Interaction	35.479908	1	35.479908	.7274945
Within Cell	2877.43	59	48.770000	

With one and 59 degrees of freedom, $F_{.95} = 4.016$. The null hypothesis that there was no change in mean scores for Consideration must be accepted, and the alternate hypothesis cannot be accepted.

Similarly, the results of 2 X 2 analysis of variance conducted on the Structure scores is presented below:

Sources of Variation	Sums of Squares	Deg. of Freedom	Mean Squares	F
Training	151.41606	1	151.41606	2.388828
Pretest	153.25977	1	153.25977	2.417915
Interaction	191.44214	1	191.44214	3.020302
Within Cell	3739.72	59	63.38508	

With one and 59 degrees of freedom, $F_{.95} = 4.016$. The null hypothesis that there was no change in the mean scores for Structure must be accepted, and the alternate hypothesis must be rejected. A complete description of all analyses of variance calculations is contained in Appendix B. Comparison of the means shows that the major effect of training is due to the difference in B_4 and B_2, not in B_5 and B_6. Hence, the interaction effect exceeds that of either individual (X or No X) and pretest or no pretest. However, the interaction is not significant.

A second statistical procedure to be applied to the data resulting from the Solomon Four-Group Design is a median test using the chi-square (X^2) technique. This procedure is presented below:

H_o: The experimental group of supervisors (B_2) and the control group of supervisors (B_4) are independent samples from a common population.

H_a: The experimental group of supervisors and the control group of supervisors are not from a common population.

Test Statistic:

$$X^2 = \frac{n(|bc-ad| - n/2)^2}{(a+b)(c+d)(a+c)(b+d)}$$

Decision Rule: Reject the H_o if $X^2 > 3.841$ at the .05 level of significance for 1 d.f.

Computations: See Appendix C.

Decision: Accept H_o in both cases.
Reject H_a in both cases.

Conclusion: The experimental and control groups are independent samples from a common population. The observed number of plus signs and the observed number of minus signs are very nearly equal.

The most widely used and accepted test of significance on data provided by the Pretest-Posttest Control Group Design is the t-test (Campbell and Stanley, 1966, pp. 22-23). A comparison will be made between the pretest-posttest differences in the experimental group and the pretest-posttest difference in the the control group. Two sets of t-tests will be conducted, one for the group of Superiors and one for the group of Subordinates.

To further strengthen the significance of the t-tests to be performed, a direct comparison between the experimental and control groups will be made using the chi-square median test.

The application of the t-test to the data collected through the SBD, in the form of the Pretest-Posttest Control

Group Design, is presented below:

H_o: $\mu_1 - \mu_2 = 0$; where μ_1 is the mean of the pretest group and μ_2 is the mean of the posttest group.

H_a: $\mu_1 - \mu_2 \neq 0$

Test Statistics:

$$t = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Decision Rule:

Superiors A_1 and A_2:	Reject the H_o if $t > 1.771$ at the .05 level of significance for 13 d.f.
Superiors A_3 and A_4:	Reject the H_o if $t > 1.812$ at the .05 level of significance for 10 d.f.
Subordinates C_1 and C_2:	Reject the H_o if $t > 1.6775$ at the .05 level of significance for 50 d.f.
Subordinates C_3 and C_4:	Reject the H_o if $t > 1.6866$ at the .05 level of significance for 38 d.f.

Computations: See Appendix D.

Decision: Accept the H_o in all cases except for the C_3 and C_4 Subordinates on the Consideration dimension.
Reject the H_a in all cases except for the C_3 and C_4 Subordinates on the Consideration dimension.

Conclusion: For the most part, there was no difference in pretest versus posttest mean scores as a result of training. The significant result obtained for the control group of Subordinates was most probably due to a random error falling outside of the .05 level of significance. This single result does not affect the decision in a major way.

A second statistical procedure to be applied to the data resulting from the Pretest-Posttest Control Group Design is a median test using the chi-square technique. The results of this procedure indicate acceptance of the null hypothesis, that the groups are indeed independent samples from a common population, in all but one instance, Consideration scores for the group of Subordinates. The specific median test calculations are contained in Appendix E, and are the same as the calculations of the median test as applied to the data resulting from the Solomon Four-Group Design (p. 43 and Appendix C).

Analysis

The presentation and analysis of the data have shown that the Leadership Program did not have a significant effect on the attitudes of the test subjects. There was very little change, if any, between the pretest scores and posttest scores of both the experimental and control groups of Supervisors, as measured by the LOQ. Similarly, the same relationship between pre- and post-post measurements was observed of the experimental and control groups of Superiors

and Subordinates, as recorded by the SBD scores. However, all test groups involved in the project were shown to be independent random samples from the same population.

There are several reasons supporting the choice of analysis of variance, t-tests, and median tests as the statistical tools used to analyze the data obtained in this study. The foremost criterion in the selection of statistical methodology and techniques depends upon identification of the most appropriate procedures for the situation at hand. The two most critical aspects of the situation, which led directly to the choice of the three statistical procedures identified above, were that the data collected were assumed to be interval in nature and that several test groups were present, making comparison of scores possible. Also, each individual was considered to be an estimate of a population.

Analysis of variance was used because it allows for the simultaneous testing of differences among several population means. The 2 X 2 analysis of variance used in this study allowed for evaluation of the main effect of the training, the main effect of (the) pretesting, and the interaction of the testing and the training. Two major assumptions were made in using analysis of variance; that the populations were normally distributed and that their variances were approximately equal.

The t-test was used because it is a procedure that is sensitive to differences in population means. The t-test was used, therefore, to evaluate the differences between the experimental groups and the control groups, on a pretest-posttest basis. The assumptions made in using the t-test were that the populations were normally distributed, the samples were from a common population, and that the samples were independent.

The chi-square median test was used primarily because it showed whether or not samples were independent random samples from a common population. The median test is a nonparametric statistic which does not require the assumption of a normally distributed population. However, the other assumptions made in using the t-test were also applied to the median test. Sound experimental design and procedure dictated, in part, using the chi-square median test in conjunction with the t-test.

Several interesting relationships can be identified by studying the research data contained in Table 3. One point of interest is the range of standard deviations observed in this data. The standard deviations ranged between 6.3 and 23.3 on the Consideration dimension, and between 4.1 and 11.5 on the Structure dimension. While these differences are to be expected, the rather large variations indicate significant individual differences in responses on the

two dimensions of the questionnaires. The information contained in Appendix A supports this contention.

Second, certain expectations concerning the relationships of the mean scores as recorded by the two questionnaires in the form of the two experimental designs can be made. Concerning the Solomon Four-Group Design, we would expect the following relationships to naturally hold true:

$B_2 > B_1$ $\qquad$ $B_1 = B_3 = B_6$

$B_2 > B_4$ $\qquad$ $B_3 = B_4$

$B_5 > B_6$ $\qquad$ $B_1 = B_4$

$B_5 > B_3$ $\qquad$ $B_4 = B_6$

Regarding the Pretest-Posttest Control Group Design, we would expect these relationships to hold true:

$A_2 > A_1$ $\qquad$ $C_2 > C_1$

$A_2 > A_4$ $\qquad$ $C_2 > C_4$

$A_1 = A_3 = A_4$ $\qquad$ $C_1 = C_3 = C_4$

$A_3 = A_4$ $\qquad$ $C_3 = C_4$

With very few deviations in either direction, all of these relationships can in fact be seen as not holding true. Furthermore, the magnitude of these relationships is such that none of the statements can be considered as being statistically significant.

A final point of consideration is a comparison of the mean scores obtained in this study with tables of normative data. For the Supervisors' scores obtained on the LOQ, comparison with normative data for a group of 1,048 general supervisory personnel and 460 first-line supervisors suggests that the Supervisors scored average on the Consideration scale and low on the Structure scale (Fleishman, 1969, p. 13). Relating both the Superiors and Subordinates to normative data, their scores were also average on the Consideration scale and low on the Structure scale (Fleishman, 1972, p. 13).

In the author's opinion, several factors may have contributed to the overall rejection of the research hypotheses of this study. These factors are mentioned only as possible influences, since a detailed examination and analysis is beyond both the scope and purpose of this study. One possible contributing factor was the experimental mortality of the Supervisors B_1 and B_2 groups, where the group's size changed from 17 at the pre-measurement phase to 14 at the post-measurement phase. In a similar manner, the difference in sample sizes for the Superiors groups, experimental and control, also could have had an effect. In conjunction, another limiting factor is the somewhat small sample sizes associated with several of the groups, both experimental and control.

A third possible factor influencing the test subjects themselves is the climate and atmosphere within the City of Tucson's governmental organization. Specifically referenced here is a current environment made up of and dealing with tensions, politics, layoffs, cut-backs, elections, and budget cuts. A further influence on the outcome of this project was the relatively short time span between the post-measurement phase and the post-post measurement phase of only one month. The trade-off here is between allowing the trainees time to implement the concepts learned in training back on the job versus the fade-out that would naturally be expected as the time between post- and post-post measurements became longer.

Finally, the reactive or interactive effect of testing may have influenced the test subjects' responsiveness to the training program. In other words, the respondents may have been affected at the post-measurement by their recollection of the pre-measurement.

CHAPTER V

CONCLUSIONS

Summary

The purpose of this study was to test out a research design by means of evaluating the effectiveness of the Leadership Program. The intent was to determine whether or not the training program resulted in significant changes in attitudes, as measured by the LOQ, and/or behavior, as measured by the SBD. The issue was to evaluate whether or not the goals and objectives of the program were being achieved through measurement of the attitudes and behavior of the program's participants.

By way of summary, the results indicate that while all of the groups involved were shown to be independent random samples from common populations, no significant differences were observed between the post-training mean scores and the pre-training mean scores. Neither changes in attitudes or changes in behavior were observed as a result of the Leadership Program.

The behavior and attitudes of an individual in a business organization are, to a certain extent, shaped and influenced by those people above and below him or her in

the organization. In a training program, the ultimate goal is behavior change. This goal is partially achieved through changes in attitudes. Therefore, the success of any training program depends, to some degree, on those people in the organization immediately above and below the individual receiving the training.

According to Keith Davis and William Scott (1969, p. 145), "The objective of human relations training is to increase executives' ability in the management of human resources." Furthermore, the worth of any training program in the area of supervisory human relations is proved by evidence of the results achieved from the program. The evidence in this study calls for a rejection of the research hypotheses and indicates little direct effect of the training program on the participants.

Recommendations

Positive Aspects

The most positive aspect of the model presented in this study is the basic experimental research design used. The design is thorough, complete, and logical in its approach to the evaluation of a supervisory human relations training program. The research design is able to control all of the sources of internal invalidity, and even some of the sources of external invalidity, thereby enhancing

the predictive and evaluative abilities of the model. Moreover, the soundness of the design has been proved in many other training evaluation studies over the past two decades.

A second positive feature of the model is the specific instrumentation used. Both the LOQ and SBD are valid and reliable measures of attitudes and resulting behavior in the area of human relations and supervisory leadership. The statistical accuracy and foundation of these two instruments have been documented many times since their inception during the Ohio State University Studies on Leadership in the 1950s. These instruments have been used in a wide variety of settings, and have been shown to be particularly effective and efficient when used to measure and evaluate the results of management and supervisory development programs.

The statistical methodology used in this study is a further positive aspect of the model. Careful consideration was given to the choice of specific statistical tools to be used in the evaluation and analysis of the data. The statistical methodology employed was chosen as being most appropriate and most applicable to the overall research situation. Both a parametric and a nonparametric statistical technique were used to analyze data obtained from the Supervisors and from the Superiors and Subordinates. The

soundness of the statistical methodology was greatly enhanced through the use of these two types of techniques in combination with one another.

A final and somewhat more general positive outcome of this study is in the contribution it makes to the field of training evaluation research. It is a specific study conducted in a particular research setting designed to evaluate a unique management development program in the area of supervisory human relations. However, the model is not without some drawbacks, weaknesses, and less than positive aspects. These limitations will be discussed in the following section.

Negative Aspects

One negative aspect of the model presented in this study is its lack of longitudinal data collection. This problem can be corrected through the institution of further data collection and more controlled experimental verifications. The best approach would be to make post-post evaluations at periods of six, nine, twelve, and eighteen months after the completion of training. More extensive evaluation of the trainees back on the job is necessary since the objective of training is to produce a lasting change in behavior that will carry-over into the actual work environment. A more longitudinal approach to data collection is, therefore, recommended. Furthermore, from this form of data

collection, time series analysis and other related statistical procedures could be performed, thus enhancing the analysis and interpretation of the data.

Another negative feature of this model is its inability to identify the scores of specific test subjects. This problem can be rectified by identifying each test subject, on both the LOQ and SBD, at all pre-training, post-training, and post-post training measurements. This process would allow for identifying individual differences in learning and relative individual gains from training. The identification of test subjects could be made numerically on each questionnaire or by requesting the test subjects to actually record their names on the questionnaire cover. The major drawback of this recommendation is the possible dysfunctional reactions from the trainees. The fact that they can now be individually identified might affect the honesty of their responses to the questionnaire items which, in turn, would affect the end results obtained from the evaluation of the training program.

A third weakness of the model concerns the test subjects themselves. A more careful analysis of the similarities and dissimilarities of these individuals should be made in terms of the impact on attitudes and behavior. It is critical to determine the extent of the similarity between experimental and control groups in order to allow

proper interpretation of results. The characteristics of the experimental group versus the control group can very definitely influence their attitudes and behavior. In conjunction with determining the extent of similarity is the fact that, in this study, many of the trainees volunteered for the training program which indicates at least some interest in human relations on their part. This fact could very well influence the results of the study since the attitudes of these individuals are already favorable to and inclined toward human relations concepts, resulting in insignificant changes in attitudes, if any, as a result of training. On the other hand, their willingness to participate might suggest greater receptivity to the course material, and therefore increased likelihood of attitude and/or behavior changes.

Another drawback of this model is the comparatively small amount of time spent in its original inception. The development process is long range in nature and in order for the process to be evaluated properly, a significant amount of effort must be expended in planning and outlining the evaluation model. The model must be planned, constructed, and analyzed before the study ever begins. One negative consequence of this lack of lead time is a failure to adequately operationalize the specifications of the desired behavioral changes as a result of training.

A final negative aspect of this model concerns the fact that no design can precisely evaluate change. The model provided in this study purports to measure effects of a training program in the area of human relations. However, the attitudes and changes under evaluation are vague, nebulous, and ambiguous at best. This fact makes evaluation very difficult and limits the traditional evaluation methods.

The Leadership Program

The Training Division's supervisory human relations training program does provide trainees with an increased awareness of human relations and an increased knowledge and understanding of supervisory techniques. However, as a result of this training evaluation study, the Leadership Program does not seem to be able to change attitudes and subsequent behavior toward a more considerate and human relations orientation. Therefore, two major recommendations will be made that are designed to improve the effectiveness of the program.

First of all, a review of the content and structure of the Leadership Program should be undertaken. The program seems to be too general in coverage and too long in duration. Each module is different and each module is taught by a different instructor. The longer the duration of the program the more difficult it becomes to maintain rapport

between the trainer and trainees. Perhaps a careful consideration of the goals and objectives would prove beneficial in the overall review of the program.

This program review should consider what types of leader attitudes and behavior result in better efficiency, more effectiveness, higher productivity, and improved interpersonal supervisor-superior relationships. Specific identification could be made of topics which relate to and lead to more considerate behavior. Following up on this review, identification could be made of what to teach and how to teach it. New and even more effective methods of instruction could be developed in conjunction with a more efficient organization of the program's content.

The second recommendation involves the general supervisor-superior climate faced by the trainee back on the job after completion of the Leadership Program. The ideas, concepts, and techniques taught in the training program lack environmental support. The behavior taught in training must be made functional to the trainees by structuring the organizational climate such that rewards and punishments hinge on the transfer of training from the classroom to the job. According to Robert J. House (1967, p. 11), ". . . if development is to be successful, it must be geared not only to the participant's needs and learning abilities, but also to the particular requirements and

and practices of the organization in which he manages." A critical aspect of real environmental support for training is the state of the supervisor-superior relationship.

Therefore, it is recommended that a human relations training program for superiors be considered. Considering the significant impact of the superior climate (the "back-home" environment) on the supervisors, this recommendation would enhance the effectiveness of the supervisory human relations training program. Moreover, in a truly supportive back-home environment any changes resulting from the Leadership Program would be made more permanent and lasting.

Implications for Training Evaluation

There are several major implications for training evaluation research as a result of this study. The need for additional research is one implication for the field of training evaluation research. Not only to replicate this study, but to conduct more research of a significant nature into the evaluation of training programs. The contribution of meaningful knowledge and results is a necessity if the field as a whole is to advance and progress into the future.

Another implication is that conclusions are realistically limited by the nature of the training, the size of samples, the characteristics of the trainees, the independent variables involved, and so on. Furthermore, in order

to arrive at conclusions, a working assumption is generally made that changes in attitude with respect to specific human relations concepts must precede any behavioral changes.

To be effective, development must consider the numerous variables present in the trainee's environment. Moreover, the selection of learning content and teaching methods must be based on a careful analysis of the goals of the program in conjunction with the needs of the participants. Both the short and long range effects of change must also be considered. Specifically, the City of Tucson should make sure that it really needs and desires the qualities to be developed in the Leadership Program. If this is not the case, then the training effort is a waste of time.

Persons conducting training evaluation research should avoid erroneous and wasted evaluation conclusions by properly designing the study, selecting the statistical methodology carefully, analyzing the data properly and completely, and spending enough time planning the entire project. According to Daniel M. Goodacre (1957, p. 535), "A carefully conducted experimental evaluation will provide information not only on the over-all value of a particular program but, if properly designed, on the specific strengths and weaknesses of content and method as well."

The situational approach to leadership is not just a style high in Consideration and high in Structure. At the same time, more Consideration and less Structure is not necessarily a good combination. Leaders must be capable of multiple behaviors and behavior patterns. In order to develop more effective leaders, training must broaden in order to permit these individuals to develop the skills, ability, and desire to make significant contributions to the organization. It follows, therefore, that according to Edwin B. Flippo (1971, p. 223):

> . . . each organization must design its own particular program to suit the climate of the firm, the organizational level for which training is required, the particular characteristics of the personnel to be developed, the recognized specific training requirements, and the availability of economic resources. There is no one best program of management development.

The final major implication for training evaluation is that management development is far from a panacea. The use of training evaluation models in management development has many limitations. For the organization, management development and models of evaluation are most effective when used in combination with other types of performance-change programs.

APPENDIX A

RESEARCH DATA FOR ALL TEST GROUPS

Supervisors--Experimental Group B_1

C	S	C^2	S^2
46	37	2116	1369
46	40	2116	1600
47	42	2209	1764
48	42	2304	1764
48	44	2304	1936
50	45	2500	2025
53	47	2809	2209
55	47	3025	2209
55	47	3025	2209
56	48	3136	2304
57	48	3249	2304
57	49	3249	2401
58	50	3364	2500
58	52	3364	2704
62	58	3844	3364
64	60	4096	3600
68	61	4624	3721

$\Sigma C = 928.$ $\quad \Sigma S = 817.$

$\mu C = 54.588235$ $\quad \mu S = 48.058823$

$\phi C = 55.$ $\quad \phi S = 47.$

$\Sigma C^2 = 51334.$ $\quad \Sigma S^2 = 39983.$

$\sigma^2 = 39.771647$ $\quad \sigma^2 = 42.290705$

$\sigma = 6.3064765$ $\quad \sigma = 6.5031304$

Supervisors--Experimental Group B_2

C	S	C^2	S^2
40	39	1600	1521
46	39	2116	1521
46	43	2116	1849
48	43	2304	1849
52	43	2704	1849
52	44	2704	1936
52	45	2704	2025
55	46	3025	2209
55	47	3025	2116
57	48	3249	2304
59	50	3481	2500
60	52	3600	2704
60	53	3600	2809
66	60	4356	3600

$\Sigma C = 748.$
$\mu C = 53.428571$
$\phi C = 53.5$
$\Sigma C^2 = 40584.$
$\sigma^2 = 44.244928$
$\sigma = 6.6516861$

$\Sigma S = 654.$
$\mu S = 46.714285$
$\phi S = 45.5$
$\Sigma S^2 = 30792.$
$\sigma^2 = 17.204142$
$\sigma = 4.1477876$

Supervisors--Control Group B_3

C	S	C^2	S^2
45	11	2025	121
46	33	2116	1089
49	39	2401	1521
50	41	2500	1681
50	42	2500	1764
51	42	2601	1764
53	44	2809	1936
54	47	2916	2209
55	47	3025	2209
56	49	3136	2401
65	49	4225	2401
66	51	4356	2601
71	58	5041	3364

$\Sigma C = 711.$ $\Sigma S = 553.$

$\mu C = 54.692307$ $\mu S = 42.538461$

$\phi C = 53.$ $\phi S = 44.$

$\Sigma C^2 = 39651.$ $\Sigma S^2 = 25061.$

$\sigma^2 = 58.828461$ $\sigma^2 = 118.24853$

$\sigma = 7.6699713$ $\sigma = 10.874213$

Supervisors--Control Group B_4

C	S	C^2	S^2
43	12	1849	144
47	33	2209	1089
49	35	2401	1225
50	36	2500	1296
51	38	2601	1444
52	38	2704	1444
53	40	2809	1600
54	42	2916	1764
54	44	2916	1936
55	45	3025	2025
57	50	3249	2500
61	50	3721	2500
70	53	4900	2809

$\Sigma C = 696.$
$\mu C = 53.538461$
$\phi C = 53.$
$\Sigma C^2 = 37800.$
$\sigma^2 = 41.325461$
$\sigma = 6.4284882$

$\Sigma S = 516.$
$\mu S = 39.692307$
$\phi S = 40.$
$\Sigma S^2 = 21776.$
$\sigma^2 = 99.597692$
$\sigma = 9.9798643$

Supervisors--Control Group B_5

C	S	C^2	S^2
43	36	1849	1296
46	39	2116	1521
48	43	2304	1849
57	43	3249	1849
57	47	3249	2209
60	50	3600	2500
63	51	3969	2601
65	52	4225	2704
66	54	4356	2916

$\Sigma C = 562.$
$\mu C = 56.2$
$\phi C = 57.$
$\Sigma C^2 = 32166.$
$\sigma^2 = 58.16$
$\sigma = 7.6262703$

$\Sigma S = 463.$
$\mu S = 46.3$
$\phi S = 47.5$
$\Sigma S^2 = 21749.$
$\sigma^2 = 31.21$
$\sigma = 5.586591$

Supervisors--Control Group B_6

C	S	C^2	S^2
40	35	1600	1225
44	37	1936	1369
46	38	2116	1444
46	39	2116	1521
47	39	2209	1521
48	39	2304	1521
48	40	2304	1600
48	41	2304	1681
48	43	2304	1849
50	43	2500	1849
50	44	2500	1936
52	44	2704	1936
52	45	2704	2025
53	45	2809	2025
53	47	2809	2209
56	48	3136	2304
57	48	3249	2304
57	49	3249	2401
57	53	3249	2809
59	54	3481	2916
60	55	3600	3025
61	57	3721	3249
63	58	3969	3364
63	59	3969	3481
64	60	4096	3600

$\Sigma C = 1381.$
$\mu C = 53.115384$
$\phi C = 52.5$
$\Sigma C^2 = 74419.$
$\sigma^2 = 41.025153$
$\sigma = 6.405088$

$\Sigma S = 1214.$
$\mu S = 46.692307$
$\phi S = 45.$
$\Sigma S^2 = 58080.$
$\sigma^2 = 53.674576$
$\sigma = 7.3262934$

Superiors--Experimental Group A_1

C	S	C^2	S^2
41	28	1681	784
45	34	2025	1156
59	36	3481	1296
74	37	5476	1369
75	38	5625	1444
76	43	5776	1849
82	43	6724	1849
84	45	7056	2025
86	47	7396	2209
86	47	7396	2209
87	47	7569	2209
90	52	8100	2704
90	53	8100	2809
90	56	8100	3136

$\Sigma C = 1065.$
$\mu C = 76.071428$
$\phi C = 82.$
$\Sigma C^2 = 84505.$
$\sigma^2 = 249.20921$
$\sigma = 15.786361$

$\Sigma S = 606.$
$\mu S = 43.285714$
$\phi S = 43.$
$\Sigma S^2 = 27048.$
$\sigma^2 = 58.347$
$\sigma^2 = 7.6385207$

Superiors--Experimental Group A_2

C	S	C^2	S^2
55	25	3025	625
65	31	4225	961
65	36	4225	1296
70	36	4900	1296
72	39	5184	1521
74	40	5476	1600
75	41	5625	1681
76	42	5776	1764
78	44	6084	1936
80	45	6400	2025
83	46	6889	2116
85	51	7225	2601
87	51	7569	2601
91	53	8281	2809

$\Sigma C = 1056.$
$\mu C = 75.428571$
$\phi C = 75.5$
$\Sigma C^2 = 80884.$
$\sigma^2 = 87.959214$
$\sigma = 9.3786573$

$\Sigma S = 580.$
$\mu S = 41.428571$
$\phi S = 41.5$
$\Sigma S^2 = 24832.$
$\sigma^2 = 57.387785$
$\sigma = 7.5754725$

Superiors--Control Group A_3

C	S	C^2	S^2
42	27	1764	729
54	30	2916	900
62	33	3844	1089
75	35	5625	1225
75.6	41	5715.36	1681
82	46	6724	2116
83	48	6889	2304
85	55	7225	3025
92	56	8464	3136
101	57	10201	3249
104	61.4	10816	3769.96

$\Sigma C = 855.6$
$\mu C = 77.781818$
$\phi C = 82.$
$\Sigma C^2 = 70183.36$
$\sigma 2 = 330.29427$
$\sigma = 18.173999$

$\Sigma S = 489.4$
$\mu S = 44.490909$
$\phi S = 46.$
$\Sigma S^2 = 23223.96$
$\sigma 2 = 131.82818$
$\sigma = 11.481645$

Superiors--Control Group A_4

C	S	C^2	S^2
49	21	2401	441
55	26	3025	676
70	39	4900	1521
71	40	5041	1600
79	41	6241	1681
80	43	6400	1849
81	45	6561	2025
84	47	7056	2209
95	48	9025	2304
97	52	9409	2704
100	56	10000	3136

$\Sigma C = 861.$
$\mu C = 78.272727$
$\phi C = 80.$
$\Sigma C^2 = 70059.$
$\sigma^2 = 242.38018$
$\sigma = 15.568563$

$\Sigma S = 458.$
$\mu S = 41.636363$
$\phi S = 43.$
$\Sigma S^2 = 20146.$
$\sigma^2 = 97.867818$
$\sigma = 9.8928164$

Subordinates--Experimental Group C_1

C	S	C^2	S^2
14	10	196	100
14	20	196	400
29	21	841	441
30	21	900	441
36	22	1296	484
41	25	1681	625
41	27	1681	729
50	27	2500	729
50	29	2500	841
51	29	2601	841
51	29	2601	841
52	31	2704	961
53	31	2809	961
53	31	2809	961
55	32	3025	1024
57	32	3249	1024
59	33	3481	1089
59	34	3481	1156
60	34	3600	1156
64	35	4096	1225
64	35	4096	1225
65	35	4225	1225
65	35	4225	1225
66	36	4356	1296
70	36	4900	1296
71	37	5041	1369
72	37	5184	1369
73	37	5329	1369
77	38	5929	1444
78	39	6084	1521
80	39	6400	1521
80	40	6400	1600
82	40	6724	1600
82	40	6724	1600
86	40	7396	1600
87	40	7569	1600
89	42	7921	1764
90	43	8100	1849
92	43	8464	1849
93	44	8649	1936
95	45	9025	2025
95	45	9025	2025
96	45	9216	2025

Subordinates--Experimental Group C_1--Continued

C	S	C^2	S^2
97	46	9409	2116
99	46	9801	2116
100	47	10000	2209
100	48	10000	2304
100	49	10000	2401
100	51	10000	2601
101	59	10201	3481
102	61	10404	3721

$\Sigma C = 3566.$
$\mu C = 69.921568$
$\phi C = 71.$
$\Sigma C^2 = 277044.$
$\sigma^2 = 543.2096$
$\sigma = 23.306857$

$\Sigma S = 1871$
$\mu S = 36.686274$
$\phi S = 37.$
$\Sigma S^2 = 73311.$
$\sigma^2 = 91.587862$
$\sigma = 9.5701547$

Subordinates--Experimental Group C_2

C	S	C^2	S^2
12	15	144	225
25	19	625	361
26	22	676	484
36	26	1296	676
41	26	1681	676
43	30	1849	900
45	30	2025	900
46	30	2116	900
46	31	2116	961
50	31	2500	961
55	32	3025	1024
58	32	3364	1024
59	34	3481	1156
60	34	3600	1156
63	35	3969	1225
63	35	3969	1225
64	35	4096	1225
65	35	4225	1225
65	36	4225	1296
65	37	4225	1369
66	37	4356	1369
68	37	4624	1369
68	38	4624	1444
69	39	4761	1521
70	40	4900	1600
71	40	5041	1600
73	40	5329	1600
75	40	5625	1600
75	40	5625	1600
75	40	5625	1600
75	40	5625	1600
75	41	5625	1681
75	41	5625	1681
76	42	5776	1764
80	43	6400	1849
80	44	6400	1936
82	44	6724	1936
84	44	7056	1936
85	45	7225	2025
87	45	7569	2025
90	45	8100	2025

Subordinates--Experimental Group C_2--Continued

C	S	C^2	S^2
90	46	8100	2116
94	47	8836	2209
95	48	9025	2304
95	49	9025	2401
100	50	10000	2500
100	50	10000	2500
101	53	10201	2809
101	56	10201	3136
103	60	10609	3600

$\Sigma C = 3529.$
$\mu C = 69.196078$
$\phi C = 70.$
$\Sigma C^2 = 265935.$
$\sigma^2 = 426.3145$
$\sigma = 20.647384$

$\Sigma S = 1964.$
$\mu S = 38.509803$
$\phi S = 40.$
$\Sigma S^2 = 79530.$
$\sigma^2 = 76.406784$
$\sigma = 8.7410974$

Subordinates--Control Group C_3

C	S	C^2	S^2
45	28	2025	784
52	29	2704	841
55	29	3025	841
65	29	4225	841
67	31	4489	961
68	31	4624	961
70	31	4900	961
70	32	4900	1024
75	32	5625	1024
78	33	6084	1089
78	33	6084	1089
80	34	6400	1156
80	35	6400	1225
80	36	6400	1296
84	37	7056	1369
87	37	7569	1369
88	38	7744	1444
88	39	7744	1521
90	40	8100	1600
90	41	8100	1681
91	41	8281	1681
91	41	8281	1681
92	42	8464	1764
93	42	8649	1764
93	43	8649	1849
94	44	8836	1936
94	45	8836	2025
96	45	9216	2025
96	48	9216	2304
96	50	9216	2500
97	51	9409	2601
99	51	9801	2601
100	52	10000	2704
101	53	10201	2809
103	54	10609	2916
104	54	10816	2916
104	55	10816	3025
104	64	10816	4096
107	64	11449	4096

Subordinates--Control Group C_3--Continued

$\Sigma C = 3345.$	$\Sigma S = 1614.$
$\mu C = 85.76934$	$\mu S = 41.384615$
$\phi C = 90.$	$\phi S = 41.$
$\Sigma C^2 = 295759.$	$\Sigma S^2 = 70400.$
$\sigma^2 = 227.20333$	$\sigma^2 = 92.44182$
$\sigma = 15.073265$	$\sigma = 9.6146669$

Subordinates--Control Group C_4

C	S	C^2	S^2
61	24	3721	576
63	30	3969	900
66	32	4356	1024
67	33	4489	1089
68	34	4624	1156
68	34	4624	1156
70	34	4900	1156
70	34	4900	1156
72	37	5184	1369
74	37	5476	1369
74	37	5476	1369
76	38	5776	1444
78	39	6084	1521
78	39	6084	1521
79	39	6241	1521
80	39	6400	1521
80	40	6400	1600
80	41	6400	1681
81	41	6561	1681
81	41	6561	1681
81	42	6561	1764
82	42	6724	1764
82	42	6724	1764
83	42	6889	1764
84	42	7056	1764
84	43	7056	1849
86	43	7396	1849
86	44	7396	1936
87	46	7569	2116
88	47	7744	2209
88	48	7744	2304
88	49	7744	2401
90	49	8100	2401
91	49	8281	2401
91	52	8281	2704
92	52	8464	2704
96	52	9216	2704
97	54	9409	2916
101	62	10201	3844

Subordinates--Control Group C_4--Continued

ΣC	= 3143.	ΣS	= 1623.
μC	= 80.589743	μS	= 41.615384
ϕC	= 81.	ϕS	= 41.
ΣC^2	= 256781.	ΣS^2	= 69649.
σ^2	= 89.421538	σ^2	= 54.031564
σ	= 9.4562962	σ	= 7.3506165

APPENDIX B

ANALYSIS OF VARIANCE CALCULATIONS

The following analysis of variance calculations are based on Summers and Peters (1973, pp. 276-281) and Winer (1971, pp. 402-404).

For Consideration

Cell Frequencies:

	No X	X	Total
Pretest	13	14	27
No Pretest	26	10	36
Total	39	24	63

Cell Totals:

	No X	X	Total
Pretest	696	748	1444
No Pretest	1381	562	1943
Total	2077	1310	3387

Cell Means:

	No X	X	Total
Pretest	53.54	53.43	106.97
No Pretest	53.12	56.20	109.32
Total	106.66	109.63	216.29

$\bar{n}_h$ = The harmonic mean of the number of observations per cell.

$\bar{A}_i$ = The estimate of the mean for the i rows.

$\bar{B}_j$ = The estimate of the mean for the j columns.

$\bar{G}$ = The grand mean in the population.

SS_a = The variation due to the training.

SS_b = The variation due to the pretesting.

SS_{ab} = The variation due to interaction.

$SS_{w.\ cell}$ = The pooled within-cell variation.

d.f. = The degrees of freedom for a given source of variation.

p = The number of rows.

q = The number of columns.

$$\bar{n}_h = \frac{(2)\ (2)}{1/13 + 1/14 + 1/26 + 1/10}$$

$$= 13.946369 \text{ observations per cell}$$

$$\bar{A}_i = \frac{53.54 + 53.43}{2}$$

$$= 53.485$$

$$\bar{A}_2 = \frac{53.12 + 56.20}{2}$$

$$= 54.660$$

$$\bar{B}_1 = \frac{53.54 + 53.12}{2}$$

$$= 53.330$$

$$\bar{B}_2 = \frac{53.43 + 56.20}{2}$$

$$= 54.815$$

$$\bar{G} = \frac{53.485 + 54.660 + 53.330 + 54.815}{4}$$

$$= 54.0725$$

$$SS_a = \bar{n}_h q \Sigma (\bar{A}_1 - \bar{G})^2 = 19.254702$$

$$SS_b = \bar{n}_h p \Sigma (\bar{B}_j - \bar{G})^2 = 30.754878$$

$$SS_{ab} = \overline{n}_h \Sigma(\overline{AB}_{ij} - \overline{A}_i - \overline{B}_j + \overline{G})^2 = 35.479908$$

$$SS_{w.\ cell} = \Sigma X^2 - \frac{(\Sigma X)^2}{n} = 2877.43$$

2 X 2 ANOVA

Sources of Variation	Sums of Squares	Degrees of Freedom	Mean Squares	F
Training	19.254702	1	19.254702	.3948062
Pretest	30.754878	1	30.754878	.6306105
Interation	35.479908	1	35.479908	.7274945
Within Cell	2877.43	59	48.770000	

$F_{.95(1,59)} = 4.016$

Accept the null hypothesis
Reject the alternate hypothesis

For Structure

Cell Frequencies:

	No X	X	Total
Pretest	13	14	27
No Pretest	26	10	36
Total	39	24	63

Cell Totals:

	No X	X	Total
Pretest	516	654	1170
No Pretest	1214	463	1677
Total	1730	1117	2847

Cell Means:

	No X	X	Total
Pretest	39.69	46.71	85.40
No Pretest	46.69	46.30	92.99
Total	86.38	93.01	179.39

$\bar{n}_h$ = 13.946369

$\bar{A}_1$ = 43.200

$\bar{A}_2$ = 46.495

$\bar{B}_1$ = 43.190

$\bar{B}_2$ = 46.505

$\bar{G}$ = 44.8475

SS_a = 151.41606

SS_b = 153.25977

SS_{ab} = 191.44214

$SS_{w.\ cell}$ = 3739.72

2 X 2 ANOVA

Sources of Variation	Sums of Squares	Degrees of Freedom	Mean Squares	F
Training	151.41606	1	151.41606	2.388828
Pretest	153.25977	1	153.25977	2.417915
Interaction	191.44214	1	191.44214	3.020302
Within Cell	3739.72	59	63.38508	

$F_{.95(1,59)}$ = 4.016

Accept the null hypothesis
Reject the alternate hypothesis

APPENDIX C

MEDIAN TEST CALCULATIONS

For Consideration

X	f
70	1
66	1
61	1
60	2
59	1
57	2
55	3
54	2
53	1
52	4
51	1
50	1
49	1
48	1
47	1
46	2
43	1
40	1

Median = 52.5

Experimental		Control	
X	Sign	X	Sign
66	+	70	+
60	+	61	+
60	+	57	+
59	+	55	+
57	+	54	+
55	+	54	+
55	+	53	+
52	-	52	-
52	-	51	-
52	-	50	-
48	-	49	-
46	-	47	-
46	-	43	-
40	-		

	-	+	Totals
Experimental	7 a	7 c	14
Control	6 b	7 d	13
Totals	13	14	27

$$X^2 = \frac{27(|42-49| - 27/2)^2}{(13)(14)(14)(13)}$$

$$= .0344387$$

For Structure

X	f
60	1
53	2
52	1
50	3
48	1
47	1
46	1
45	2
44	2
43	3
42	1
40	1
39	2
38	2
36	1
35	1
33	1
12	1

Median = 43.5

Experimental		Control	
X	Sign	X	Sign
60	+	53	+
53	+	50	+
52	+	50	+
50	+	45	+
48	+	44	+
47	+	42	-
46	+	40	-
45	+	38	-
44	+	38	-
43	-	36	-
43	-	35	-
43	-	33	-
39	-	12	-
39	-		

	-	+	Totals
Experimental	5 a	9 c	14
Control	8 b	5 d	13
Totals	13	14	27

$$X^2 = \frac{27(|72-25| - 27/2)^2}{(40)(45)(45)(40)}$$

$$= .009352$$

APPENDIX D

T-TEST CALCULATIONS

The following t-test calculations are based on Hamburg (1970, pp. 351-354) and Summers and Peters (1973, pp. 203-222).

For Superiors Experimental

Consideration =

$$t = \sqrt{\frac{76.1 - 75.4}{\frac{249.2}{14} + \frac{88}{14}}}$$

$$= .1426324$$

Structure =

$$t = \sqrt{\frac{43.3 - 41.4}{\frac{58.3}{14} + \frac{57.4}{14}}}$$

$$= .6609321$$

For Superiors Control

Consideration =

$$t = \sqrt{\frac{78.8 - 78.3}{\frac{330.3}{11} + \frac{242.4}{11}}}$$

$$= -.0692951$$

Structure =

$$t = \sqrt{\frac{44.5 - 41.6}{\frac{131.8}{11} + \frac{97.9}{11}}}$$

$$= .6346201$$

For Subordinates Experimental

Consideration =

$$t = \sqrt{\frac{69.9 - 69.2}{\frac{543.2}{51} + \frac{426.3}{51}}}$$

$$= .1605496$$

Structure =

$$t = \sqrt{\frac{36.7 - 38.5}{\frac{91.6}{51} + \frac{76.4}{51}}}$$

$$= - .9917517$$

For Subordinates Control

Consideration =

$$t = \sqrt{\frac{85.8 - 80.6}{\frac{227.2}{39} + \frac{89.4}{39}}}$$

$$= 1.8250728$$

Structure =

$$t = \sqrt{\frac{41.4 - 41.6}{\frac{92.4}{39} + \frac{54.0}{39}}}$$

$$= - .1032266$$

APPENDIX E

MEDIAN TEST CALCULATIONS FOR SBD DATA

The following median test calculations are based on Edwards (1967, pp. 334-337).

For Superiors

Consideration

X	f
100	1
97	1
95	1
91	1
87	1
85	1
84	1
83	1
81	1
80	2
79	1
78	1
76	1
75	1
74	1
72	1
71	1
70	2
65	2
55	2
49	1

Median = 78.5

Experimental		Control	
X	Sign	X	Sign
91	+	100	+
87	+	97	+
85	+	95	+
83	+	84	+
80	+	81	+
78	-	80	+
76	-	79	+
75	-	71	-
74	-	70	-
72	-	55	-
65	-	49	-
65	-		
55	-		

	-	+	Totals
Experimental	9 a	5 c	14
Control	4 b	7 d	11
Totals	13	12	25

$X^2 = .9680423$

Structure

X	f
56	1
52	1
51	2
48	1
47	1
46	1
45	2
44	1
43	1
42	1
41	2
40	2
39	2
36	2
31	1
26	1
25	1
21	1

Median = 42.3

Experimental		Control	
X	Sign	X	Sign
53	+	56	+
51	+	52	+
51	+	48	+
46	+	47	+
45	+	45	+
44	+	43	+
42	-	41	-
41	-	40	-
40	-	39	-
39	-	26	-
36	-	21	-
36	-		
31	-		
25	-		

	-	+	Totals
Experimental	8 a	6 c	14
Control	5 b	6 d	11
Totals	13	12	25

$X^2 = .0314789$

For Subordinates

Consideration

X	f
103	1
101	3
100	2
97	1
96	1
95	2
94	1
92	1
91	2
90	3
88	3
87	2
86	2
85	1
84	3
83	1
82	3
81	3
80	5
79	1
78	2
76	2
75	6
74	2
73	1
72	1
71	1
70	3
69	1
68	4
67	1
66	2
65	3
64	2
63	3
61	1
60	1
59	1
58	1
55	1
50	1
46	2

Experimental		Control	
X	Sign	X	Sign
103	+	101	+
101	+	97	+
101	+	96	+
100	+	92	+
100	+	91	+
95	+	91	+
95	+	90	+
94	+	88	+
90	+	88	+
90	+	88	+
87	+	87	+
85	+	86	+
84	+	86	+
82	+	84	+
80	+	84	+
80	+	83	+
76	+	82	+
75	+	81	+
75	+	81	+
75	+	81	+
75	+	80	+
75	+	80	+
75	+	80	+
73	0	79	+
71	-	78	+
70	-	78	+
69	-	76	+
68	-	74	+
68	-	74	+
66	-	72	-
65	-	70	-
65	-	70	-
65	-	68	-
64	-	68	-
64	-	68	-
63	-	67	-
63	-	66	-
60	-	63	-
59	-	61	-

X	f
45	1
43	1
41	1
36	1
26	1
25	1
12	1

Median - 73.0

Experimental		Control	
X	Sign	X	Sign
58	-		
55	-		
50	-		
46	-		
46	-		
45	-		
43	-		
41	-		
36	-		
26	-		
25	-		
12	-		

	-	+	Totals
Experimental	27 a	23 c	50
Control	9 b	30 d	39
Totals	36	53	89

$X^2 = 7.4614524$

Structure

X	f
62	1
60	1
56	1
54	1
53	1
52	3
50	2
49	3
48	2
47	2
46	2
45	3
44	4

Experimental		Control	
X	Sign	X	Sign
60	+	62	+
56	+	54	+
53	+	52	+
50	+	52	+
50	+	52	+
49	+	49	+
48	+	49	+
47	+	49	+
46	+	48	+
45	+	47	+

X	f
43	3
42	6
41	5
40	8
39	5
38	2
37	6
36	1
35	5
34	6
33	1
32	1
31	2
30	4
26	2
22	1
24	1
19	1
15	1

Median = 40.5

Experimental		Control	
X	Sign	X	Sign
45	+	46	+
45	+	44	+
44	+	43	+
44	+	43	+
44	+	42	+
43	+	42	+
42	+	42	+
41	+	42	+
41	+	42	+
40	-	41	+
40	-	41	+
40	-	41	+
40	-	40	-
40	-	39	-
40	-	39	-
40	-	39	-
39	-	39	-
38	-	38	-
37	-	37	-
37	-	37	-
37	-	37	-
36	-	34	-
35	-	34	-
35	-	34	-
35	-	34	-
35	-	33	-
35	-	32	-
34	-	30	-
34	-	24	-
32	-		
32	-		
31	-		
31	-		
30	-		
30	-		
30	-		
26	-		
26	-		
22	-		
19	-		
15	-		

	-	+	Totals
Experimental	32 a	19 c	51
Controls	17 b	22 d	39
Totals	49	41	90

$X^2 = 2.5427656$

LIST OF REFERENCES

Biggs, Donald A., Huneryager, S. G., and Delaney, James J. "Leadership Behavior: Interpersonal Needs and Effective Supervisory Training." Personnel Psychology 19 (Autumn 1966):311-320.

Blumenfeld, Warren S., and Holland, Max G. "A Model for the Empirical Evaluation of Training Effectiveness." Personnel Journal 50 (August 1971):637-640.

Bunker, Kerry A., and Cohen, Stephen L. "The Rigors of Training Evaluation: A Discussion and Field Demonstration." Personnel Psychology 30 (Winter 1977): 525-541.

Campbell, Donald T., and Stanley, Julian C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally & Company, 1966.

Carron, Theodore J. "Human Relations Training and Attitude Change: A Vector Analysis." Personnel Psychology 17 (Winter 1964):403-424.

Catalanello, Ralph F., and Kirkpatrick, Donald L. "Evaluating Training Programs--The State of the Art." Training and Development Journal 22 (May 1968):2-9.

Davis, Keith, and Scott, William G. Human Relations and Organizational Behavior. New York: McGraw-Hill Book Company, 1969.

DiVesta, Francis J. "Instructor-Centered and Student-Centered Approaches in Teaching a Human Relations Course." Journal of Applied Psychology 38 (October 1954):329-335.

Edwards, Allen L. Statistical Methods. New York: Holt, Rinehart, and Winston, Inc., 1967.

Elkins, Aaron. "Some Views on Management Training." Personnel Journal 56 (June 1977): 305-306, 311.

Fleishman, Edwin A. "Leadership Climate, Human Relations Training, and Supervisory Behavior." Personnel Psychology 6 (Summer 1953):205-222.

_____. Manual for the Leadership Opinion Questionnaire. Chicago: Science Research Associates, Inc., 1969.

_____. Manual for the Supervisory Behavior Description Questionnaire. Washington, D.C.: American Institute for Research, 1972.

Fleishman, Edwin A., and Harris, Edwin F. "Patterns of Leadership Behavior Related to Employee Grievances and Turnover." Personnel Psychology 15 (Spring 1962):43-56.

Fleishman, Edwin A., Harris, Edwin F., and Burtt, Harold E. Leadership and Supervision in Industry: An Evaluation of a Supervisory Training Program. Columbus, Ohio: Bureau of Educational Research, Ohio State University, 1955.

Flippo, Edwin B. Principles of Personnel Management. New York: McGraw-Hill Book Company, 1971.

Goodacre, Daniel M. "Evaluating a Human Relations Training Program." Personnel 33 (May 1957):534-538.

Hamburg, Morris. Statistical Analysis for Decision Making. New York: Harcourt, Brace & World, Inc., 1970.

Hand, Herbert H., and Slocum, John W. "Human Relations Training for Middle Management: A Field Experiment." Academy of Management Journal 13 (December 1970): 403-410.

_____. "A Longitudinal Study of the Effects of a Human Relations Training Program on Managerial Effectiveness." Journal of Applied Psychology 56 (October 1972):412-417.

House, Robert J. "A Commitment Approach to Management Development." California Management Review VII (Spring 1965):15-28.

_____. Management Development. Ann Arbor, Michigan: University of Michigan, 1967.

Katzell, Raymond A. "Testing a Training Program in Human Relations." Personnel Psychology 1 (Autumn 1948): 319-329.

Miles, Mathew B. "Changes During the Following Laboratory Training: A Clinical Experimental Study." The Journal of Applied Behavioral Science 1 (1965): 215-242.

Nixon, George. People, Evaluation and Achievement. Houston: Gulf Publishing Company, 1973.

Pfeiffer, J. William, Heslin, Richard, and Jones, John E. Instrumentation in Human Relations Training. La Jolla, California: University Associates, Inc., 1976.

Solomon, Richard L. "An Extention of Control Group Design." Psychological Bulletin 46 (March 1949):137-150.

Stroud, Peggy V. "Evaluating a Human Relations Training Program." Personnel 36 (November-December 1959): 52-60.

Summers, George W., and Peters, William S. Basic Statistics in Business and Economics. Belmont, California: Wadsworth Publishing Company, Inc., 1973.

Valdez, Joel D. "Memorandum to Department Heads, Division Heads, and Supervisory Personnel." Tucson, Arizona: City of Tucson, 21 January, 1977.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill Book Company, 1971.