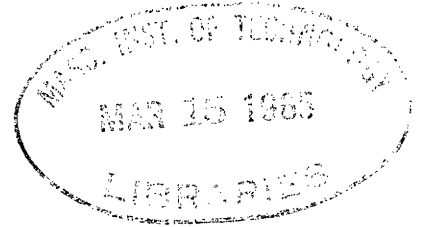


CONTRIBUTION TO THE EVALUATION OF A
MANAGEMENT TRAINING PROGRAM



by

MICHAEL IAN VALIQUET

B.A., Royal Military College of Canada
(1957)

B.Comm., University of Toronto
(1958)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF
SCIENCE
at the
MASSACHUSETTS INSTITUTE OF
TECHNOLOGY
1964

Signature of Author
Alfred P. Sloan School of Management,
November 23, 1964.

Certified by
Thesis Supervisor.

Accepted by
Chairman, Departmental Committee
on Graduate Students.

0-5

CONTRIBUTION TO THE EVALUATION OF A MANAGEMENT TRAINING PROGRAM

MICHAEL IAN VALIQUET

Submitted to the Alfred P. Sloan School of Management on November 23, 1964, in partial fulfillment of the requirements for the degree of Master of Science.

The purpose of this thesis is to evaluate one aspect of a specific internal development program conducted by the General Electric Company: the degree to which the values underlying the training objectives have been translated by participants into actual changed attitudes and behavior on the job. The stated program goals were improved ability in adapting to change, increased self-understanding and the capacity for open, meaningful working relationships - goals through which both collaboration and conflict could be rendered productive, ultimately enhancing the efficiency and profitability of the firm.

A random sample of 60 experimental subjects was chosen from lists of participants at four different company locations. A "matched-pair" control group was then obtained by asking each experimental to nominate another GE manager who had not participated in the training. The primary data source was an "open-ended perceived-change" questionnaire completed by five "describers" (peers, superiors, and subordinates) for each subject. Next, twenty-one criterion variables, inductively derived in earlier research, and considered sensitive to the attitudinal and behavioral change goals of the program, were imposed on the data to determine total change scores for each subject.

The result of the inquiry indicate that: (a) a significantly greater proportion of program participants than non-participants evidence increased change in behavior and attitudes toward the stated training values; (b) a significantly higher number of changes verified on the basis of concurrence between two or more describers was obtained for experimental subjects than controls; (c) the most important changes for participants occurred in the direction of improved democratic leadership, risk-taking, adaptability, and insight into self and role. Sub-analysis of the data on the basis of training location indicated that the program had been substantially more successful in one of the departments participating, and less successful in another. No significant difference was found in the overall ratings by class of describer (peer, superior, or subordinate) although notable differences occurred in certain content categories.

Thesis Adviser: Warren G. Bennis
Professor of Industrial Management

OTTAWA, Canada
November 20, 1964

Professor Philip Franklin
Secretary of the Faculty
Massachusetts Institute of Technology
Cambridge 39, Massachusetts

Dear Professor Franklin:

In accordance with the requirements for graduation, I herewith submit a thesis entitled "Contribution to the Evaluation of a Management Training Program".

I would like to express my sincere appreciation to my thesis committee members, Professor Warren G. Bennis and Professor David E. Berlew, for their advice and encouragement. Special thanks are due Professor Bennis, without whose original inspiration and enduring interest, this thesis would not have been possible, and to Professor Douglas Bunker of the Harvard Business School, who granted me liberal reference to and use of his research. Last, but not least, I would like to thank my wife, Helen, for handling the myriad details associated with a large research population.

Sincerely yours,

Michael I. Valiquet

CONTENTS

Chapter 1	- Introduction	p.1
Chapter 2	- The Evaluation of Training	p.8
Chapter 3	- The Business Effectiveness Program	p.20
Chapter 4	- Research Design and Methods	p.30
Chapter 5	- Results and Conclusion	p.55
Appendix A	- Experimental Questionnaire	p.76
Appendix B	- Describers' Questionnaire	p.80
Appendix C	- Controls' Questionnaire	p.82
Appendix D	- Inductively Derived Categories for Content Analysis	p.86
Appendix E	- Scoresheet for Mentions by Class of Subject and Content Category	p.88
Appendix F	- χ^2 Computation for Subject Choice of Describers	p.89
Appendix G	- χ^2 Comparison of Peers, Superiors, and Subordinates with respect to the Total Number of Changes in Category Sets "A" and "B" mentioned by Each Group	p.90
Appendix H	- χ^2 Comparison of Total Change by Location based on number of subjects' responses from each	p.91

TABLES AND DIAGRAMS

Conceptual Model of the Business Effectiveness Program	p.22
Program Schedule	p.24
Tables:	
1 Distribution of Experimental and Control Subjects According to Total-Change Scores	p.58
1A Distribution of Experimental and Control Subjects According to Self-Scores	p.59
2 Distribution of Experimental and Control Subjects With Respect to Change Scores in Category "A"	p.61
3 Distribution of Experimental and Control Subjects With Respect to Change Scores in Category "B"	p.61
4 Difference Between Experimental and Control Subjects on the Basis of "Verified" Changes	p.62
5 Difference Between Experimental and Control Subjects With Respect to Total Number of "Verified" Changes Expected for Each Group	p.64
6 Differences Between Experimental and Control Groups in Proportion of Subjects Reported as Changed	p.65
7 Total Changes Recorded for Each Criterion Variable by Class of Descriptor	p.70
8 Total and Average Change Scores for Experimental Subjects by Location	p.72

Restricted Distribution

This thesis has been given limited reproduction with the permission of the School of Industrial Management. However, this copy is issued with the understanding that none of the data resulting from this investigation will be used for advertising or publicity purposes, and that the thesis is solely for the confidential use of the organization or individual to whom it is addressed.

CHAPTER I

INTRODUCTION

"Without research and experimentation, education remains the only major profession without the guts to look at itself."

- Dean Francis Keppel¹
Harvard Graduate School of Education

Dean Keppel's indictment of the teaching profession is perhaps less true today than ever before in the history of education. Educators in virtually every avenue of society - university, industry and government - have in recent years demonstrated an increasing willingness to expose their theories and practices to rigorous, scientific investigation.² With respect to industry, much of this increased inquiry into training processes and outcomes has stemmed from the added impetus which management development has received in the post-war years and the concomitant desire to spend the "training dollar" in the most efficient manner possible.

1 Now Commissioner of Education, U.S. Office of Education

2 Allport, G. W., "How Shall We Evaluate Teaching?" in A Handbook For College Teachers, edited by Bernice Brown Cronkhite, Harvard University Press, Cambridge, 1950.

1.1 Purpose

The primary purpose of this thesis is to evaluate one aspect of a specific internal development program conducted by the General Electric Company: the degree to which the values underlying the program have been translated by participants into actual changed attitudes and behavior on the job. Since the basic hypothesis of training is that of transfer (the transmissibility of what is learned in the stimulus situation to the customary work situation) verification of the direction and extent of such transfer becomes an important link in any overall program evaluation.

In the particular training studied, the stated program goals were improved ability in adapting to change, increased self-understanding and the capacity for open, meaningful working relationships - goals through which both collaboration and conflict can be rendered productive, ultimately enhancing the efficiency and profitability of the firm. The major strategy thus becomes to develop a research design and methodology which will provide valid data for assessing the impact of the program on the attitudes and behavior of participants in terms of these training objectives. In brief, this involves selecting criteria which are at once measurable and relevant to the stated goals, developing adequate measuring devices,

and finally applying these instruments to data gathered with respect to the individual's on-the-job behavior and attitudes.

1.2 Research Design and Methods

A random sample of 60 experimental subjects was chosen from lists of participants in the program at four different company locations. These particular departments were chosen because in each of them the training had begun at least a year previously, giving the participants an opportunity to test and internalize the new values, attitudes, and behavioral skills in their working environment, and allowing the researcher to tap whatever residual effects have survived the state of early post-training euphoria as well as any erosive constraints in the environment.

A "matched-pair" control group was obtained by asking each experimental subject to nominate another manager in GE who held an identical or closely similar functional role to that of the participant, but who had not participated in the training. This permits the researcher to negate the effects of large discrepancies in ability or opportunity on the potential for change.

The primary data source was an "open-ended perceived-change" questionnaire completed by five "describers" for each subject. These "describers" were

chosen at random from seven to ten names submitted by each subject of people with whom the subject had worked closely for at least a year. Subjects were asked to select a variety of peers, superiors, and subordinates deemed qualified to assess the subject's on-the-job behavior and any perceived change in that behavior over the past year.

The open-ended question posed was: "Over a period of time people may change in the ways they work with other people. Do you believe that the person you are describing has changed his/her behavior in working with people over the last year as compared with the previous year in any specific ways?" A similar question was used to elicit responses from the subjects.

Content analysis of the sometimes generous verbal responses to the questionnaire required an objective method for counting and classifying the data to permit statistical analysis. To this end, twenty-one criterion variables or content categories, inductively derived in earlier research, and considered sensitive to the attitudinal and behavioral change goals of the program were imposed on the data. The scoring task involved assigning each specific mention of change to one of the twenty-one content categories, recording a score of "One" in each category in which there occurred one or more

mentions. The total change score for each subject was then obtained by adding the individual scores on each of the five responses. Finally, by combining all values (zero or one) in the matrix of categories and descriptors, a variety of scores were obtained.

1.3 Hypotheses Tested and Results

In keeping with the conceptual framework of the program, and if the training has been effective, we might expect the following results:

- (a) A significantly greater proportion of experimental subjects than controls showing greater total change toward program goals, as determined both from self-ratings and the ratings of others;
- (b) A significantly greater proportion of experimentals than controls demonstrating behavioral and attitudinal changes on the job;
- (c) Changes in behavior and attitudes among the subjects will tend to be varied and individual; that is, no one category of change tested will predominate all others. This hypothesis is based on the belief that the individual tends to assimilate training according to his own preferences and needs;

- (d) A significantly higher number of changes "verified" by concurrence of two or more describers for experimental subjects than control subjects.

In addition, it was felt of interest to obtain some measure of the varying impact of the training in each of the four company locations, and to test whether different types or frequency of change mentions occur with respect to the relationship between the subject and the describer (i.e.-whether superior, subordinate, or colleague). Also, the results were compared with those of Professor Douglas R. Bunker³, who performed a larger experiment of a similar nature.

The results of the inquiry indicate that:

- (a) A significantly greater proportion of program participants than non-participants evidence increased change in behavior and attitudes toward the stated training values as perceived by peers, superiors, and subordinates.
- (b) The behavior and attitude changes tend to be varied and individual, indicating perhaps, that information relevant to the achievement of individual and conjoint goals tends to be assimilated according to the individual's own preferences and needs.

³ Professor Douglas R. Bunker, "The Effect of Laboratory Education Upon Individual Behavior", National Training Laboratories Subscription Service, No. 4, 1963.

- (c) A significantly higher number of changes verified on the basis of concurrence between two or more describers was obtained for experimental subjects than controls.
- (d) The most important changes for program participants occurred in the direction of improved democratic leadership, risk-taking, adaptability, and insight into self and role.
- (e) Sub-analysis of the data on the basis of training location indicated that the training had been substantially more successful in one of the departments participating, and less successful in another.
- (f) No significant difference was found in the overall ratings by peers, superiors, and subordinates, although notable differences occurred in certain content categories.

The Evaluation of Training

2.1 The Newness of Evaluation

During the last two decades there has been a notable proliferation of time, effort and money invested in various forms of management training and development. Indoctrination and training programs conducted at various points in the manager's career; systematic job rotation involving changes in the nature of the functions performed (e.g., moving from production into sales) as well as changes in the individual's physical location and superiors; performance appraisal programs including various amounts of testing, general personality assessment, and counseling both within the organization and through the use of outside consultants; apprenticeships, systematic coaching, junior management boards, and various specialized projects and intern programs designed to facilitate practice by the young manager in functions he is to perform in later years; participation in specialized seminars, conferences, and training programs, including professional association meetings, human relations workshops, advanced management programs conducted in business schools or by professional associations, as well as regular academic courses at universities are among the many educational devices

used together with elaborate schemes of selection, appraisal, and placement to ensure an adequate inventory of capable management personnel.¹

The training described above is usually aimed at providing trainees with experiences which will allow them to appropriately modify their behavior, attitudes, values, skills and knowledge in a direction leading to more effective action in groups and organizations. Unfortunately, a large part of such educational investment is based on the assumption that training ipso facto changes the attitudes and behavior of individual managers so exposed, and that organizational improvement must ultimately follow. But too often, as Tannenbaum and others have pointed out, the blind faith of sponsoring management is involved.² Much of the management training and development carried on today is accepted on its face validity - the objectives appear desirable and the methods employed appear sound. But face validity alone is an inadequate criterion

1 For an even larger inventory of these activities, see E. H. Schein, "Management Development as a Process of Influence", reprinted in Industrial Management Review, May 1961, M.I.T.

2 Tannenbaum, R., Weschler, I., and Massarik, F., Leadership and Organization, McGraw-Hill, 1961, p.22

against which to measure the differential results of alternative training activities; only through the rigorous and objective measurement of such activities can management more fully assess the risks and benefits which attach to the various possible approaches, and hence, make more effective personnel decisions.

In recognition of this situation, there is evidence in more recent years of an increasing growth of interest in and demand for evaluation of management training and development. This new spirit of evaluation is captured well by Allport when he remarks:

"For those who are charged with the conduct of our social institutions the postwar period has turned into a period of self-examination. Evaluation is in the air. No longer do we assume complacently that our schools, churches, courts, social services, industries, communications, government - as now run - are equal to the tasks assigned to them."³

While it is difficult to demonstrate the expanded interest in training and development with statistics,⁴ few deny the

3 Allport, G.W., "How Shall We Evaluate Teaching?" in a Handbook for College Teachers, edited by B.B.Cronkhite, Harvard University Press, 1950.

4 Financial criteria can rarely be employed since expenditures on training are often debited to a variety of accounts and no accounting procedure is established for making such an overall estimate. Additionally, training departments in many firms carry on a variety of tasks not directly related to their normal training duties, which complicates the prospects of obtaining reliable estimates even further. Also, on-the-job training carried on in the line divisions on an ad hoc or semi-formal basis is both difficult to account for and to classify.

increasing tendency to regard training as a permanent and generalized function within the firm. And since training involves considerable investment, managers and employers are evincing considerable interest in whether the results justify the outlays.

Trainers themselves have been showing greater interest in formulating theories underlying their activities and in controlling and adapting their methods. Evaluation of the latter is essential to facilitate the choice of techniques appropriate to given needs; it constitutes both "feedback" to the trainer and is a method of "reflecting the image" of the training to the trainees themselves. The latter are often concerned with the psychological investment which may be required of them in certain types of training, particularly training "in depth" such as the T-Group method of sensitivity training. To dispel the myths and uncertainties connected with such training, practitioners are turning to rigorous, evaluative research.

In addition, trade unions and management associations, whose natural function is to make recommendations to their members concerning the propriety of, and conditions which should attach to various training systems, are demonstrating greater concern for training processes and outcomes.

Thus, a variety of parties, who hitherto have been forced daily to evaluate training on the basis of their own empirical or intuitive judgments, require a clearer understanding of the nature and extent of training results, an understanding which can come only from the careful, objective measurement and interpretation of those results.

2.2 The Concept of Evaluation

To evaluate something, in the strictest sense of the word, is to assess its value. But where the evaluation of training is concerned, the term takes on a somewhat broader meaning. Here it implies not only an assessment of the value of training results, but an attempt to discover the nature of the changes brought about by the training. Evaluation should also be differentiated from "validation". Any training which produces an intended change is considered valid; to validate training, therefore, is at most to establish that it tends to produce the desired results. Evaluation, on the other hand, attempts to elucidate even those changes which may have been unplanned.⁵

5 This point is well made in one statement of the primary problem in evaluation: "to provide objective, systematic, and comprehensive evidence on the degree to which the program achieves its intended objectives plus the degree to which it produces other unanticipated consequences, which when recognized would also be regarded as relevant to the agency." Hyman, H.H., Wright, C.R., and Hopkins, T.K., Applications of Methods of Evaluation, U. of California Press, 1962.

The latter distinction becomes important when we realize that each class of persons interested in training results tends to interpret them in terms of its own particular problems, values, and sets of expectations. Thus, the president will probably be interested in the overall increases in productivity associated with the training, the chief of personnel will look for a reduction in employee turnover, while the production manager may seek greater acceptance of his improvement programs. The staff trainer, or outside consultant, for his part, may interpret the results in terms of improved organizational climate and human relations within the firm. In view of these varied and diffuse expectations, it may be neither possible nor desirable to confine the evaluation to a limited number of a priori objectives. Moreover, the questions of who set the objectives and with what degree of precision further complicate the problem of determining after the fact what the objectives were to begin with. It is partly for this reason that in evaluating the specific program which was of interest in this thesis, no attempt has been made to differentiate between results due to the trainers' activities and those due to the on-going impact of the training on the organization itself. However, such differentiation is felt to be of crucial importance only where, for scientific purposes, one is attempting to isolate the effects of the training in

an effort to compare their conformity with well established standards.⁶ Otherwise, in the interests of greater objectivity and ease of analysis, it seems better to make an overall evaluation of what has been achieved through the integrated use of various forms of action: training, organizational action, and structural change.⁷

2.3 Difficulties in Training Evaluation

While the concept of evaluation, even in the broader sense defined above, is fairly straightforward in the stating, the procedures can be involved and fraught with difficulty. Objectives must be defined operationally so that they can be measured; criteria for the evaluation must be selected and measuring instruments for these developed; and, finally, procedures for the measurement and comparison of results

6 A distinction is sometimes drawn between substantive evaluation and procedural evaluation. The latter is chiefly concerned with "the conformity of practice to certain established standards considered essential to the achievement of desired effects", whereas substantive evaluation is concerned with the "net" effects of the training activity. It is the latter which is the chief concern of this thesis. See further Mahoney, T.A., Jerdee, T.H., and Korman, A., "An Experimental Evaluation of Management Development", Personnel Psychology Vol 13, No.1, Spring 1960, p.83.

7 Evaluation of Supervisory and Management Training Methods, Organization for Economic Co-Operation and Development Publications, Paris, France, No. 15,091, June 1963, p.37

developed. Each of these steps is critical to the usefulness and validity of the evaluation, and as will become apparent in what follows, subject to many pitfalls, and what Miller has aptly described as the "innate cussedness of things" in research on the human object.⁸

The difficulty in defining objectives has already been alluded to in 2.2 above. In the first place, there may be a significant discrepancy between management's perception of the a priori objectives of the training, and that of the "outside" consultant-trainer. The latter may feel, for example, that it is top management that needs training in re-examining its fundamental attitudes, rather than employing stop-gap measures at lower levels to remedy its internal dysfunctionings. But, in the belief that management may not yet be "ready" to receive this information, the consultant may find it necessary to compromise by formulating his objectives in vague classical concepts and terminology which are bound to satisfy management, but which permit the consultant to pursue his objectives. While this ambiguous state of affairs is conducive to neither a satisfactory training milieu nor good conditions for evaluating training results, it may be a

8 F. Miller, "Resistentialism in Applied Social Research", Human Organization, 12:5-8, 1954

necessary transitional state in arriving at a level of human relations within the firm which will allow the clarification of common objectives in an unclouded atmosphere, free of secret strategies.

The situation outlined above is compounded in the case of the staff trainer who is a member of the firm. The dissonant perceptions of management and himself as to the nature of the problem may also require compromise, but for obvious reasons his is an even more precarious cakewalk - requiring as it so often does, flexibility, hidden persuasion and even "shell game" tactics with respect to training objectives. Many of his problems derive from the fairly universal phenomenon in which every hierarchical level of management feels that it is the level immediately below it which is most in need of training. Thus, the more timorous staff trainer may be forced to address his initial efforts to the relative "fringe" areas of the problem, while his more intrepid colleague uses influence to start at a higher level, although risking thereby a reputation as a "pawn of management".

Another source of difficulty in the evaluation of training arises from the nature of training itself. Specifically, training employs as its medium an increase in insight into one's way of

doing or being (behavior and attitudes), and into the meaning of these. Thus, a pertinent criterion in evaluating any training activity is the increase in insight or perception as well as the changed behavior of the individual after the training. But training can have a dual effect on the trainee: one due to the explicit or rational content of the training (i.e., training for a specific objective) and the other due to its symbolic content (e.g., the trainee may feel he has been placed on the program because "somebody up there likes me", or, conversely, that management is dissatisfied with his performance and is using the training to give him one last chance to produce "or else"). Both these effects are likely to be operant with varying degrees of intensity in any management training activity.⁹ We are thus led to ask, in the case where the symbolic content is the preponderant effect, whether the training is not merely an illusion, since the causal relationships are unclear, and the requirements of increased insight has not been used as the attitude-modifying agent. However, as was

9 For a fuller explanation of this phenomenon see Evaluation of Supervisory and Management Training Methods, op.cit., pp. 34-35

argued with respect to differentiating among the various forms of organizational change, perhaps the distinction is essential only where one is attempting to isolate the pure training effects, and is irrelevant to an overall evaluation of a program where many forms of action are at play.

A further class of problems arising in evaluative studies stems from the nature of the social and psychological sciences themselves. In the first place, whenever an effort is made to evaluate the results of training activity, the methodological problems inherent in any attempt to study human phenomena are invariably present. Specifically, and despite valiant efforts on the part of researchers to avoid it, it is almost certain that the response of a subject will be modified by the fact that he knows he is being observed; and so the threat of the notorious "Hawthorne" effect is ever imminent. Even where elaborate control groups are set up to provide the researcher with some measure of command over extraneous variables, the control group may become contaminated by communication from the experimental group (if they are drawn from the same organization), and thus not yield a fair estimate of changes owing to "zero" or "nontreatment". Insulation from such vicarious influence is a difficult, if not

impossible, condition to meet in programs involving organized groups of people in close interaction.

The foregoing far from exhausts the plethora of problems besetting the evaluator. A host of extraneous factors, concurrent with the program or set in motion by the evaluation itself, may influence the gross changes observed. Hence, the evaluation must be designed in such a way as to eliminate these factors if it is to demonstrate the exclusive impact of the program. The world of events, the normal process of adult growth, the inherent instability that measuring instruments show upon repeated use, and the practice effect from sheer retesting are but a few of the myriad factors which may complicate the measurement of the criterion variables. While it is beyond the scope of this thesis to examine these in detail, further reference will be made later to such problems as actually impinge on the evaluative study at hand.

CHAPTER 3

The Business Effectiveness Program¹

One of the first, and often, slighted tasks facing the researcher in any evaluative study is to adequately describe the environment in which the training has taken place. Frequently the program is described in terms of physical time and space - the where, when, how often and how many of the activity. But only rarely do such descriptions go beyond physical facts into the realm of psychological and sociological variables, even though the latter are of primary importance where the training may involve changes in attitudes, behavior, or values. Again, since the effects of the activity are being measured in terms of such variables, it seems logical that the program should be described in corresponding terms. In addition to the customary time-space parameters, therefore, an effort is made here to present

¹ The bulk of the material presented in this chapter is based on two internal papers concerning the G.E. Business Effectiveness Program: "The Business Effectiveness Program Research: Progress and Prospects" by Warren G. Bennis, M.I.T., 1963 and "Business Effectiveness Work" by H. D. Meader, Business Effectiveness Consulting Service, August 22, 1963.

the psychological and sociological variables of the stimulus situation.

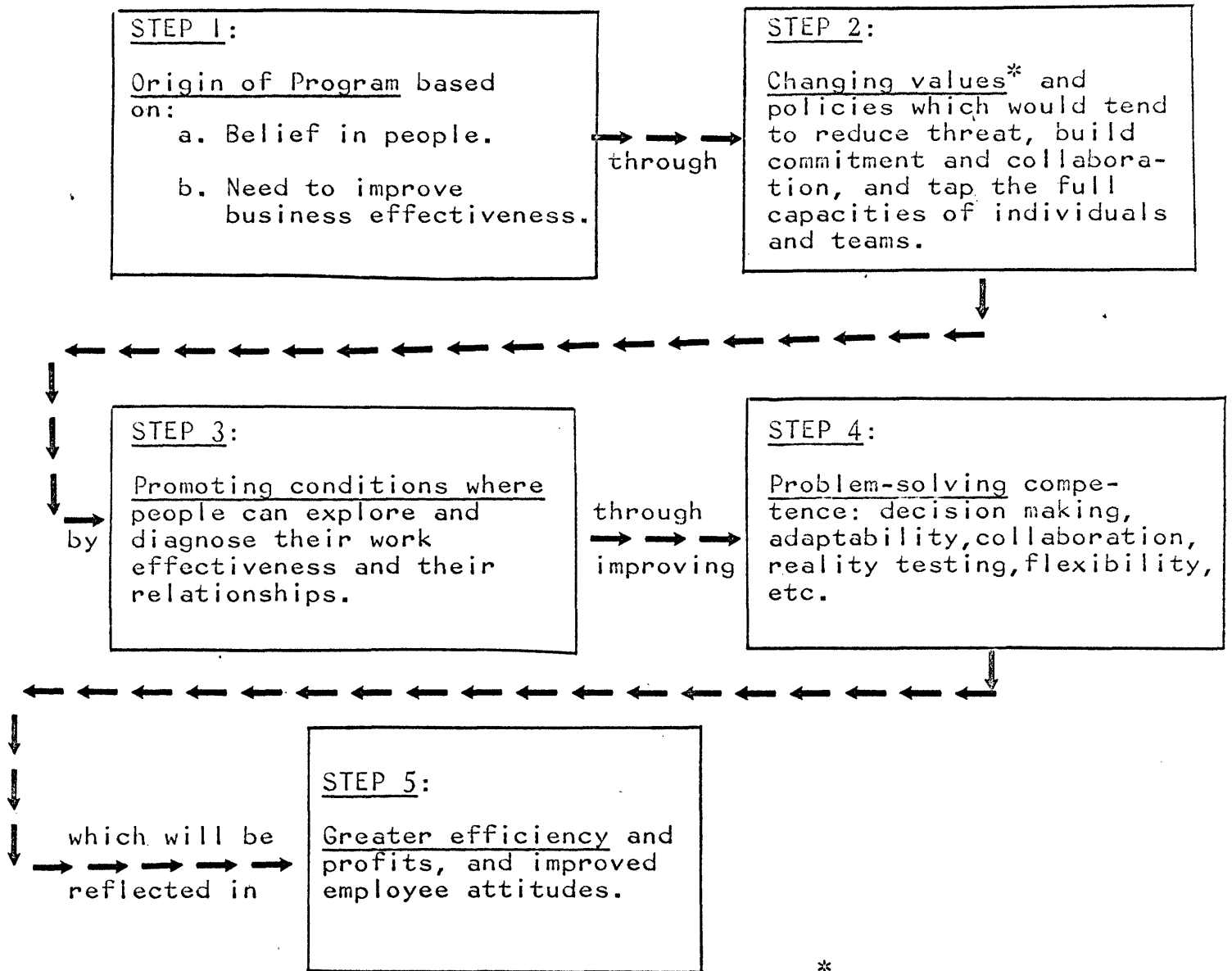
The Business Effectiveness Program is an organizational change program instituted by the General Electric Company in response to a variety of needs. Some of these were induced by society, some by the technical and environmental demands of the company itself, and some by a new vision of leadership and motivation designed to cope with an era of rapid industrial and social change and to develop new and better forms of human collaboration.

The conceptual framework of the B.E.P. (see Figure 1) was designed to promote conditions where managers could "explore, diagnose, and take remedial action with respect to their work effectiveness and their relationships"², and thereby more fully utilizing their talent and energies. As such, the program involves the identification of the basic causes of ineffectiveness, the determination of the changes required to correct this ineffectiveness and the impact of such changes on G.E. personnel, the development of the requisite competencies to make the change, and, finally, the actual implementation of the changes

2 Bennis, op.cit., p.1.

FIGURE 1

CONCEPTUAL MODEL OF BUSINESS EFFECTIVENESS PROGRAM



*Such as openness, consensus, management of conflict, self-control, distribution of influence.

on a controlled schedule for specific positions and teams in such a way as to build employee understanding, commitment and collaboration in the changes.

Operationally the program comprised three phases: analysis, planning and decision-making, and the implementation of planned changes. Before describing these in greater detail, certain characteristics of the B.E.P. which distinguish it from other "change programs" should be emphasized. Firstly, the change agents involved in effecting the desired changes in effectiveness were drawn from within in the company, rather than from outside consultants, as is so often the case. Secondly, these change agents or "catalysts", as they have been called, were not professional behavioral scientists and, therefore, were able to commence their consulting process without the benefit or detriment (whichever the case may be) of the behavioral scientist stereotype. Thirdly, the catalysts themselves received a uniform 10-week training program in "change-agentry". Finally, the benefits of using change agents well-versed in the technical skills and organizational milieus in which they were to operate was considered to more than offset the problems of the trainer who is a staff member of a firm, as discussed in Chapter 2.

It was envisaged that a trained catalyst would work with the various management teams in a department of the company for a period of about one year. During this time he would act as a consultant, trainer and "role-model" for management by "living" the values of the program. The approximate schedule of the catalyst's activities in any one department is outlined below:

	<u>Approx. No. of Weeks</u>
(1) Analysis & Diagnosis	
-Top Management & throughout department	6
-Top Management Meeting	1
-Sub-Section Managers	8
-Multi-Functional Sub-Section Manager Meeting	1
(2) Planning & Decision-Making	5
-Functional team Manager Meetings and Preparations	10
(3) Implementation	<u>15</u>
TOTAL	<u>46 weeks</u>

PHASE I - ANALYSIS

The purpose of the analytical phase is to develop employee motivation through understanding and involvement. This phase of the program, in turn, may be divided into three steps:

1. Diagnosis

The catalyst, working with the managers at all levels in the department, attempts to examine factual information concerning current performance levels, major weaknesses in such operating procedures as production loading, manpower planning, and work measurement, and managerial attitudes regarding policies, objectives, problem-solving, decision-making, productivity, performance, organizational factors, motivation and human relationships.

2. Four-day Meetings

The second step in the analytical phase is a four-day meeting involving the general manager and section managers held at a site other than the department location to avoid the distractions of routine operations. At the outset of this relatively unstructured meeting in which the managers themselves set the agenda, the General Manager initiates the meeting by emphasizing the need for employee effectiveness and by attempting to create a climate wherein frank and open discussion of problems commonly withheld or ignored can take place. Throughout these discussions the catalyst provides a common "frame of reference" against which the managers can identify and develop a list of symptoms of ineffectiveness and their underlying causes, explore

the consequences of their management behavior on others, and become more sensitive to their own needs and those of others in their business relationships. One of the decisions customarily made is to carry on similar meetings at the Sub-Section Manager level.

3. Sub-Section Manager Meetings

The third step in the analytical phase involved sub-section manager discussions similar to those carried on with the general manager and section managers. These meetings are structured to provide multi-functional participation in groups of 8 to 12.

PHASE II - PLANNING AND DECISION-MAKING

Following the sub-section meetings, a series of one or two day functional team meetings held at, or near, the department location initiate Phase II. Here individual work groups are given the opportunity to integrate their skills so that they can function more effectively as teams. Discussions flow naturally from the results of the previous meetings and focus on planning change projects to improve procedures, standards, plant loading, relationships, etc.

PHASE III - IMPLEMENTATION

Implementation of the changes planned by the functional teams takes place on a manageable, project by project basis over the period of a year or more.

As explained above, the training environment must be described in terms of the significant psychological and sociological variables as well as the physical. To this end, and with training goals of changed values and improved work performance, it is apparent that any research used to determine the effectiveness of the program must be concerned with two types of factors or variables. These shall be referred to as "output" factors and "behavioral" factors.

By "output" factors is meant such typical business variables as productive efficiency, sales volume, growth rate, profit rate, etc., which are sought for their own economic value and are usually considered as having trade-off value in relation to each other. The "behavioral" or dynamic variables, as they are sometimes called, include such factors as communication effectiveness, problem-solving ability, interpersonal competence, pride in firm, group cohesiveness, level of performance goals, amount of control, delegation, mutual trust, and so on.

The importance of both classes of variables for understanding the training activity and its effects may become clearer with the following example. One may wish to measure the effects of a certain training program on the time devoted to exercising controls over production, etc., which represents the objective of the training. This could be done by direct measurement of the ultimate objective at different periods (e.g., before training, just after training, or several months after training). But this very same time variation may reflect a profound change in such behavioral variables as superior-subordinate relationships, anxiety reduction, etc. - dynamic factors which may themselves be directly related to the training process, but which will remain hidden if the ultimate "output" variable is used as the sole criterion of evaluation. Thus, if one really wants to understand how the training works, both behavioral and output variables must be brought to the fore. In cognizance of this fact, the total research on the subject program is designed to measure the effects of training on a variety of behavioral and output factors - problem-solving competence, values, attitudes, structural change, performance, employee turnover, absenteeism, and so on.

The present study, for its part, however, focuses solely on the effect of the program on individual change, i.e., the degree to which the concepts and values underlying the training have been translated by participants into actual changed attitudes and behavior on the job. Since the basic hypothesis of training is that of "transfer" - the transmissibility of what is learned in the training situation to the usual on-the-job situation - verification of the direction and extent of such transfer becomes an important link in any overall evaluation of the program.

In view of this aim, this study is confined to the measurement of such attitudinal and behavioral variables as communication, risk-taking, self-control, delegation, self-confidence, sensitivity to group behavior and the feelings of others, and insight into own behavior with respect to the job. Each of these will be defined and discussed in greater detail in the ensuing chapter. Suffice it to say here that these factors seem intuitively to encompass the behavioral goals of the B.E.P. - openness, consensus, management of conflict, self control, distribution of influence, etc. - goals which hopefully would tend to reduce threat, build commitment and collaboration, and tap the full capacities of individuals and teams, thereby improving organizational effectiveness.

CHAPTER 4

Research Design and Methods

As noted previously, while management and academic interest in training evaluation has gained added impetus in recent years, scant attention has been paid to the longer range consequences of training in the work environment. Not least of the reasons for this is the fact that whenever an effort is made to evaluate the results of training activity, the methodological problems inherent in any attempt to study human phenomena are also present. Many of these problems have been outlined in Chapter 2. Others which almost invariably creep in are the difficulty of obtaining comparable control subjects against which to measure the change, if any, in the experimental subjects; the necessity of selecting criteria which are at once measurable and operationally meaningful; and the difficulty of trying to separate the pure training effects from the "normal", or base rate of change and growth effects to which virtually all people are subject.

The research design and methodology used here attempts, as far as possible, to deal with these problems in a way which will provide valid data for assessing the impact of the B.E.P. on the attitudes and behavior of management where the program has been

introduced. Basic elements of this design are adapted from two similar studies.¹

4.1 Characteristics of the Sample

A random sample of 60 experimental subjects was chosen from lists of participants in the B.E.P. at four different locations: Electronic Specialty Capacitor Product Section, Irmo, S.C.; the Medium Transformer Department, Rome, Ga.; the Hanford Atomic Products Operation, Richland, Wash.; and the Chemical Processing Department, Richland, Wash. These particular departments were chosen because in each of them the B.E.P. had begun at least a year previously. The sample comprised managers from the General Manager to Sub-Section Manager levels and was distributed over the four locations roughly equally.

4.2 Matched-Pair Controls

A "matched-pair" control group was obtained by asking each experimental subject to nominate another manager in G.E. who held an identical or closely similar

1 Miles, Matthew B., "Human Relations Training: Processes & Outcome", Journal of Counselling Psychology vol.7, No. 4 (1960) pp 301-306.

Bunker, Douglas R., "The Effect of Laboratory Education Upon Individual Behavior", National Training Laboratories Subscription Service, No. 4, 1963.

functional role to that of the participant, but who had not participated in the B.E.P. This allows the researcher to negate the possibility of large discrepancies in ability or opportunity affecting the potential for change.

4.3 The Questionnaire

The primary data source was an "open-ended perceived-change" questionnaire completed by five "describers" for each experimental and each matched control subject. The five "describers" were chosen at random from seven to ten names (submitted by each experimental and control subject) of people with whom the subject had worked closely for more than a year, i.e., peers, superiors and subordinates, who were deemed qualified to assess the subject's on-the-job behavior and any change perceived in that behavior over the past year.

The open-ended question took the following form:

"Over a period of time people may change in the ways they work with other people. Do you believe that the person you are describing has changed his/her behavior in working with people over the last year as compared with the previous year in any specific ways? Yes _____ No _____ If yes, please describe."

A similar question was used to elicit descriptions of their own behavior from the subjects themselves.

The major strategy then becomes to assess the types of changes making up each subject's "total-change score" (the total change perceived by the describers) and see whether any significant difference exists between the experimental and control groups. The method of assessing the change and arriving at a "total-change score" is described below under "Criterion Development".

4.4 Design for Data Collection

Ordinarily, management and researcher alike must be wary of "a posteriori traces", or, after-the-fact measures of training effectiveness, because the lapsing of time permits many intervening variables to change, thereby contaminating the results. But since the B.E.P. is an on-going or continuous type of training program, this very time lag works more to our advantage than disadvantage. Firstly, because the focus of the study is change on-the-job, participants must have had the opportunity to test and internalize the new values, attitudes and behavioral skills in their working environment, the so-called "incubation" period of the training. Thus, by collecting data a year or more

after the training has begun, we are able to tap whatever durable effects have survived the waning of early training enthusiasm, the process of selection and incubation, and the erosive effects of constraints in the working environment.

The question thus arises whether describer bias induced by awareness of the subject's participation or non-participation in the B.E.P. would affect the results. A corollary question is whether the describers could easily contrive a set of situationally-relevant behavior changes and whether they would be motivated to do so. These questions will be dealt with in greater detail in the discussion of results, but for the moment it can be noted that judging from the high number of "no change" as well as positive change responses obtained from describers, a tentative "no" answer is proffered. Also, the high number of verified changes (changes confirmed by concurrence among the reports of two or more describers), the fact that the 5 describers of each subject were chosen at random from a list of seven to ten possibles, and the discriminatory power of the large number of change categories into which the data were classified, all yield confidence in the objectivity of the describer responses.

Again, since the B.E.P. is an on-going activity involving both training and organizational change, and since the object of the research is to assess the overall effectiveness of the program on individual change, rather than what could be attributed to the training alone, it seemed reasonable to ignore such factors as relocations and promotions, changes in organization structure, and participation in other training which probably operated with mixed effects on the criterion for both experimental and control subjects.

4.5 Criterion Selection²

The large volume of verbal material contained in the responses to the open-ended questionnaire required an objective method for classifying and counting the responses to permit statistical comparison. The content categories chosen for classifying the responses involved 21 constructs which had been inductively derived in an earlier study by Bunker and which had been found to be both organizationally relevant and personally meaningful to a large number of respondents from diverse organizational settings. (See Appendix "D")

2 The author is indebted to Professor Bunker for permission to use the scoring system developed and tested by he and his colleagues at the National Training Laboratories.

These constructs or variables seemed intuitively to encompass the attitudinal and behavioral change goals of the B.E.P. - openness, consensus, management of conflict, self control, distribution of influence, etc.- goals which hopefully would tend to reduce threat, build commitment and collaboration, and tap the full capacities of individuals and teams, thereby improving organizational effectiveness. Moreover, having been inductively derived and tested over a lengthy period in a broad spectrum of organizational settings, it was felt that these constructs could be most confidently imposed on our heterogeneous research population.

The scoring task thus involved assigning each mention of a specific change to one of the 21 content categories. But before examining this procedure further, the descriptive specifications of the 21 categories used in classifying the data should be defined in greater detail to illustrate some of the finer distinctions which have been made. While some of these may seem tenuous and arbitrary, they were nevertheless deemed necessary in the interests of easier scoring and comparison with Professor Bunker's results. The content categories fall into three comprehensive sets: (A) overt changes in behavior, (B) inferred changes in insight and attitude, and (C) global or non-specific changes.

The descriptions scored in the "A" categories tend to refer to overt behavior or action - the subject does more or less of something, or acts more often in a particular way. Set "B", on the other hand, includes responses based on first-order inferences as to the internal state of the subject based on behavioral cues (e.g., attributions of insight or sensitivity to group processes.) The "C" category set is used as an empirical dustbin for all gross inferences concerning character and non-specific change references. It should be noted that while precise category fit, according to the scoring conventions to be described below, was demanded for category sets "A" and "B", less rigid requirements were imposed on set "C" on the grounds that it comprised changes which were either doubtful or irrelevant to the research task. With this lone qualification, then, scoring in each category depends upon an explicit statement of qualitative or quantitative difference in the subject's on-the-job behavior or attitudes. Changes may be positive or negative reflecting increases or decreases in quantity and greater or lesser utility.

A. OVERT BEHAVIORAL CHANGES

A-1. Communication

Under the heading of "communication" there can be three kinds of changes: sending, receiving and an unspecified change encompassing both. To qualify for a change in this category, a change mentioned by a describer must specify a change in the quantity and/or quality of the subject's behavior as a sender or as a receiver (or both) of communication. Wherever possible, a differentiation is made between sending or receiving as opposed to an unspecified change statement.

Statements concerning the subject's "sending" behavior are usually fairly explicit in that the subject is usually perceived as communicating something to someone in an improved manner; he is seen as expressing himself oftener or better. Similarly, changes in receiving communication are explicit in that they involve quantitative or qualitative changes in how the subject listens to or perceives messages. A rather difficult distinction must be made between this category sub-heading and that of B-5, "Tolerance of New Information", described below. The latter category deals more directly with what the subject listens to, i.e., his evaluation or selection of what he will "hear", rather than how he receives the

communication, which is the act of listening. Nevertheless, since they are so inextricably related, we would expect a significant correlation between the two.

The "unspecified" statement of change in communication is one which does not explicitly discern between sending and receiving, and may, in fact, comprehend both. An example of a response falling into this subcategory would be, "He communicates better with his subordinates". Here we cannot readily ascertain in which of the two aspects the subject has improved. If, however, both are mentioned, they are scored separately.

Certain statements relating to communication, however, carry more than "improved communication" as their intended message, especially when read within their full context. For example, the statement: "His contributions are more helpful in our work group", implies more than ameliorated communication on the part of the subject; it implies that his "contributions" fulfill a specific function in the group, and that they stem from a diagnosis of the situation. Such statements are therefore more correctly categorized under "Functional Flexibility". (See category A-5 below).

A-2. Relational Facility

Changes in the subject's behavior which make him easier to get along with, evidence that he has improved his relations with others are scored as "relational facility". Evidence of relational facility may be found in such statements as "irritates co-workers less", "increased ability to work with people"; "more pleasant and sociable", "easier to talk with", "more cooperative in relations with others", and so on. While the actions here described seem fairly straightforward, some difficult distinctions must be made to distinguish them from closely allied categories. For example, statements which refer to self control or deliberate control of emotions are coded under A-6, "Self-Control", even though they make the subject easier to get along with.

One of the more onerous distinctions to be made is between Relational Facility and Tolerance of New Information. Thus a statement such as, "more open to the opinions of others", seems equally acceptable as an expression of improved relationships or as evidence of an attitude toward others' opinions. How, then, should it be classified? The distinction must be made on the basis of whether the change is "attitudinal" or "actional". Is the statement about something the subject is, or something he does? It seems likely that being

"more open to the opinions of others", while certainly facilitating relationships, expresses attitudinal rather than operational change, and is therefore more reasonably classified under B-5, "Tolerance of New Information".

General statements referring to deliberate control of emotions, or self-control, are coded under category A-6, "Self-Control", even though this quality may be readily construed as making the subject easier to get along with.

A-3. Risk-Taking

This variable pertains to changes in the subject's overt behavior which indicate that where he was formerly passive or protective, he now seems more willing to expose himself to uncertainty or opposition.³ Statements falling into this category would be: "more willing to take a stand on issues"; "more willing to be opposed"; "more willing to try new ideas"; "experiments more"; "more assertive"; "more creative solutions to

³ It should be emphasized that for this variable, as well as for all others which were evaluated, the direction of change need not be positive; the scoring system accounted for negative change in the criterion variables as well.

problems"; etc. These statements carry as their main theme the extension of one's self, experimentation, or the freedom to act in new areas. While a high correlation between this variable and "Self-Confidence" (B-6 below) might be expected, "Risk-Taking" is the operational or actional demonstration of the attitude of confidence. Again, where a describer's response is specifically related to the concept of adaptability, of action as a result of diagnosis, or of action as the result of a perceived need, it implies more than mere risk-taking and is consequently coded as "Functional Flexibility". (See A-5 below)

A-4. Increased Interdependence

Statements falling in this category are those which express or imply an increased effort on the part of the subject to create a democratic environment, i.e., to encourage the sharing of responsibility and the participation of subordinates in decision-making processes. As with the other variables in Category "A", the statement must express an action on the part of the subject, rather than an attitude. Examples of such statements are: "encourages participation in decision-making"; "lets others do more thinking and experimenting"; "involves others more"; "gives greater leeway to subordinates"; "less autocratic"; "less likely to dominate a discussion".

Because of the comparative ease with which responses can be fitted into this category, care must be taken that certain statements are correctly placed. For example, "more willing to share information" seems intuitively to allude to the sharing of responsibility or participation in decisions. However, its main theme is "Communication" and for this reason would be scored in the more relevant category. For a statement to qualify as "increased Interdependence" it must have as its central theme encouraging participation, the sharing of responsibility, or the improvement of democratic action and involvement in work situations.

A-5. Functional Flexibility

This category comprehends statements specifically related to the concept of adaptability, i.e., action as the result of diagnosis. Both aspects of the concept must be present: the diagnosis, which may be implied (in terms of perceived needs of individuals, groups, or situations), and the action, which must be explicit, to alter or fulfill the perceived needs. The major criterion for admission of a statement to this category is evidence that a subject is better able to change his behavior to suit a changed situation, that is, he shows a flexibility and adaptability

in his interactions with his environment. The distinguishing feature is the subject's ability to perceive or diagnose a situation, and respond to it on the basis of such perception or diagnosis.

Examples are "more helpful to team members"; "finds it easier to accept change, and is more flexible in groups"; "takes group roles more easily"; "gives more help to his staff with personal and work problems"; "contributions in group more helpful".

A-6. Self-Control

Unlike "Risk-Taking", which implies the release of checks and stops, where the latter have been perceived as dysfunctional, "Self-Control" implies an increase in these controls in such a way that the subject becomes more functional and improves his relations with others. Statements expressing behavior changes which imply a former excessive expression of emotion, judgments or ideas, and indicate that the subject is presently controlling these to better advantage, are placed in this category. Examples are: "checks feelings more carefully"; "less quick with his judgment"; "more self-discipline".

Several difficult and, perhaps, tenuous distinctions are made with respect to this variable. Firstly, many statements may seem to imply controls

without actually mentioning them, as in "he irritates co-workers less" or "seems more tactful in meetings". These statements, however, do not refer to specific acts of self-control; they describe other actions which merely suggest certain controls may be in operation, and are therefore scored in more immediately obvious categories.

Another problem arises with such words as "patient" and "calm", which are often used in ways implying actions of self-control but, in fact, are referring to an attitude toward others. Thus, the statement "more patient with others" carries a strong implication of control, but is more correctly classified as an attitude, "Acceptance of Others" (B-4 below). Similarly, "more calm" implies some measure of self-control, but again seems more reasonably viewed as an attitudinal change rather than an actional one, and so fits more accurately into category B-7, "Comfort". On the other hand, a statement such as "progress in developing the ability to remain calm" seems to emphasize the action which a subject takes to maintain a "calm" attitude and would consequently be scored as "Self-Control" rather than "Comfort".

While the distinction is perhaps tenuous statements scored in this category are often distinguishable by being negative, either in their grammatical form

or in their message, i.e., the subject improves through the negation of some undesirable characteristic.

B. Inferred Attitudinal Changes

B-1. Awareness of Human Behavior

Statements falling into this category concern the subject's intellectual understanding of other people's behavior. In contrast to what is referred to as "sensitivity" in various other categories below, "awareness" implies only the conscious understanding and detached intellectual digestion of behavior, as opposed to the subjective involvement connoted by sensitivity. This concept is usually triggered by such words as "aware", "appreciate", "conscious", "analytic" and "insight". To be so scored, both the process (the intellectualization, awareness, analysis, or insight) and the object of the process (human beings - their behavior, actions or needs) must be present in the statement. Also included in this category are statements in which the awareness pertains ostensibly to only specific individuals or groups of individuals, as in "more insight into reasons for his subordinates' actions" and "more aware of individual behavior in groups".

B-2 Sensitivity to Group Behavior

This category comprises increased understanding or perception of the group as an entity - its

behavior, roles, and needs. The key word is "group"; without it the statement must be classified elsewhere. If mention or strong implication is made of awareness of individual behavior, it is classified under B-1 above. Some illustrative statements of increased sensitivity to group behavior are: "more analytical of group processes"; "better able to perceive needed roles in the group"; "more aware of sub-currents in the group"; "more conscious of group processes".

B-3 Sensitivity to Others' Feelings

This variable implies an attitude of greater empathy with or compassion for other people, as in "more sensitive to others' feelings", "more sensitive to the needs of others", "more capacity for understanding feelings", "more conscious of the problems of others". For scoring purposes, a distinction is made between this category and B-1, "Awareness of Human Behavior", with respect to the process involved in each. The point has already been made above, but bears repeating. If a detached intellectual understanding is implied, the statement is scored under B-1. If, on the other hand, the needs or feelings of others are "sensed" sympathetically, i.e., the subject seems personally involved, then the statement is scored here.

B-4. Acceptance of Other People

Included in this category are mentions of increased consideration of people as individuals, acceptance of them as they are. Not included are statements of any action displaying this attitude. Thus, a response such as "he gives his subordinates more responsibility" can certainly be seen as characteristic of greater acceptance of individual differences, but it is concerned primarily with action, with an outward display of the attitude, rather than with the attitude per se. It would therefore be scored as A-4, Increased Interdependence. On the other hand, statements such as "more able to tolerate shortcomings" and "more considerate of individual differences" are concerned with the attitude of the subject, rather than the action he takes, and so are scored in this category.

B-5. Tolerance of New Information

This variable deals with the subject's attitude toward the ideas and opinions of others. Two types of statements are generally anticipated. The first is one that states directly that the subject is better able to accept the opinions of others, e.g., "more able to see the other person's point of view"; "more willing to accept the recommendations of others"; "takes our ideas into consideration more". The other

type of statement assigned to this category, admittedly, for lack of a better fit elsewhere, is one in which the openness to new information is merely implied. Thus, the statement "I have tried to think more objectively and dispassionately", which at first blush seems unsuitable for this category, upon further analysis yields a different conclusion. The first part of the sentence ("have tried to think...") indicates that the change is attitudinal rather than action. The latter portion ("...more objectively and dispassionately"), or its equivalent form, "less subjectively and emotionally", implies a reaction to something - specifically, the views and opinions of others.

The crucial distinction between "action" and "attitude" so frequently made in the scoring of other variables applies equally here. The operational act of "listening" must be differentiated from the closely related attitude which supports better listening; a statement such as "increased effort to listen better" must not be confused with "tries to understand others' opinions more". The former deals with the act of listening (A-I-R Communication), while the latter is concerned more with the attendant attitude, i.e., increased tolerance and admittance of others' views.

Statements apparently encompassing both concepts must be studied with care; for example, "listens with greater understanding", while perhaps descriptive of the subject's attitude, would be scored as Communication on the basis of the governing theme "listens".

B-6. Self-Confidence

Included here are all specific references to the confidence or security of the subject, e.g., "more poised and confident"; "leads discussions more confidently"; "more sure of himself"; "feels more secure in his position", etc.

B-7. Comfort

This variable is concerned with apparent anxiety reduction in the subject; that is, where he was once nervous or insecure in specific situations, he is now seen as coping with them. Such statements as "more relaxed", "more calm", "more comfortable in meetings", "stands up to pressure better", "not as tense as he used to be", all indicate decreases in anxiety, insecurity or tension and are correctly placed in this category.

B-8. Insight Into Self and Role

This category deals with the way the subject perceives himself and/or his job. It is therefore considered one of the more important criteria

of the program, impinging as it does so directly on the individual's work effectiveness. Indications of improved insight into either self or role would be: "better acceptance of job challenge"; "shows more respect for task-oriented jobs"; "more aware of his own assets and limitations"; "better grasp of his place in the organization"; "closer identification with the leadership role"; "has left his drawing board and started to manage"; "seems more aware of the increased scope of our operation"; and so on.

C. Global Changes

The five categories included in this set arise from definitional requirements rather than any intrinsic importance with respect to program goals. They serve as a "catchall" for the many non-specific and marginally scorable descriptions of vague behavioral changes and gross changes in character which may be received from describers. Much of this is probably due to a well-known research phenomenon. When respondents are asked to accommodate a researcher and wish to oblige, but have no specific or concrete behavioral referents on which to base their replies, they tend to furnish vague and global descriptions. In addition, it is likely that these global descriptions incorporate a large part of the "normal" or base rate of change and growth evident in most everyone.

C-1. Miscellaneous and Broad Changes

Statements descriptive of this category are "better outlook on life"; "broader view of the human scene"; "more realistic in his approach to problems"; "now has a broader background to support his intuitive thinking", and so on.

C-2. Character Changes

These encompass such global changes in character as "more mature", "more human", "more sensitive", "shows greater sense of responsibility", etc.

C-3. Miscellaneous Operational Changes

This classification includes all non-specific statements of operational change which cannot be fitted into the "A" category set - "makes better use of consultant services", "uses broader concepts", "better prepared to deal with problems", etc.

C-4. Use of Special Skills and Tools

Statements exemplifying this category are: "more positive methods of keeping people on the subject"; "applies Business Effectiveness Program Skills to the training of his own staff"; "shows improved technique in his planning discussions".

C-5. Consequences of Change

Describer responses stating nebulous results of change or, where the change itself is not readily identifiable as to direction or intent, are placed in this category. For example, "staff-line relationships are improved", "makes us feel more secure in our work situation", "he spends more time in his office and less time at our desks", are all statements which indicate results of apparent change, but tracing the causal factors would require sheer guesswork. They are therefore relegated to the global set in an attempt to preserve the reliability and discriminating power of the more pertinent categories.

4.6 Scoring

Turning now to the scoring task, for each describer response a score of one was assigned for each category in which there was one or more mentions. The total-change score for each subject was then obtained by adding up the individual scores on each of the five describers' responses. Next, by combining all values (zero or one) in the matrix of categories and describers for all subjects, a variety of change scores can be obtained, e.g., the total-change score for experimental subjects vis-a-vis controls; separate scores for category sets A, B, and C, for self-ratings as opposed to describer ratings, for different experimental groups

by location; and a "verified" change score developed by counting the number of observations on a particular subject where two or more describers occur.

The completed questionnaires were stripped of group identification to ensure a blind-process and thereby avoid the "halo" effect referred to in Chapter 2. Also, they were independently re-coded to check drift in the use of the categories.

CHAPTER 5

Results and Conclusion

5.1 Questionnaire Response

The rate of return on mailed questionnaires is notoriously low and the resultant problem of subject self-selection can be extremely destructive to otherwise well-conceived research strategies. In the particular research design used in this study the difficulty was compounded by the fact that the researcher was dependent on the good graces of the experimental group to provide an adequate control group and a sufficient number of describers. The simplicity of the questionnaire, accompanying explanatory letters and numerous reminders combined to give decent response statistics. Of the random sample of 60 experimental subjects who were originally canvassed, 39 or 65 percent replied, and of the 31 control subjects provided by these, 18 or 58 percent returned questionnaires.¹

¹ These figures compare favorably with Bunker's experience of roughly 75 percent and 66 percent return on the original experimental and control groups respectively. One problem with which Bunker did not have to contend was the frequent omission of control names from experimentals' returns on the grounds that all managers in the individual's immediate work area had participated in the program.

After others were eliminated because of recent promotions and relocations, or lack of describer response, 49 or 54 percent of both the original experimental and control groups were included in the study. Of approximately 450 describers who received questionnaires over fifty percent returned usable responses. In light of these returns, while not as astounding as Bunker's experience with an even larger research population, subject self-selection can be eliminated as an important source of error variance with some confidence.

5.2 Change Scores

Combining cell values (zero or one) in the matrix of categories and describers for each subject (see Appendix "E") a variety of scores were obtained. The most comprehensive of these were the total-change score based on the matrix sum and the "verified" change score mentioned in the last chapter. Separate change scores were also obtained for category sets A, B, and C, for subject self-ratings as opposed to describer ratings, for different experimental groups by location, and for subject ratings by class of describer (i.e., peers, superiors and subordinates).

5.3 Group Comparisons²

The initial comparison of experimental and control subjects is made on the basis of total-change scores. This analysis, presented in Table I below with the distribution of the total research population divided into thirds,³ indicates that a significantly greater proportion of experimental subjects (participants) than controls (non-participants) were in the middle and top thirds of the distribution of change scores. The probability of obtaining a value of chi-square as large as 11.37 if the two groups were not different is less than .001. While the value of chi-square is substantially smaller than that obtained by Bunker, the results are comparable at his stated level of significance.

-
- 2 Wherever possible in the ensuing analysis, comparison will be made with Professor Bunker's results, highlighting any apparent differences. See Bunker, D.R., op.cit.
 - 3 This procedure has been followed for the purposes of comparison with Professor Bunker's experiment. Cell values have been combined to permit use of the chi-square test. To provide a more stringent check on such a small research population (N = 49) and cell values, the "exact" procedure for calculating chi-square was employed where necessary and found not to reverse any decisions concerning the null hypotheses. See Croston and Cowden, Applied General Statistics, Prentice-Hall, 1955, p.686.

TABLE 1

DISTRIBUTION OF EXPERIMENTAL AND CONTROL
SUBJECTS ACCORDING TO TOTAL-CHANGE SCORES.

	<u>EXPERIMENTAL</u> <u>SUBJECTS</u>	<u>CONTROL</u> <u>SUBJECTS</u>	<u>TOTAL</u>
UPPER TWO-THIRDS (Total Scores: 4 to 17)	28	5	33
LOWER THIRD (Total Scores: -1 to 3)	<u>6</u>	<u>10</u>	<u>16</u>
N	34	15	49

$$\chi^2 = 11.37$$

$$(d.f. = 1)$$

$$P(\chi^2 = 11.37) < .001$$

When the test is repeated on the basis of self-scores, the difference is significant only to the .01 level (Table IA). This may merely reflect the normal propensity of people for an enhanced self-image. Since the consequent lack of objectivity could "muddy" the results, not too much stock is placed in them.

TABLE IA

DISTRIBUTION OF EXPERIMENTAL AND
CONTROL SUBJECTS ACCORDING
TO SELF-SCORES

	<u>EXPERIMENTAL</u> <u>SUBJECTS</u>	<u>CONTROL</u> <u>SUBJECTS</u>	<u>TOTAL</u>
UPPER HALF (total scores: 2 to 4)	25	5	30
LOWER HALF (total scores: -1 to 1)	13	13	26
	<hr style="width: 50%; margin: 0 auto;"/> 38	<hr style="width: 50%; margin: 0 auto;"/> 18	<hr style="width: 50%; margin: 0 auto;"/> 56

$$\chi^2 = 7.09$$

$$P(\chi^2 = 7.09) < .01$$

$$d.f. = 1$$

When this same type of analysis is executed solely on the basis of the changes recorded by describers in category set "A", the results are (coincidentally) identical (see Table 2). This indicates that participants in the training program demonstrated significantly greater behavioral change toward program goals than did non-participants. When this procedure is repeated independently for the attitudinal category set (see Table 3), a similar pattern occurs, although this time only to the .01 level of significance. A somewhat post hoc interpretation of this result is that the participants, while not necessarily "sold on the management philosophy advanced during their training sessions, nevertheless feel free to experiment with the new modes of behavior. This would account for the higher frequency of actional as opposed to attitudinal change, and such a supposition would be supported by the significantly higher level of "Risk-Taking" observed in experimental subjects (see Table 6 below). Only when the experimental control comparison is extended to the "global" change data (i.e., category set "C") does no significant difference between the two groups occur.⁴ This empirical result lends credence to the hypothesis advanced in the last chapter that the "C" category changes constitute an important component of the "normal" or "base" rate of change observable in nearly all human beings, and emphasizes the discriminating power of the other change categories.

4 Of the total research population, five experimental and nine control subjects were described as having changed along these lines. The two-cell table yields an insignificant chi-square value of 1.14.

TABLE 2
DISTRIBUTION OF EXPERIMENTAL AND CONTROL SUBJECTS
WITH RESPECT TO CHANGE SCORES IN CATEGORY "A".

	<u>EXPERIMENTALS</u>	<u>CONTROLS</u>	<u>TOTAL</u>
UPPER TWO-THIRDS (change scores: 0 and 1)	28	5	32
LOWER THIRD (change scores: 2 - 10)	6	10	17
	34	15	49
N			

$$\chi^2 = 11.37$$

$$(d.f. = 1)$$

$$P(\chi^2 = 11.37) < .001$$

TABLE 3
DISTRIBUTION OF EXPERIMENTAL AND CONTROL SUBJECTS
WITH RESPECT TO CHANGE SCORES IN CATEGORY "B".

	<u>EXPERIMENTALS</u>	<u>CONTROLS</u>	<u>TOTAL</u>
UPPER TWO-THIRDS	22	4	26
LOWER THIRD	12	11	23
	34	15	49
N			

$$\chi^2 = 7.07$$

$$(d.f. = 1)$$

$$P(\chi^2 = 7.07) < .01$$

Lest the indications up to this point of the analysis be deemed more artifactual than substantive, the raw data are now subjected to a more demanding test. Table 4 presents the difference between experimental and control groups on the basis of "verified" change scores, i.e., where a subject had one or more specific observations of change confirmed by concurrence among the responses of two or more describers.

TABLE 4

DIFFERENCE BETWEEN EXPERIMENTAL AND CONTROL
SUBJECTS ON THE BASIS OF "VERIFIED" CHANGES

	<u>EXPERIMENTALS</u>	<u>CONTROLS</u>	<u>TOTALS</u>
One or more changes verified by describer concurrence	25 (73%)	2 (13%)	27
No Verification	<u>9</u> (27%)	<u>13</u> (87%)	<u>22</u>
N	34	15	49

$$\chi^2 = 15.22$$

$$(d.f. = 1)$$

$$P(\chi^2 = 15.22) < .001$$

In Chapter 4 the questions arose as to whether describer bias induced by awareness of the subject's participation or non-participation in the program would affect the results, whether the describers could easily contrive a set of situationally-relevant behavioral changes, and whether they indeed would be motivated to do so. While a tentative negative answer was advanced on the basis of the high number of "no change" as well as positive change responses received from describers, the random process of describer selection, and the discriminating power of the many change categories, the Table 4 results permits even greater confidence in the objectivity of the majority of describer responses. Over two-thirds of the experimentals, as compared with less than one-third of the controls, had one or more specific changes confirmed by the reports of two or more describers. This difference in the proportion of verifications adds substance to the interpretations made of Tables 1, 2 and 3, and compares closely with findings of Bunker.

Table 5 presents the comparison of experimental and control groups with respect to the number of verified changes expected and observed.

The expectation is based on the total number of changes mentioned for all subjects in each group.

TABLE 5

DIFFERENCE BETWEEN EXPERIMENTAL AND CONTROL
SUBJECTS WITH RESPECT TO TOTAL NUMBER OF
VERIFIED CHANGES EXPECTED FOR EACH GROUP

<u>GROUP</u>	<u>OBSERVED FREQUENCY</u>	<u>EXPECTED FREQUENCY</u>	<u>$\frac{(O-E)^2}{E}$</u>	<u>χ^2</u>
EXPERIMENTAL	48	41	1.19	4.79
CONTROL	4	11	3.60	

$$P(\chi^2 = 4.79) < .05 \quad (\text{Two-tailed test})$$

Here we find that verified changes occur more frequently for experimentals than controls even when our expectation is based on the distribution of total changes among the two groups. And we must concur with Bunker's findings that "these data permit us to place some confidence in the pooled observations of the several observers for each subject and further indicate that the verified-change score is an even more powerful discriminator between experimentals and controls than the total change score."⁵

5 op.cit., p.12

Finally, in Table 6 is presented an analysis by scoring category of the differences in proportions of subjects reported as changed in each of the two groups.

TABLE 6

DIFFERENCES BETWEEN EXPERIMENTAL AND CONTROL GROUPS
IN PROPORTION OF SUBJECTS REPORTED AS CHANGED

<u>SCORING CATEGORY</u>	<u>PROPORTION PERCEIVED AS CHANGED</u>		<u>DIFFERENCES</u>
	<u>EXPERIMENTALS</u>	<u>CONTROLS</u>	
A-1 Sending	.2941	.0667	.2274*
A-1 Receiving	.1765	.0667	.1098
A-2 Relational Facility	.3529	.1333	.2196*
A-3 Risk-Taking	.3235	.0667	.2568**
A-4 Increased Interdependence	.5588	.1333	.4255**
A-5 Functional Flexibility	.3529	.0667	.2862**
A-6 Self Control	.3529	.1333	.2196*
B-1 Awareness of Behavior	.1765	.0667	.1098
B-2 Sensitivity to Group Process	.1176	.0667	.0509
B-3 Sensitivity to Others' Feelings	.3235	.1333	.1902
B-4 Acceptance of Others	.3235	.0667	.2568*
B-5 Tolerance of New Information	.3235	.1333	.1902
B-6 Confidence	.1176	.0667	.0509
B-7 Comfort	.2059	.0667	.1392
B-8 Insight into Self and Role	.4412	.1333	.3079**

** $P < .01$

* $P < .05$

Some interesting similarities and differences with respect to Professor Bunker's results are brought to light. The

latter found that seven of the eight categories in set "B" discriminated at the .05 level or better, while only four of the seven categories in set "A" had comparable power. In the present case the situation is somewhat reversed with some interesting qualitative differences. Six of the eight behavioral categories discriminate at the .05 level or better, while only two attitudinal categories demonstrated similar power. Bunker found no significant difference in his experimental and control groups with respect to "Risk-Taking", while it is found here to be significant at the .01 level. This, perhaps, may be explained by the fact that Bunker's experimental group was exposed to a stimulus situation of shorter duration and remote from its operational environment, whereas the subject program was relatively continuous and carried out for the most part within the operational environment. What is implied here is that gradual and protracted training within a generally supportive atmosphere may better lend itself to experimentation and exposing one's self and ideas, than returning from a "cultural island" to the erosive effects of a constricted environment. Too, there is the fact that the Business Effectiveness training was accompanied by other forms of change (e.g., organizational, structural) which may have reinforced a spirit of experimentation.

Another significant divergence from Bunker's results lies in the discriminating power of Category A-5, "Functional Flexibility". Where the latter again discovered this factor to be relatively insignificant, it is here found to be of major importance. Reasoning similar to that advanced for the high incidence of "Risk-Taking" appears relevant. Recalling that a score for "Functional Flexibility" requires evidence of both diagnosis and the ability to act to fulfill perceived needs, it seems reasonable to expect higher incidence of the second criterion in an organizational climate which has been carefully nurtured at all management levels to accept the reality of change, and the necessity of adapting to it.

The findings with respect to Increased Interdependence are similar to those of Bunker. In both cases, the variable is found to distinguish between program participants and non-participants at the .01 level. This result may reflect both the cognitive preferences of describers as well as the relevance of the variable to the organization.

The one other qualitative difference which bears mention is the significant increase in "Insight into Self and Role" found in program participants. Where

the NTL participants in Bunker's study were distinguishable from controls at only the .05 level, the result in the present study is significant at the .01 level. Again it may be that an improved perception of one's own attributes and limitations, particularly with respect to the job, is more readily acquired in the actual working environment, rather than as has until recently been supposed, through social interaction with strangers (rather than "family" groups) at some secluded retreat.

5.4 Analysis of Change With Respect to Subject-Descriptor Relationships

It was thought to be of interest to examine the type and frequency of change recorded for experimental subjects with respect to the type of describer - i.e., whether for peers, superiors or subordinates. The thirty-one subjects used in the analysis were found to have chosen 26 superiors, 53 subordinates and 96 peers as their describers. Since subjects were asked to choose a mixture of peers, superiors and subordinates as their describers, the chi-square test was used to determine whether the describer mix actually chosen differed significantly from an expectation based on the choice of one superior, two peers and two subordinates. (See Appendix "F") No such significant divergence was found.

When this type of analysis is extended for subjects from the four company locations making up the

research population, however, subjects from Department "A" are found to have substituted a significantly greater number of peers for subordinates than was expected. A possible interpretation of this result would be that the training had been less successful at this location, so that subjects felt somewhat more threatened by the prospect of subordinate ratings than peer ratings. This possibility will be pursued further in the sub-analysis by location (5.5 below).

The chi-square analysis was again applied to see whether there exists any significant difference among the total change scores attributable to each class of describer for criterion variable sets "A" and "B". (See Appendix "G"). While no significant difference was found, in each category as a whole, examination of the "type" of change most frequently noted by each class (Table 7) yields some interesting results.

The largest category scored by subordinates proves to be "Increased Interdependence" (A-4). This is as might be expected since the variable involved is one which impinges directly on the largest segment of a subordinate's working environment. It is a truism that most subordinates consider themselves, perhaps justifiably,

the best judges of their superior's management style. Peers, on the other hand, score Relational Facility and insight into Self and Role more frequently than other variables. The higher incidence of scores for relational facility could be construed as indicating that subjects have improved their relations with peers more than they have with their superiors and subordinates. In turn, this may indicate a lack of confidence in applying new found behavioral skills vertically, so that experimentation with them is still being restricted to the less risky peer level.

TABLE 7

TOTAL CHANGES RECORDED FOR EACH CRITERION VARIABLE
BY CLASS OF DESCRIBER

	<u>PEERS</u>	<u>SUPERIORS</u>	<u>SUBORDINATES</u>
A-1-S	4	3	5
A-1-R	3	1	3
A-2	13	3	5
A-3	7	4	3
A-4	9	7	13
A-5	10	2	7
A-6	8	4	6
B-1	2	3	2
B-2	1	-	1
B-3	4	3	3
B-4	7	3	2
B-5	6	4	5
B-6	3	2	1
B-7	3	1	4
B-8	13	6	5

The equally high occurrence of scores for Insight into Self and Role may be due to a natural human tendency to use equals as behavioral self-referents. Also, having for the most part taken the same training, peer-describers may have been more sensitive to this type of attitudinal change.

5.5 Analysis of Change With Respect to Company Location

As mentioned previously, when the subjects' choice of describers was studied on the basis of company location, it was found that Department "A" had chosen a significantly larger number of peers than any other class of describer.⁶ One possible explanation offered was that the training at this location may have been less successful, so that the selection of peers instead of superiors or subordinates was deemed less threatening. Such an interpretation would be supported by the fact that substantially fewer subjects from Department "A" responded to the questionnaire than from any of the others, and is further reinforced by perusal of Table 8.

⁶ Locations have been disguised as A, B, C, and D on the grounds that such information should be confidential to the company.

TABLE 8

TOTAL AND AVERAGE CHANGE SCORES FOR EXPERIMENTAL
SUBJECTS BY LOCATION

<u>DEPARTMENT</u>	<u>TOTAL CHANGE SCORE</u>	<u>NO. OF S RESPONDING</u>	<u>AVERAGE SCORE</u>	<u>NO. OF DESCRIBERS</u>
A	28	6	4.6	30
B	26	5*	5.2	25
C	57	12	4.8	53
D	98	11	8.9	49

$$\chi^2 = 16.76 \text{ (d.f. = 3) } P(\chi^2 = 16.76) < .001$$

*Four other subjects were eliminated due to zero describer response.

Here it is evident that Department "A" had a lower average change score than any other location. Further, when chi-square is computed for the difference in total scores based on the distribution of subject response, the value obtained is significant beyond the .001 level, with locations "A" and "C" contributing most to the negative aspect of the divergence from expected values. (See Appendix "H").

Department "D", on the other hand, demonstrates a significantly high average score and contributes most to the value of chi-square. The inference here is that change in the direction of program goals was most successful at this location.

5.6 Conclusions

Based on the foregoing, the following conclusions seem in order:

1. A significantly greater proportion of program participants than non-participants evidence greater change in behavior and attitudes toward the stated training values as perceived by peers, superiors and subordinates.
2. Referring again to Table 6, we find that only the most popular scoring category (A-4) includes more than half of the experimental subjects. This indicates, perhaps, that information relevant to the achievement of individual and conjoint goals tends to be assimilated according to the individual's needs and preferences.
3. A significantly higher number of changes verified on the basis of concurrence between two or more describers was obtained for program participants than non-participants.
4. The greatest change in the experimental group occurred in the direction of improved democratic leadership or participative management, risk-taking, adaptability, and insight into self and role.

5. Sub-analysis of the data on the basis of training location indicated that the training had been substantially more successful in one of the departments participating, and less successful in another.
6. No significant difference was found in the overall ratings by peers, superiors, and subordinates, although notable differences occurred in certain content categories.

From the above it would appear that the "Action research" format of the GE program faces up to the phenomenon of "resistance to change" so well-documented by behavioral scientists. When it has been suggested to an individual either explicitly or implicitly that he needs to change his managerial attitudes and behavior, his sense of identity and perceived status with respect to others in the organization is threatened. Thus, as Schien points out,⁷ since attitudes are generally organized and integrated about the individual's self-image, and result in stabilized, characteristic modes of behavior toward others, any suggestion of the need for change implies not only some "criticism of the person's image of himself, but also threatens the stability of his

7 Schien, op.cit.

working relationships because change at this level implies that the expectations which others have about him will be upset, thus requiring the development of new relationships." It is for this reason, he goes on to explain, that training of this kind commonly arouses resistance, or at best, produces only temporary change as the expectations of co-workers tend to return the individual to his original mold.

The program studied in this thesis seems to have been able to thwart these psychological resistances, particularly the waning of the training "honeymoon" so much in evidence in programs otherwise conducted. This is not to say that all participants learned to accelerate their adaptive processes and obtain better control over them. To hope for as much would be quixotic. The data does indicate, however, that the majority of participants were able to integrate their newly acquired attitudes and behavioral skills into their personality and ongoing work relationships.

Experimental Questionnaire

SUBJECT: Evaluation Research Project

As part of an M.I.T. Research Project designed to increase understanding of organizational effectiveness, we are conducting inquiries to get accurate descriptions of the way people actually behave on the job. We believe that this research can be of great help to your Company in evaluating its efforts to date and in making plans for future programs.

Specifically, we are trying to study the effectiveness of General Electric's "Business Effectiveness Program". To do this we are comparing the changes that take place in the people who took part in the program to the changes that take place in people who didn't have an opportunity to take part in the program. To keep these two groups as equal as possible, we asked each participant to name a person who is in a position similar to his but who has not taken part in the B.E.P. You are one of these people named.

The following questionnaire is an attempt to find out in what ways you have changed your behavior in working with and managing people. We would like you to be as frank and objective as possible. Please be assured that your response and those of others taking part in this study shall be kept in strict confidence, and that in any report of research results anonymity will be preserved.

The last two pages of the enclosed material explain why we are asking you to supply the names of others in your organization who can help with the research project.

We hope that you will be able to help us complete this study. As we are working with a fairly small sample, your reply is extremely important to the value and success of the project. Previous experience with the attached forms has shown that it takes about 20 minutes to fill them out.

Thank you for your help and cooperation.

Appreciatively,

Michael Valiquet
Research Assistant

BEHAVIOR DESCRIPTION QUESTIONNAIRE

1. Over a period of time people may change in the ways that they work with other people. During the last year do you believe you have changed your behavior in working with people in any specific ways?

_____ Yes _____ No

2. If yes, please describe:

3. Job Title: _____

4. Have you had the same job and held the same position since January of 1963? If no, please explain:

5. In the last two years, have you attended any management training courses, or participated in activities similar to the Business Effectiveness Program? If yes, please explain:

6. Birth Date: ____/____/____

SUBJECT: Additional Persons to Help with Research

Now that you've completed the questionnaire, and can see what it is like, we would like to ask your assistance in locating additional persons to help.

The questionnaire was designed to measure the way you see yourself in working with and managing people. An equally important element for thorough research is the way others see you managing people. Accordingly, we would like to be able to get responses from people who work closely with you. They would fill out the same questionnaire you have just finished, but respond in terms of the way they see you. They will be your 'describers'. It is up to them to describe your on-the-job behavior as carefully and as objectively as possible.

We do not make this request lightly. It very important to get concrete, accurate descriptions of managerial behavior, both from you and the people who work with you.

If you are willing to help us with this, please list on the next page the people with whom you work most directly. These may include subordinates, superiors, and peers in your working organization, or others you work with regularly. Our experience has shown that the most accurate descriptions come from describers who have worked with the people they are describing for more than a year. It is better to include a variety. The lists should have five to seven names and addresses -- more, if you prefer.

From this list, we will draw a random sample of describers and mail them the questionnaire, asking them if they are willing to fill it out. Note that you need not do anything beyond supplying the names and addresses -- whether they agree to help or not is up to them. Using this method, you will not know who the actual describers are. We also are going to take steps so that when the describers' questionnaires come in to us, we will not know who they are from. In other words, names will be unimportant; what we will look at are the answers.

We appreciate your help on this very much indeed. If there are any questions this raises in your mind, please feel free to write us.

TO: M.I.T. Evaluation Research Project

FROM: _____ Date _____

The following list of names and addresses include the people with whom I work most directly. You may select a random sample of names and ask them to serve as describers.

1. _____ 5. _____

2. _____ 6. _____

3. _____ 7. _____

4. _____

Describers' Questionnaire

SUBJECT: Evaluation Research Project

As part of an M.I.T. Research Project designed to increase understanding of organizational effectiveness, we are conducting inquiries to get accurate descriptions of the way people actually behave on the job. We believe that this research can be of great help to your Company in evaluating its efforts to date and in making plans for future programs.

_____ is a subject in this study. In other words, he has answered various questions about the way he works with and manages people in an on-the-job situation. In the interest of research, he has suggested you and several other of his co-workers as people qualified to evaluate his on-the-job behavior. We at M.I.T. are not concerned with your personal evaluation as such, but we are interested in how groups of these evaluations compare with others. Your evaluation combined with many others is a way of forming this comparison. His name appears on the questionnaire so that we can group the descriptions of the several people who will be evaluating him. We hope that you will be as frank and objective as possible.

In the analysis of these questionnaires, there will be no mention of names; only general conclusions will be drawn about your Company's organizational effectiveness programs. You will notice that there is no place for your name on the questionnaire; this is so that your response will remain anonymous. We are asking, however, that you send in a separate postcard at the same time you send back this questionnaire so we can know you completed reply is represented. Both the stamped envelope and the postcard accompany this letter.

The questionnaires are a vital part of this research project, and as we are working with a fairly small sample, your reply is extremely important to its value and success.

Thank you for your help and cooperation.

Appreciatively,

Michael Valiquet
Research Assistant

BEHAVIOR DESCRIPTION QUESTIONNAIRE

1. Over a period of time people may change in the ways they work with other people. Do you believe that the person you are describing has changed his/her behavior in working with people over the last year as compared with the previous year in any specific ways?

Yes No

2. If YES, please describe:

3. About how frequently have you had contact with this person?

All Day Several times a day Once a day

4. In what capacity do you have contact with this person? (Subordinate, superior, or peer)

It is important that we now have the name of the person you are describing and your birth date. We will use the latter as a coding device to distinguish your report from the others. This will in no way link your description with your name.

5. The person I am describing is: _____

6. My birth date is: / /
Month Day Year

Controls' Questionnaire

SUBJECT: Evaluation Research Project

As part of an M.I.T. Research Project designed to increase understanding of organizational effectiveness, we are conducting inquiries to get accurate descriptions of the way people actually behave on the job. We believe that this research can be of great help to your Company in evaluating its efforts to date and in making plans for future programs.

Specifically, we are trying to study the effectiveness of General Electric's "Business Effectiveness Program". To do this we are comparing the changes that take place in the people who took part in the program to the changes that take place in people who didn't have an opportunity to take part in the program. To keep these two groups as equal as possible, we asked each participant to name a person who is in a position similar to his but who has not taken part in the B.E.P. You are one of these people named.

The following questionnaire is an attempt to find out in what ways you have changed your behavior in working with and managing people. We would like you to be as frank and objective as possible. Please be assured that your response and those of others taking part in this study shall be kept in strict confidence, and that in any report of research results anonymity will be preserved.

The last two pages of the enclosed material explain why we are asking you to supply the names of others in your organization who can help with the research project.

We hope that you will be able to help us complete this study. As we are working with a fairly small sample, your reply is extremely important to the value and success of the project. Previous experience with the attached forms has shown that it takes about 20 minutes to fill them out.

Thank you for your help and cooperation.

Appreciatively,

Michael Valiquet
Research Assistant

BEHAVIOR DESCRIPTION QUESTIONNAIRE

1. Over a period of time people may change in the ways that they work with other people. During the last year do you believe you have changed your behavior in working with people in any specific ways?

_____ Yes _____ No

2. If yes, please describe:

3. Job Title: _____

4. Have you had the same job and held the same position since January of 1963? If no, please explain:

5. In the last two years, have you attended any management training courses, or participated in activities similar to the Business Effectiveness Program? If yes, please explain:

6. Birth Date: _____/_____/_____

SUBJECT: Additional Persons to Help with Research

Now that you've completed the questionnaire, and can see what it is like, we would like to ask your assistance in locating additional persons to help.

The questionnaire was designed to measure the way you see yourself in working with and managing people. An equally important element for thorough research is the way others see you managing people. Accordingly, we would like to be able to get responses from people who work closely with you. They would fill out the same questionnaire you have just finished, but respond in terms of the way they see you. They will be your "describers". It is up to them to describe your on-the-job behavior as carefully and as objectively as possible.

We do not make this request lightly. It is very important to get concrete, accurate descriptions of managerial behavior, both from you and the people who work with you.

If you are willing to help us with this, please list on the next page the people with whom you work most directly. These may include subordinates, superiors, and peers in your working organization, or others you work with regularly. Our experience has shown that the most accurate descriptions come from describers who have worked with the people they are describing for more than a year. It is better to include a variety. The lists should have five to seven names and addresses -- more, if you prefer.

From this list, we will draw a random sample of describers and mail them the questionnaire, asking them if they are willing to fill it out. Note that you need not do anything beyond supplying the names and addresses -- whether they agree to help or not is up to them. Using this method, you will not know who the actual describers are. We also are going to take steps so that when the describers' questionnaires come in to us, we will not know who they are from. In other words, names will be unimportant; what we will look at are the answers.

We appreciate your help on this very much indeed. If there are any questions this raises in your mind, please feel free to write us.

TO: M.I.T. Evaluation Research Project

FROM: _____ Date _____

The following list of names and addresses include the people with whom I work most directly. You may select a random sample of names and ask them to serve as describers.

1. _____

5. _____

2. _____

6. _____

3. _____

7. _____

4. _____

Inductively Derived Categories for Content Analysis*

A. Overt Operational Changes - Descriptive

1. Communication

S. Sending - shares information, expresses feelings, puts ideas across

R. Receiving - more effort to understand, attentive listening, understands

2. Relational Facility - cooperative, tactful, less irritating, easier to deal with, able to negotiate

3. Risk-Taking - willing to take stand, less inhibited, experiments more

4. Increased Interdependence - encourages participation, involves others, greater leeway to subordinated, less dominating, lets others think

5. Functional Flexibility - more flexible, takes group roles more easily, goes out of way, contributions more helpful, less rigid

6. Self-control - more self-discipline, less quick with judgment, checks temper

B. Inferred Changes in Insight and Attitudes

1. Awareness of Human Behavior (intellectual comprehension) more conscious of why people act, more analytic of others' actions, clear perceptions of people

2. Sensitivity to Group Behavior - more conscious of group process, aware of subcurrents in groups

3. Sensitivity to Others' Feelings - more capacity for understanding feelings, more sensitive to needs of others

4. Acceptance of Other People - able to tolerate shortcomings, considerate of individual differences, patient

*

Scoring depends upon an explicit statement of qualitative or quantitative difference. Changes may be positive or negative reflecting increases or decreases in quantity and greater or lesser utility. Precise category fit according to scoring conventions is required for sets of categories A and B.

5. Tolerance of New Information - willing to accept suggestions, considers new points of view, less dogmatic, less arbitrary
 6. Self-Confidence
 7. Comfort - relaxed, at ease (must be specific as to setting or activity)
 8. Insight into Self and Role - understands job demands, more aware of own behavior, better adjusted to job
- C. Global Judgments - gross characterological inferences, non-comparable references to special applications of learning, and references to consequences of change

Computation of χ^2 for Subject Choice of Describers

Type of Describer	f_o	f_e 1:2:2	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
Superior	26	31	- 5	25	.81
Peer	76	62	+14	196	3.16
Subordinate	53	62	- 9	81	1.31
Total	155	155	0		5.27

$$P(\chi^2 \geq 5.27) > .05$$

Chi-square Values for Subject Choice of Describers by Location

	"A"	"B"	"C"	"D"
Superior	5	3	9	13
Peer	20	8	28	18
Subordinate	5	14	16	18
Total	30	25	53	49
χ^2	9.58	2.80	3.69	1.07
$P(\chi^2)$	< .01	> .05	> .05	> .05

Chi-square Comparison of Peers, Superiors, and Subordinates with respect to the Total Number of Changes in Category Set "A" Mentioned by Each Group.

Class of Describer	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	χ^2
Superiors	24	21	+ 3	9	.43
Peers	54	59	- 5	25	.42
Subordinates	43	41	+ 2	4	.10
Total	121	121	0		.95
			d.f. = 2	$P(\chi^2 = .95) >$.05

Chi-square Comparison of Peers, Superiors, and Subordinates with respect to the Total Number of Changes in Category Set "B" Mentioned by Each Group.

Class of Describer	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	χ^2
Superiors	22	14	8	64	4.57
Peers	38	41	- 3	9	.22
Subordinates	23	28	- 5	25	.89
Total	83	83	0		5.68
			d.f. = 2	$P(\chi^2 = 5.68) >$.05

1 Expectation is based on overall distribution of describers: 76 Peers, 26 Superiors, and 53 Subordinates.

Chi-square Comparison of Total Change by
Location based on number of subjects'
responses from each.

Department	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	χ^2
A	28	37.6	- 9.6	95.04	2.53
B	26	31.4	- 5.4	29.16	.93
C	57	73.	-16.0	256.	3.50
D	98	67.	+31.0	961.	9.80
Total	209	209.0	0		16.76

$$P(\chi^2 = 16.76) < .001$$

$$d.f. = 3$$

BIBLIOGRAPHY

- Andrews, Kenneth R.: "Is management training effective?"
Harvard Business Review - Jan-Feb 1957
and Mar-Apr 1957.
- Anshen, Melvin: "Executive Development: In Company vs.
University Programs"
Harvard Business Review - Sept-Oct 1954.
- _____ : "Better Use of Executive Development
Programs" Harvard Business Review -
Nov-Dec 1955.
- Argyris, Chris.: Interpersonal Competence & Organizational
Effectiveness. Homewood, Illinois; Irwin-
Dorsey, 1962.
- _____ : "Top management dilemma: company needs
vs. individual development", Personnel:
Sept.55.
- _____ : "Executive Leadership: an appraisal of
a manager in action" Harper 1953.
- _____ : "The Organization: what makes it healthy?"
Harvard Business Review, Nov-Dec 1958.
- Bennett, Willard E.: "Master Plan For Management Development"
Harvard Business Review - May-June 1956.
- Bourgeois, M.: "Contribution to Evaluation of an
Executive Development Program" M.S. thesis,
M.I.T.-1960.
- Bradford, L.,
Gibb, J., and
Benne, K., (eds.): T-Group Theory and Laboratory Method,
Wiley, 1964.
- Buchanan, Paul C.: "A system for evaluating supervisory
development programs", Personnel,
Jan 1955.
- _____ : "Testing the Validity of an Evaluation
Program" Personnel, Nov-Dec 1957.
- Bunker, Douglas R.: "The Effect of Laboratory Education Upon
Individual Behavior", National Training
Laboratories, Subscription Service.

- Burke, Richard L. and Bennis, Warren G.: "Changes in Perception of Self & Others During Human Relations Training", Human Relations, 1961, 14,2.
- Canter, Ralph R. Jr.: "A Human Relations Training Program" Journal of Applied Psychology, vol 35, no. 1, Feb 1951.
- _____ : "The use of extended control - group designs in Human Relations Studies", Psychological Bulletin vol 48, no. 4, July 1951.
- Cantor, Nathaniel F.: "The learning process for managers", Harper 1958.
- Carson, Robert C. and Lakin, Martin: "Some Effects of Group Sensitivity Experience", Duke University Paper presented SEPA Symposium, Miami Beach, Florida 1963.
- Castle, Peter F.: "The evaluation of Human Relations Training for Supervisors", Occupational Psychology, Oct 1952.
- Croston, F.E. and Cowden, D.J.: Applied General Statistics, Prentice-Hall, 1955 (second edition)
- Edwards, A. L. Experimental design in psychological research Rinehart, 1950.
- Festinger, Leon and Katz, Daniel: Research methods in the Behavioral Sciences", Dryden Press N.Y., 1953.
- Fleishman, Edwin A., Harris, Edwin F., Burtt, Harold E.: "Leadership & Supervision in Industry: an Evaluation of a Supervisory Training Program", Ohio State University 1955.
- Fryer, D.H.: "Training with special reference to its evaluation", Personnel Psychology, Spring 1951.
- Goodacre, D.M.: "The experimental evaluation of Management Training - principles & practices", Personnel, May 1957.
- Guilford, J.P.: Fundamental Statistics in Psychology and Education, McGraw-Hill, 1956

- Herzog, E.: Some Guidelines for Evaluative Research, U.S. Department of Health, Education, and Welfare, Social Security Administration, Children's Bureau, No. 375, 1959.
- Hyman, H., Wright, C., and Hopkins, T.: Application of Methods of Evaluation, University of California Press, 1962.
- Katz, Robt. F.: "Skills of an Effective Administrator", Harvard Business Review, Jan-Feb 1955.
- Kreck & Crutchfield: Theory & Problems of Social-Psychology, McGraw-Hill, 1948.
- Krugman, Herbert: "Management development training: who profits most?", Personnel, May-June 1959.
- Likert, Rensis: "Motivational approach to management development", Harvard Business Review, July-Aug, 1959.
- Mahoney, T.A., Jerdee, T.H., and Korman, A.: "An Experimental Evaluation of Management Development", Personnel Psychology, Vol.13 no. 1, Spring 1960.
- McKinney, A.C.L.: "Progressive levels in the Evaluation of Training Programs", Personnel, Nov-Dec 1957.
- McTavish, R.D.: "Evaluation of a unique two-year executive development plan", Advanced Management, May 1954.
- Miles, Matthew B.: "Human Relations Training, Processes & Outcomes", Journal of Counselling Psychology, Vol.7, no. 4, 1960.
- Nuttall, Richard: "An analysis of Growth in Management Development" - M.S. thesis M.I.T., 1960.
- Oslerberg, Wesley and Lindbom, Theodore: "Evaluating human relations training for supervisors", Advanced Management Vol. 28, Sept 1953.
- Pshry, Barry L.: "Changes in Perceptions of the On-the Job Problems Following Laboratory Training", Boston University and National Training Laboratories, 1962.

- Personnel Research
Dept.: "Findings of Research into Senior Management Seminars" - Hydro Electric Power Commission of Ontario, June 1962 (internal paper)
- Report No. 15,091, "Evaluation of Supervisory and Management Training Methods" - Organization for Economic Co-operation & Development, Paris, France, 1963
- Riecken, H.: The Volunteer Work Camp: A Psychological Evaluation, Addison-Wesley Press, 1952.
- Savitt, Morris: "Is management training worthwhile", Personnel, Sept-Oct 1957.
- Schein, E. H.: "Management Development as a Process of Influence", Industrial Management Review, M.I.T., May 1961.
- Tannenbaum, R.,
Weschler, I.,
Massarik, F.: Leadership and Organization, McGraw-Hill, 1961.
- Walker, Helen and
Lev, Joseph: Statistical Inference, Holt, N.Y., 1953
- Winn, A.: "An experiment in training", Canadian Personnel & Industrial Relations Journal, July, 1955.
- _____: "Training in Administration & Human Relations", Personnel, Sept 1953.