An Evaluation of an Organizational Training Program: Assessing Cognitive, Affective, and Behavioral Outcomes

Thomas L. Watson

Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_diss

Part of the Education Commons

Recommended Citation

https://ecommons.luc.edu/luc_diss/2335

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License. Copyright © 1985 Thomas L. Watson
AN EVALUATION OF AN ORGANIZATIONAL TRAINING PROGRAM:

ASSESSING COGNITIVE, AFFECTIVE, AND BEHAVIORAL OUTCOMES

by

Thomas L. Watson

A Dissertation Submitted to the Faculty of the Graduate School of Loyola University of Chicago in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

May 1985
ACKNOWLEDGMENTS

The writer wishes to express his appreciation to all those who assisted in the completion of this study. To the following who contributed to this dissertation, I acknowledge my appreciation and thanks.

To my advisor, Dr. Todd Hoover, for his assistance in formulating the original concept for the study, for his critiquing of this manuscript, and for his continued support and encouragement throughout the doctoral program.

To my committee members, Dr. Barney Berlin and Dr. Diane Schiller, for their guidance and assistance in finalizing this study.

To Dr. Leroy Cougle, a trusted and valued colleague and friend, for his unwavering counsel and advice.

To Larry Coon, Dean of Hamburger University, and to the entire faculty for their willing and total cooperation in the data collection for this study.

To the parents of the writer, Mr. and Mrs. Virgil Watson, who through personal example and repeated support and encouragement have fostered the attitude of life-long learning.

To McDonald's Corporation for its support of my educational pursuits and for its full and complete
cooperation in this study.

And finally, special recognition to my wife, Susan, and my sons, Jonathan and Jeremy, whose patience, understanding, and loving support enabled the pursuit of this goal.
VITA

The author, Thomas Lee Watson, is the son of Virgil and Betty (Niland) Watson. He was born July 13, 1948, in Sandusky, Ohio.

His elementary education was obtained at Kenwood Elementary School, Bowling Green, Ohio. His secondary education was completed in 1966 at the Barboursville High School, Barboursville, West Virginia.

In September, 1966, Mr. Watson entered Bowling Green State University, receiving a Bachelor of Science in Business Administration with a major in Personnel Management.

In September, 1970, Mr Watson was granted an assistantship in the School of Business Administration at Bowling Green State University, enabling him to complete a Master of Business Administration in June, 1972.

Since 1976, Mr. Watson has been employed by McDonald's Corporation, Oak Brook, Illinois, in several different training and development capacities.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................ ii
VITA ................................................ iv
LIST OF TABLES ........................................ vii
LIST OF ILLUSTRATIONS ................................ ix
CONTENTS OF APPENDICES ................................. x

Chapter

I. INTRODUCTION ........................................... 1
   Description of the Problem ....................... 7
   Importance of the Study ......................... 11
   Statement of the Problem ....................... 16

II. REVIEW OF RELATED LITERATURE .................. 18
   Training in American Industry ............... 19
   Training Evaluation ............................ 27
   Training Models ............................... 37
   Training Evaluations in the Literature ...... 51

III. METHOD .............................................. 62
   Statement of Purpose ........................... 63
   Hypotheses ..................................... 68
   Experimental Setting ........................... 73
   Description of Instruments ................... 76
Procedural Chronology .................................... 83
Description of Sample ....................................... 89
Design of the Analysis ....................................... 92

IV. RESULTS ....................................................... 94
Initial Comparability of Groups ......................... 95
Main Results .................................................... 99
Analysis of Variance of Posttest Scores ............ 110
Need for Achievement ....................................... 113
Other Individual Difference Variables ............. 124

V. DISCUSSION .................................................... 130
Verification of the Impact of Training ............... 131
Individual Difference Variables ....................... 134
Retention Over Time ......................................... 140
Implications of the Study ................................. 143
Indications for Future Research ....................... 146

SUMMARY ........................................................... 148

REFERENCES ....................................................... 150

APPENDIX A ....................................................... 155
APPENDIX B ....................................................... 161
APPENDIX C ....................................................... 163
APPENDIX D ....................................................... 168
TABLE

| 1. Analytic Paradigm for the Study | 67 |
| 2. Analysis of Variance of Total Points Earned by Individual Difference Variables | 85 |
| 3. Comparison of Demographic Data Between Hamburger University’s Student Population Norms and Sample Groups’ | 96 |
| 4. Pretest Mean Comparisons Testing for Differences in Trained and Untrained Groups on Knowledge, Attitude, and Behavior Measures | 97 |
| 5. Means, Standard Deviations, and Sample Sizes of Pretest and Posttest Scores | 100 |
| 6. Pretest-Posttest Mean Comparisons Testing for the Main Effect of Pretesting on Knowledge, Attitude, and Behavior Outcomes (I) | 102 |
| 7. Pretest-Posttest Mean Comparisons Testing For the Main Effect of Pretesting on Knowledge, Attitude, and Behavior Outcomes (II) | 107 |
| 8. Pretest Effect on Trained vs Untrained Subjects by Subtraction of Posttest Means For Knowledge, Attitude, and Behavior Outcomes | 108 |
| 9. Analysis of Variance of Posttest Scores for Knowledge, Attitude, and Behavior Outcomes | 111 |
| 10. Raw Scores Ranges and Number of Subjects For Category Scores on the EPPT | 114 |
| 11-A. Analysis of Variance of Posttest Scores Partitioned by Need for Achievement Level For the Knowledge Outcome Measure | 116 |
| 11-B. Analysis of Variance of Posttest Scores Partitioned by Need for Achievement Level For the Attitude Outcome Measure | 117 |
TABLE

11-C. Analysis of Variance of Posttest Scores Partitioned by Need for Achievement Level For the Behavior Outcome Measure ............... 118

12. Cell Frequencies and Posttest Means for the 2 X 2 ANOVA on Behavior Posttest Scores ............ 120

13. Analysis of Variance of Academic Achievement (TOT) by Need for Achievement (EPPT) ............... 122

14. Means, Standard Deviations for Academic Achievement (TOT) Classified by Need for Achievement Level (EPPT) ............... 123

15. Analysis of Variance of Posttest Scores by Education Level and Type of Employment for Knowledge, Attitude, and Behavior Outcomes ...... 125

16. Analysis of Variance of Total Points Earned by Education Level and Type of Employment ...... 126


18. Pretest, Posttest, 45-Day Posttest Paired Mean Comparisons Testing for Retention of Gains in Knowledge and Attitude Outcomes ...... 129
### LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>DESCRIPTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Pretest to Posttest Changes in Mean Score Performance on Knowledge Measures</td>
<td>103</td>
</tr>
<tr>
<td>2.</td>
<td>Pretest to Posttest Changes in Mean Score Performance on Attitude Measures</td>
<td>104</td>
</tr>
<tr>
<td>3.</td>
<td>Pretest to Posttest Changes in Mean Score Performance on Behavior Measures</td>
<td>105</td>
</tr>
<tr>
<td>4.</td>
<td>Behavior Posttest Score Interactions Plotted by Training Level for Each Need for Achievement Level</td>
<td>120</td>
</tr>
</tbody>
</table>
# CONTENTS OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Knowledge Pretest/Posttest</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A</td>
<td>A.O.C. Knowledge Test</td>
<td>155</td>
<td></td>
</tr>
<tr>
<td>Appendix B</td>
<td>Opinion Survey</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>Appendix C</td>
<td>Management Behavior Index</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>Appendix D</td>
<td>Need for Achievement Measure</td>
<td>168</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The Modified Edwards Personal Preference Schedule</td>
<td>168</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Training and Development represents a major investment in many companies. It is estimated that private and public U.S. employers invest over $30 billion each year in the education and training of their employees.1 The Training and Development department, or Human Resource Development department as it has also been called, provides programs in a wide range of job-related knowledge and skill areas to employees from all levels of the organization. These departments employ instructors and other professionals to design, select, deliver, and administer the programs which are usually presented in employer-owned facilities dedicated solely to learning activities. In addition, these departments have often been asked by the top management in their organization to take an increasing role in diagnosing and helping to resolve organizational problems and to actively participate in long-range human resource planning.

In answering the question "Why a Training and Development Department?", Laird responds this way:

1. Organizations get outputs because people perform tasks to a desired standard.

2. Before people can perform their tasks properly, they must master the special technology used by their organization. This means acquisition of knowledge and skill. Sometimes this acquisition is needed when the employee is new to the organization; sometimes it is needed because the organization changes its technology; sometimes it is necessary if an individual is to change places within the organization either by lateral transfer or promotion.

3. Training is the acquisition of the technology which permits employees to perform to standard.2

But as Laird goes on to point out, training departments are not only expected to train people to perform their present tasks properly, they must also take the responsibility to educate certain employees so they can assume greater responsibilities in the future and to develop people and entire organizations for futures... sometimes for undefined and undefinable futures.3

Several factors have combined to contribute to the growth in the amount of employer-sponsored training and education and to the expanded role of the training department:

1. Partly to compensate for inadequacies of traditional education, not only in the basic skills of secondary school graduates whose deficiencies may


3 Ibid., p. 32.
be a barrier to their ability to perform their jobs, but also to train many college-educated employees who lack abilities in a wide range of generic areas such as communication, problem-solving and decision-making, and interpersonal relations.

2. Partly to cope with economic and social changes that affect the workplace. These changes include government-mandated equal opportunity requirements, union seniority agreements requiring promotion based on tenure rather than ability, and the impact of both the U.S. and the world's economy on the business. This impact affects the products and services it sells, the markets where it can sell them, and the manner in which they can be most efficiently and profitably produced.

3. Partly to provide upward mobility for employees, through training for more technical or managerial responsibility. This role requires the training department to become active in the planning for and development of all of the human resources within the organization.

4. Partly to cope with the changes in technology which make job skills obsolete, to respond to the growing automation and electronic sophistication of American Industry.
5. Partly for proprietary reasons (this is how we want you to do it here) or for competitive reasons (sales training for a specific product line).

6. Partly for the inconveniences of scheduling, administration, or distance problems that are presented when trying to use a more traditional educational source.  

Most of the activities conducted by corporate training departments are intended to accomplish one or more of the following immediate objectives: 1) to improve an individual's level of self-awareness, 2) to increase an individual's skill in one or more areas of expertise, and/or 3) to increase an individual's motivation to perform his or her job well.  

The obvious intended ultimate outcomes of these training activities are subsequent positive changes in trainees' knowledge, skills, and attitudes, which are expected to result in more productive behavior on the job in the achievement of organizational objectives.

In light of the previously mentioned growth in the importance of training and development in American business and industry and the staggering sums of money being

4 Craig and Evers, pp. 6-7.

invested, it would seem obvious that the results and benefits of training activities to the organization are manifest and well-documented. The opposite is true and the issue of evaluating the effectiveness of training is one of the most critical issues facing the training profession today. In many companies, training was initially instituted and encouraged as a "good faith" investment in its most important resources, its people. In these companies that attitude was founded on the belief that training's benefits are obvious and need no formal examination. Such an attitude is perhaps best represented by the well-worn cliche "Training doesn't cost, it pays."

In most companies, however, the need to begin formally evaluating the effectiveness of training programs has been well understood for some time. The conscientious and concerned Training Director recognizes the importance of evaluating training programs, although in a recent survey of Training and Development Journal readers, at least one-third consider evaluation to be their most challenging problem on the job. The challenge was heightened for many companies during the turbulent cycles in the U.S. economy in the last few years which forced the top management teams of many corporations to ask some pointed questions.

concerning the return on investment they were receiving from their corporate training departments.
DESCRIPTION OF THE PROBLEM

The issue of training evaluation has received a great deal of attention in the training literature in the last thirty years. Consistently there has been unanimous agreement on the importance of and need for regular evaluation of training activities. There has been a wide divergence in points of view, however, regarding some of the basic issues associated with the evaluation process. Some of the many issues at debate are:

1. What audiences should the evaluation serve?
2. What criteria of evaluation should be used?
3. Who should do the evaluation?
4. What measurement dimensions should be used?
5. What data sources should be used?
6. What data collection techniques should be used?
7. What type of design should be used?
8. How should the results of the evaluation be used?

The wide range of issues incorporated in these questions have received much attention in the training evaluation literature as will be seen. However, in many cases the debate has served to only confuse or convince trainers of the complexity and hopelessness of the task.

In 1970, Campbell, Dunnette, Lawler, and Weick did a
thorough review of the management training literature for
the previous twenty years in an attempt to assess the
effectiveness of existing evaluation studies of management
training. They focused only on evaluation studies that had
employed a control group or, in the absence of a control
group, both pre-training and post-training measures. In
their view, only studies meeting these criteria could be
expected to provide meaningful data about the effectiveness
of a particular training course. A total of seventy-three
studies were found which met these criteria. Several
common shortcomings of these studies were found:

1. Less than a third of the studies had measured the
effects of training on individual job performance
or on results for the organization (sales,
profit, productivity, etc.) Most focused on
training outcomes typically assessed during or
just after a course, such as trainee reactions to
the course or improvement in knowledge.

2. Very few studies had compared the relative
effectiveness of two or more techniques of train-
ing in reaching a desired objective. In other
words, most studies did not attempt to determine
which training technique was most effective in
improving a particular skill (lecture, demonstra-
tion, case study, etc.)

3. No studies had measured the influences of
individual differences on the outcomes of training. For example, no study attempted to correlate the success of training with the level of a trainee's prior experience or level of education.

4. Few studies had investigated the effect that the organizational environment has on the transfer of training to the job setting.7

In 1981, Clement did a review of the training evaluation literature from 1970 through 1980. To allow for a direct comparison of findings, he used the same criteria, focusing only on studies employing a control group or at least pre-training and post-training measures. He also examined the same four issues of concern indicated above: 1) training outcomes measured in evaluation, 2) comparisons of relative effectiveness, 3) measurement of the influence of individual differences, and 4) measurement of the influence of the organizational environment. His findings were based on twenty-six studies meeting the above criteria and are as follows:

1. 32 per cent of the post-1970 studies reviewed could not meet the criterion of a control group, thus relying solely on post-training measures of training outcomes. This situation was worse than

the pre-1970 studies, where only 13 per cent had not employed a control group.

2. Few training researchers had attempted to measure the influence of the organizational environment or the impact of individual differences upon the success of a training course and little has been added to our knowledge of what training method is most effective in reaching a given objective.\(^8\)

Clement concludes his report by stating:

> In short, evaluation practices have not improved much since 1970. Furthermore, the outlook for the 1980's shows that evaluation may continue to play a lesser role in management training....The overriding implication is that evaluation practices (by training professionals) are unlikely to change until top management demands it.\(^9\)


\(^9\) Ibid., p. 12.
The population of the United States, while constituting only six per cent of the world's population, produces one-third of the world's goods based on an economy whose economic philosophy is capitalistic and whose predominant social institution is the corporation. The training of the managers of these corporations thus takes on added importance and training and management development, involves the very heart and future success of these institutions and the nation itself.\(^{10}\) Taken in this context, the importance of effective and appropriate design and execution of training evaluation is a key ingredient which enables organizations to accomplish their objectives.

In addressing this issue, Bunker and Cohen state:

> It is our contention that stated and unstated rationalizations to the contrary, competent evaluation is the cornerstone of meeting both organizational and individual educational needs, and of improving the cost effectiveness of the training function. Although a single evaluation rarely provides answers to all of the questions relevant to return on the training dollar investment, carefully planned and controlled research enables one to monitor and justify productive training expenditures, and to avoid or reduce unnecessary

losses. In addition to enabling organizations to accomplish their objectives, two additional reasons for evaluation have been identified by Bunker and Cohen:

1. Meeting individual educational needs: It has already been seen that one of the major tasks of the training department in an organization is to fully develop its human resources to prepare them to deal more successfully with both present and future challenges. Without effective evaluation it will be very difficult to determine how well this task is being accomplished. Campbell, et al., and Clement in their respective reviews of training evaluations have targeted the importance of individual differences that trainees bring to the training session and their potential impact on the desired outcomes. As a result of effective evaluation, trainers could target those individual difference variables, such as measures of education or past experience, which could then be used to select those trainees who would benefit most.


from the training program.

2. Improving the cost effectiveness of training: As with any other organizational activity in which the resources of time, money, and personnel are to be committed in the attempt to achieve certain desired outcomes, it becomes necessary to find appropriate ways to measure the return on this investment to insure these scarce resources are being well spent. Regardless of the economic climate in which the organization finds itself, the training director must be prepared to document how well the training programs he or she oversees are achieving the desired results and at what cost.

A final reason why the subject of evaluation is so important relates to how evaluation results can be made more useful. Most of the training evaluation literature deals with the methods of evaluation: how to make it more objective, accurate, precise, and scientific. As Brinkerhoff points out, evaluation tends to be viewed as a problem, a task that must be completed. Instead, he suggests evaluation should be viewed as a solution, relative to how training can be more efficient and effective.

Evaluation is an important part of any training and development effort. It is more than an assessment of outcomes or effects. Evaluation is systematic inquiry into training contexts, needs, plans, operation, and effects. It should help collect information to decide
what's needed, what's working, and how to improve it, and what's happened as a result.\textsuperscript{13}

This view of evaluation suggests that ideally the results of training evaluation should benefit the entire field of human resource development. It suggests that trainers should share their results with one another in an attempt to improve the state of the art of training and its varied attempts to improve training outcomes and their positive impact on organizational effectiveness.

This sharing of results has long been an interest of the Training in Business and Industry Special Interest Group (SIG) of the American Educational Research Association (AERA). AERA conducts annual conventions where members hear papers presented on topics of interest including the evaluation of training programs. In addition, there has been a growing interest on the part of training professionals and representatives from higher education to collaborate with each other in a number of areas including evaluation design and implementation. An example of such collaboration is a program co-hosted by the Arthur Anderson & Company and Northwestern University entitled "Joint Ventures Between Business and Higher Education in Human Resource Evaluation." The forum, held in conjunction with the joint Evaluation Network-Evaluation

Research Society National Conference, was conducted at the Arthur Anderson & Company Center for Professional Education in St. Charles, Illinois, October 16-18, 1983. It included representative from twenty-five academic institutions and an equal number of training professionals from business and industry. The primary focus of the conference was to promote the effective use of evaluation and applied research for training and retraining the nation's workforce. Attendees agreed on the importance of sharing the results of "in-house" training evaluations and focused on ways to overcome the barriers to this sharing process. A few of the suggestions to overcome these barriers included: 1) sharing of evaluation results through publication in a newsletter and other professional journals, 2) presentations at symposiums and to the Training in Business and Industry SIG of the AERA, and 3) collaboration with universities and colleges by allowing graduate students completing coursework in evaluation to conduct evaluations of selected training programs as part of a practicum.
STATEMENT OF THE PROBLEM

It was the critical importance of this issue of training evaluation and the author's personal interest in the subject as a practicing training professional that formed the basis for this study. The question this study seeks to address is: What is effective training evaluation and how should it best be designed and implemented. In addressing this question, the study focused on the basic issues associated with the evaluation process which have been previously identified:

- Identifying the audiences the evaluation will serve
- Selecting the criteria of evaluation to be used
- Selecting the evaluators
- Identifying and selecting measurement dimensions
- Selecting an appropriate design
- Use of results

Because of the unique opportunities afforded by the author's work setting, serving as a trainer in a industrial training facility, there was a convenient and readily available setting within which to investigate the various issues related to the question of what constitutes effective training evaluation. A specific objective of the study then was to take advantage of this convenient access
to on-going training activities in a corporate training facility by selecting one of the on-going programs as the focus for the evaluation design and implementation.

The limitations of this study include the following:

1. No attempt is made to generalize beyond the immediate setting. All conclusions are limited to the specific training course and to those receiving the training.

2. Only two classes (263 students) were selected from among the sixteen classes (over 2000 students) that complete the course on an annual basis. It is believed these two classes constitute a representative sample; however, only these 263 students were formally evaluated.

3. The use of "mailed-in" instruments created the potential for missing data, subjects who either forgot or elected not to complete the instruments and mail them in. A certain degree of control is also surrendered when relying on students to complete instruments accurately and unassisted.
CHAPTER II

REVIEW OF LITERATURE

INTRODUCTION

The review of literature is presented in four sections. The first deals with training in American industry, its historical roots, its growth and present status, and its organizational role. The second section attempts to summarize the literature concerning the definition of evaluation, the many issues involved, and the rationale for why training evaluation should be completed. In the third section, several models of evaluation are presented, both traditional and popular ones with a brief description of their underlying assumptions. Finally, some examples of present day training evaluations found in the literature are presented with a summary of their respective strengths and weaknesses.
TRAINING IN AMERICAN INDUSTRY

The foundations for employer-provided education finds its roots in the Code of Hammurabi, some 4000 years ago, in which rules governing apprenticeship can be found. The medieval craft guild refined this form of training and it survives today within the craft-trade union structure and elsewhere. The basic method of apprenticeship training in which a skilled worker transmits knowledge to a learner by coaching, learner observation, and supervised practice undergirds much of industrial skills training.\(^{14}\) By the 1800's, apprenticeship alone was inadequate to meet America's growing need for skilled workers. Employers began to find other ways to train its employees. In 1872, Hoe and Company, a New York City printing manufacturer, established what was to be the first of many "factory schools" to train machinists so the firm could keep up with its expanding volume of business. Other companies were to follow suit including Westinghouse (1888), General Electric (1901), Baldwin Locomotive Works (1901), and International Harvester (1907). As interest in this activity grew and several other companies also established

schools - Western Electric, Goodyear, Ford, and National
Cash Register - some sixty representatives of thirty-four
firms established a National Association of Corporation
Schools in 1920. This organization was later to become the
American Management Association in 1923. In 1917, the
Smith-Hughes Act authorized the first federal funds for
vocational education.

During World War I the Emergency Fleet Corporation of
the U.S. Shipping Board was created and given the task of
training several hundred thousand workers to build a
"bridge of ships" to Europe. This organization, led by a
former vocational school instructor, developed a four-step
method for shipyard supervisors to use in training new
workers which was to become the central method for pro-
viding much of the industrial skills training until World
War II ("Show, Tell, Do, Check").

The years after the war saw little that was unique in
the area of industrial training, although there was an in-
creased use of correspondence schools. The Depression era
of the 1930's created two training influences that were to
have later impact: 1) the population became "training-
conscious" as a result of widespread handicrafts training

15 Craig and Evers, op. cit., p. 9.
16 Bird McCord, "Job Instruction," in Robert L. Craig
offered with federal funding in those years and 2) business recognized that economic recovery would only occur if people could be encouraged to buy their products. The National Society of Sales Training Executives was founded in 1940 to facilitate professional communications in this area.

World War II, with its demands for personnel and material, provided a major impetus to training. Nearly two million plant supervisors and foremen were trained in methods that enabled them to train an unskilled workforce, (many of whom were either women, elderly, or had some type of physical disability). Training became an integral part of the supervisory job function and there was a widespread emergence of training directors to coordinate the effort.

Led by industry executives with WWI Emergency Fleet Corporation experience, the Training Within Industry (TWI) Service, later part of the Manpower Commission, in 1940 began refining years of know-how into three major training programs to be used by plant operating personnel. They were the so-called "J" Programs - Job Instruction Training, Job Methods Training, and Job Relations Training - that helped American Industry meet its wartime production needs.

17 Steimetz, op. cit., pp. 1-10.
18 Ibid., p. 13.
Later programs were added in Job Safety Training and Program Development Training. ¹⁹

A simultaneous need to upgrade workers in college-level subjects was recognized and Engineering, Science and Management War Training (ESMWT) was launched. ESMWT programs in technology and management were conducted by colleges and universities both on and off-campus. Steinmetz says that "in many communities they became the forerunners of junior or community colleges," as well as "the strong roots" for the continuing education centers and management training centers that developed later. ²⁰

During the 1950's there was increased interest in "management development." This was based not only on the wartime emphasis on the importance of the supervisor's role in the training of employees, but also because of the advancement of management concepts in the fields of behavioral science. The importance of modern-day supervisory training had its formative growth during this period. The increasing emphasis on "human relations training," focusing on helping people to work together more effectively and productively, continued into the 1960's. Such managerial skills as leadership styles, communication skills, and problem-solving received increasing emphasis.

¹⁹ McCord, op. cit., pp. 32-34.
²⁰ Craig and Evers, op. cit., p. 11.
A popular development during this period was the use of "teaching machines" and "programmed instruction." Though the movement has lost its original "faddish" appeal, it has left behind a strong commitment to sequentially-arranged instruction based on a thorough assessment of training needs. Instructional System Design, as it was later to be called, has been used by many companies in the design of their training programs.

The rapid technological changes of the 1960's and 1970's, the gradual shift in balance from manufacturing to the service industry, and the changing composition of the American workforce all combined to place increasing demands on the important role of the training function. The modern training department is expected to assume an expanded role and in many organizations has been renamed the Human Resource Development department. Its concern has become overall workforce competence, from entry-level, hourly-paid employees through the middle ranks of supervisory management to the top levels of executive management. Within this broad range of employees are included a host of technical and managerial skills required for the organization to accomplish its objectives. The modern day Human Resource Development professional is charged with a number of additional duties beyond those traditionally associated with the training function including: indentifying and helping to solve organizational problems,
strategic planning in forecasting the short-term and long-term organizational needs for human resources, and finding new ways to increase organizational effectiveness.

A 1978 study completed by the American Society for Training and Development identified nine major activity areas for member training and/or human resource development professionals: 1) analyzing needs and evaluating results, 2) designing and developing training programs and materials, 3) delivering training and development programs, 4) advising and counseling, 5) managing training activities, 6) maintaining organizational relationships, 7) doing research to affect the training field, 8) developing professional skills and expertise, and 9) developing basic skills and knowledge.

The modern growth of corporate investment in training and education is substantial, exceeding $30 billion annually. One indication of the growth of the field is found in the ASTD membership rolls which have doubled in the last decade to over 21,000 national members with another 20,000 holding local membership in one of ASTD's 127 chapters throughout the U.S. The major portion of this growth has occurred in the 1974-1980 period.

21 Craig and Evers, op. cit., p. 16.
22 Ibid., p. 5.
Obviously, the training function varies within individual organizations based on its needs and goals. A study of management and training practices by the Bureau of National Affairs, Inc. in 1977, surveyed 113 organizations, 65 per cent with 1000 employees or more, 35 per cent with 1000 or less. For first-level supervisors, 75 per cent operated "in-house" training programs; for middle managers the figure was nearly 67 per cent. The next most commonly reported form of management training was "attendance at job-related outside seminars" (66 per cent for supervisors; 89 per cent for middle managers). "Self-training/correspondence courses" were used at approximately the same level in both cases (50 per cent for supervisors, 45 per cent for middle managers). "Attendance at professional or trade association meetings" was used in 96 per cent of the responding organizations for middle managers, 54 per cent for supervisors. "University development programs" had the lowest usage figures reported: 6 per cent for supervisors and 39 per cent for middle managers. 23

An ASTD National Report in October, 1980, surveyed "on-site training or education" practices in the manu-
facturing, utilities, banking and insurance industries for nonexempt (hourly) production/operations and

23 Ibid., p. 16.
office/clerical employees, as well as lower-level exempt (salaried) employees. It found that the most common training method used was "planned on-the-job training"—at least 73 per cent of the companies reported this for the non-exempt categories, 61 per cent for the exempt.

"Lecture, demonstration, and group discussion" was the next most frequently used method overall—45 per cent for production/operation, 59 per cent for office/clerical, and 76 per cent for exempt employees. The next most frequently reported method for production/operations was apprenticeship." At least 37 per cent of the companies reported using full-time company instructors for training all three employee categories, and at least 24 per cent reported using training vendors and suppliers to the company in training the same three employee groups. Finally, 48 per cent reported using "consultants or other outside private training specialists" to train exempt employees. 24

Clearly the role of the modern day Training and Development or Human Resource Department in an organization is a critical one, and one in which major resources of time, money, people, and equipment are being committed. But it is not clear how the results of this training is being evaluated.

TRAINING EVALUATION

Throughout the early history of training and development previously reviewed, the issue of evaluation was not a matter of serious concern. It was assumed that trained employees were a benefit to the organization and that training and development efforts were contributing to the overall accomplishment of organizational goals. Programs were generally allowed to stand on their own merits and little demand for evaluation was heard, either internally from training personnel or externally from top management or other departments within the organization. As has been seen, the historical beginnings of training focused on apprenticeship and the heavy emphasis on the acquisition of technical skills. Thus, in practice, evaluation was simply a matter of determining how well the trainee could perform the particular skill and how long it took him to perform it. With the advent of management training, human relations, and other conceptual skills building, the issue of how and what to evaluate became increasingly less clear.

The need for some type of evaluation receives nearly unanimous agreement in the literature. Tracey states the need for evaluation within the specific context of
training and development programs. Beyond the mere fact that executives and managers may be demanding systematic appraisal of training and development efforts, evaluation is important to these activities, just as it is to any other organizational element as a means of determining activity levels at any given moment and of providing a baseline for the measurement of progress. Stated simply, he submits that evaluation can determine whether the time, energy, and money expended in planning and operating training and development programs are producing results which justify the investment. Tracey defends his position on the need for evaluation by identifying three functional ways in which it is critically important:

First, the steady growth of training and development activities in most enterprises, which in total involve millions of people and many more millions of dollars, makes it essential that those responsible for the management of those activities be able to defend their programs by knowing the accomplishments and contributions of the activities to enterprise goals.

Second, evaluation provides trainers with a means of determining the efficiency, effectiveness, and utility of both management and operation. Only by appraisal is it possible to insure that programs are suited to the groups for which they are designed and that they result in the behavioral changes required for improved products or services.

Third, evaluation provides a starting point for the design of an improvement program. 25

Though much of the literature dealing with training evaluation begins with the assumption that the reasons for doing evaluation are obvious, the following list includes many of the most commonly stated and implied reasons:

1. Many groups are interested in the results.
2. To enhance survival of the training function.
3. To determine if dollars are being well spent.
4. To improve program, methods, and techniques.
5. To improve program outcomes.
6. Evaluation is intrinsically good.
7. It is required by a higher authority.
8. To determine training needs.
9. To establish better criteria for trainee selection.
10. To determine what to drop/add to training.
11. To provide a basis for cutting costs.
12. To determine if there is a pay-off.
13. To permit calculations of return-on-investment.
14. To insure against doing harm.
15. To measure exactly what is being learned.
16. To document specific behavior change on the job.
17. To identify ways training can make a greater contribution to profit and other corporate goals.
18. To verify how much of what is learned is retained.
19. To determine if specific training is valuable.
20. To determine if ultimate objectives of training are being accomplished.

21. To give trainers a sense of importance and accomplishment.

22. To suggest areas for further research.

23. To satisfy a requirement of a federal, state, or local governmental agency.\(^{26}\)

Though this list of reasons for evaluating training will generally produce agreement among training professionals, none-the-less, there are many companies who do little or no evaluation of their training activities. As has already been seen, nearly one-third of the trainers surveyed by ASTD view it as their single, greatest challenge. A list of the most commonly cited reasons for not evaluating training would include: lack of time, lack of staff, lack of money, need to convince supervisors, lack of effective methods, the overwhelming magnitude of the task, and the risk of discovering training outcomes are minimal or not measurable.\(^{27}\) Perhaps two of the real reasons accounting for the lack of training evaluation

\(^{26}\) This list is a consolidation of a variety of sources, the primary one being: Alden Sullivan, "An Analysis of Management Training Program Evaluation Practices in American Industry," Doctoral Dissertation, (George Washington University, (June, 1970), p. 73.

\(^{27}\) Ibid., pp. 75-77.
concerns the fact that 1) there is a perception of "limited benefit" from an evaluation, and 2) there is no real incentive to do evaluation. Few training departments are asked to "document" their results, or even to nominally assess them, so that the amount of evaluation that may be done in a particular organization is left solely to the discretion of the training director.

There is no single, integrated, and generally accepted model of evaluation that is uniformly embraced by the training profession with the possible exception of assessing trainee reactions to the program. Though serving some immediate practical purposes such as providing immediate feedback to the training staff on how well their efforts were received, these "feel-good" or "happiness" indexes (as they are sometimes derisively described in the literature) fall short of what is generally defined as acceptable training evaluation. The major criticism of what little training evaluation studies can be found in the literature can be summarized as follows: they have not been objective, systematic, comprehensive, or scientifically accurate, often violating most of the requirements of basic experimental design. Thus, the literature seems to be suggesting that most evaluations "prove nothing" about the efficacy of training. In response, many training professionals have countered that "educational studies" and "experimental research" is unrealistic, impractical, and
irrelevant to their "real-world" needs. A major source for this apparent disagreement is in the confusion over the definition of the word "evaluation."

Some writers suggest that the two terms "evaluation and "research" are essentially synonymous and equal; Suchman combines the two in his use of the term "evaluative research." He defines this as "the specific use of the scientific method for the purpose of making an evaluation." In his view, there are more similarities than differences in the techniques that can appropriately be used by the evaluator and those that can be used by the researcher. He states:

Hence, evaluative research has no special methodology of its own. As "research" it adheres to the basic logic and rules of scientific method as closely as possible. Its canons of "proof" and its laws of inference are the same as those of any research project. It utilizes all available techniques for the collection and analysis of data and employs a wide variety of research designs....In other words, evaluative research is still research and it differs from nonevaluative research more in objective or purpose than in design or execution.

In Suchman's view, evaluation as a goal is differentiated from evaluation research as a means of attaining that goal.

Hemphill takes a significantly different view in the

---


29 Ibid., p. 81.
treatment of the research and evaluation question. He states:

It is to be regretted that evaluation studies have earned the reputation of being poorly conceived and executed research. Despite the fact that precision, care, discipline, and logical thought are the marks of 'good' evaluation as well as 'good' research, there is no requirement that evaluation studies must be judged on the same basis as that on which research studies are now conventionally judged. It is suggested that the criterion of worth of an evaluation study (program) is to be found in its contribution to a rational decision process. 30

Stufflebeam, et al., further address the issue:

Perhaps the greatest challenge facing the evaluator is overcoming the idea that evaluation methodology is identical to research methodology. Equating them forces several constraints inimical to the purposes of evaluation and makes it impossible to meet certain of the needs served by good evaluation.... Evaluations are not designed to establish universal laws but to make possible judgments about the phenomenon. 31

Tyler, in addressing the issue of how research and evaluation studies differ, indicates that evaluation studies differ in the manner in which value questions are resolved - especially value questions that help determine choices about what information is sought. In his view, the "ideal" research study is one in which:


1. Problem selection and definition are the responsibility of the individual doing the research.

2. Tentative answers (hypotheses) to the problem may be derived by deduction from theories or by induction from an organized body of knowledge.

3. Value judgments by the researcher are limited to those implicit in the selection of the problem.

4. Given the statement of the problem and the hypothesis, the research can be replicated.

5. The data to be collected are determined largely by the problem and the hypothesis.

6. Relevant variables can be controlled or manipulated and systematic effects of other variables can be eliminated by randomization.

The Evaluation study may be described in terms of characteristics almost the reverse of those outlined above:

1. The problem is almost completely determined by the situation in which the study is conducted. Many people may be involved in its definition, and, because of its complexity, the problem is initially difficult to define.

2. Precise hypotheses usually cannot be generated; rather the task becomes one of testing generalizations from a variety of research studies, some of which are basically contradictory. There are many gaps which in the absence of verified knowledge must be filled by reliance on judgment and experience.

3. Value judgments are made explicit in the selection and the definition of the problem as well as in the development and implementation of the procedures of the study.

4. The study is unique to a situation and seldom can be replicated, even approximately.

5. The data to be collected are heavily influenced if not determined by feasibility. Choices, when possible, reflect value judgments of decision-makers or of those who set policy. There are often large differences between data for which the collection is feasible and data which are of most value to the decision makers.

6. Only superficial control of a multitude of variables important to interpretation of results is possible. Randomization to eliminate the systematic effects of these variables is extremely
difficult or impractical to accomplish. 32

Perry categorizes the differing schools of thought as to how evaluation should be defined into four major groupings:

Evaluation as Performance of Objectives - Evaluation is the process of comparing performance data with clearly specified objectives stated in behavioral terms.

Evaluation as Measurement - This definition was the norm following the growth and popularity of the measurement movement in education during and after World War I. The focus was on use of valid and reliable measurement instruments and techniques.

Evaluation as Professional Judgment - This approach utilizes the opinions of experts to support decisions. Criteria established by the experts themselves form the basis on which the evaluation is made.

Evaluation for Decision-Making - Evaluation's purpose is to provide the best possible basis for informed decision-making by program managers. The evaluator's role is to identify and collect information to support the process of choosing the best from a variety of decision alternatives. 33


Mahler discussed three levels or types of evaluation design:

1. **Common Sense Evaluation** - The unsystematic reporting of facts, inferences, or feelings about the training.

2. **Systematic Evaluation** - The planned use of quantitative methods decided in advance, but not necessarily experimental in nature.

3. **Experimental evaluation** - Quantitative measurement under controlled experimental conditions.34

---

The varied views on training evaluation have produced a variety of models. Several major models will be reviewed, generally categorized by the definition of evaluation with which they are most associated.

PERFORMANCE OF OBJECTIVES - Ralph Tyler's work originally supplied the theoretical basis on which this type of evaluation is based. His model characterized evaluation as the process of comparing performance data with clearly specified objectives stated in behavioral terms. Suchman proposes a model of this type. He views the program objectives which are to be evaluated as the hypothesis to be tested in basic research. From this point of view, an evaluation project is a study of change, in which the program being evaluated is the stimulus and the desired change is the dependent variable. He suggests that the project be presented as a series of hypotheses to be tested, which state that "Activities A, B, and C will produce results X, Y and Z." He further states that "the most identifying feature of evaluation research is the presence of some goal or objective whose measure of
attainment constitutes the main focus of the research problem." 35

A modern-day practitioner of this philosophy and a recognized "giant" in the training and development field today is Robert Mager. Though essentially an instructional design specialist, he has had a profound impact on the field through his insistence on well-contructed, measurable, and behaviorally-grounded objectives. His course, "Criterion Referenced Instruction," has been used by many companies to train its instructional design staff. Its strong emphasis on clearly establishing the precise behavioral outcomes sets up a form of evaluation in which the trainee outcomes -- measures of knowledge, attitude, or behavior -- are compared to original course objectives.

DECISION-MAKING - Alkin represents the decision-oriented point of view. He defines evaluation as the "process of ascertaining the decision areas of concern, selecting appropriate information, and collecting and analyzing information in order to report summary data useful to decision makers in selecting among alternatives." 36

35 Suchman, op. cit., p. 27.
Five areas of evaluation are identified and described:

1. Systems assessment - Determining range and specificity of program objectives.
2. Program planning - Produces ways to reach objectives.
3. Program implementation - Determines degree to which planning descriptions and the program actually implemented correspond.
4. Program improvement - Provides data on the extent to which objectives are being achieved.
5. Program certification - Determines whether objectives are being reached.  

Another type of decision-making is the one originally developed by Stufflebeam and leading educators of the National Study Committee on Evaluation of Phi Delta Kappa. Called the CIPP Model, it is an acronym formed from the four basic components found in this model: context, input process, and product. The four components of evaluation in the model are derived from four types of decisions: planning decisions, structuring decisions, implementing decisions, and recycling decisions. Context evaluation is used to provide a rationale for determining objectives. It defines the relevant environment, identifies unmet needs and unused opportunities, and diagnoses specific problems. Input evaluation provides information to determine how to utilize resources to best meet program goals. It is used for deciding if outside assistance is necessary, determining the general strategy to be used, and planning and

37 Ibid., p. 25.
designing the program. Process evaluation provides feedback to persons responsible for implementation. It is accomplished through monitoring potential sources for failure, providing information for preplanned decisions during implementation, and describing what actually takes place. Product evaluation measures and interprets the attainment of objectives. It should measure intended as well as unintended outcomes. A 1982 survey of ASTD members nationally found that the CIPP model was preferred over the next most popular model (Kirkpatrick's) for evaluating management education.

Alden proposes another type of decision-making model in which he defines successful evaluation as a clear focus on the management decisions being considered for the training programs. He identifies four factors which need to be determined before a meaningful evaluation project can be designed:

1. Will management even consider making a decision about whether or not to change a program or how it should be changed?


2. What research questions will provide the data necessary to make the management decisions?

3. What level of data is practical to collect and important enough for management to use in the decision-making process?

4. What criteria will management use to make the decision?

His model identifies four levels of data which can be applied to almost any kind of research question: participant perception, expert opinion, measurement of behavior, and measurement of end results. Finally, Alden suggests that for organizational decision-makers to make maximum use of the evaluation results, "minimum acceptable criterion" must be established that defines when change should or must take place. He identifies two ways for establishing such criterion:

1. Absolute Standard - The criterion represents a "threshold" level that the measured factor must equal or exceed.

2. Comparative Standard - The criterion is derived from the performance of a control group. To meet the minimum acceptable criterion, the difference between the evaluation and control group measures must exceed chance probability.

Other writers in the field of evaluation propose models which are a blend of definitions. Provus, for

example, proposes an approach which combines decision-making and professional judgment schools of thought. While he maintains that evaluation should be used to determine whether a project should be continued, changed, or terminated, this process involves: 1) agreeing on program standards, 2) determining if a discrepancy exists between standards and accomplishments, and 3) using this information to pinpoint areas of the project where improvement is needed. According to Provus, this model is accomplished in five stages: design, installation, process, product, and cost.  

Scriven describes an objective/decision approach when he defines evaluation as:

"a methodological activity which consists simply in the gathering and combining of performance data with a weighted set of goal scales to yield either comparative or numerical ratings, and in the justification of (a) the data-gathering instrument, (b) the weighting, and (c) the selection of goals."  

In this model the evaluator first examines the worth of goals and then determines if they are being met. Formative evaluation (on-going) provides information on the merit of


the program; summative evaluation (end-of-project) provides information on this success. The evaluation report, composed of data-based value judgments, is used for making decisions on program improvement.

Another model combining elements of the measurement, performance of objectives, and decision-making evaluation classifications is the one proposed by Kirkpatrick. He outlines four levels or steps in the assessment of an activity:

1. Reaction: How well did the trainee like the program? It includes rating sheets that trainees complete assessing their attitudes and opinions about course content, its relevance, the effectiveness of the trainer(s), etc.

2. Learning: What principles, facts, techniques were learned? This type of evaluation uses paper-and-pencil knowledge tests assessing cognitive outcomes or skill-demonstration tests where trainees must demonstrate specific learned skills.

3. Behavior: What changes in job behavior resulted from the program? Instruments assessing specific on-the-job behaviors by trainees are used which are completed by superiors, subordinates, and peers.

4. Results: What were the tangible results in terms
of impact on organizational performance? Most commonly, such measures as increased profits, higher sales, lower turnover, etc. are used. 43

Kirkpatrick's model has had wide influence on many training professionals attempting to evaluate their programs. Brethower and Rummler have translated Kirkpatrick's four levels of reaction, learning, behavior, and results into four questions:

1. Are the trainees happy with the course?
2. Does the training course teach the concepts?
3. Are the concepts used on the job?
4. Does the application of the concepts positively affect the organization? 44

In addressing these four "levels of evaluation," they propose an additional series of questions:

1. What questions do we want answered?
2. What might we measure to answer these questions?
3. What are the dimensions of learning or performance we are trying to measure?
4. What are the sources of the data to help measure?
5. What are alternative ways of gathering data for measurement?
6. What are the evaluation criteria we want to apply to each question? 45

This combination of four levels and five categories of questions forms a matrix which can drive the design and

EXPERIMENTAL EVALUATION - The issue of the differences between evaluation and research design have already been presented. The major issue that most training professionals take with making more regular use of appropriate research design is the "real world" limitations that it presents. It is extremely difficult to engineer and enforce all of the conditions necessary to the proper use of laboratory research design in the field setting.

Brethower and Rummler propose four research designs for the "real world" and some accompanying restrictions:

1. Control group - While one group receives training, a comparable group does not.

Difficulties: Hard to find two "naturally-occurring" groups comparable on the same dimensions. Even if possible to find, other variables other than "training" are likely to change; job conditions, the economy, sales levels, thus confounding the results and making clear-cut conclusions difficult to make.

Suggestions: Use only when the evaluation is seen as important enough to take extraordinary measures to ensure continuing comparability.

2. Reversal or ABA Research Design - This type of design evaluates by: 1) taking a baseline measurement, 2) implementing the training and
measuring its impact, and 3) returning to the original condition by removing the training and remeasuring.

Difficulties: If measured performance improves, management is unwilling to stop the "training" - the perceived cause.

Suggestions: Use in cases where naturally-occurring changes will lead to removing the training or other practices being evaluated.

3. Multiple Baseline - The same program is used with different groups at different times. The design is an attempt to determine whether the change in performance is "caused" by the training or just concomitant with the training.

Difficulties: Organizations may be unwilling to spread out the introduction of something that has proven successful in the pilot setting.

Suggestions: A decision to introduce progressively by area rather than everywhere at once may be necessitated by lack of staff or by the argument that "careful high quality introduction more than compensates for the loss incurred by not installing the program all at one time."

4. Before and After Measurement - Performance is measured "in the relevant job setting" before the training and again after the training.
Difficulties: Results may not be directly attributable to training but to any other change that occurred during the same time. Any one, or combination of changes may at to confound results.

Suggestions: Use only as a "last resort."
Maintain a detailed log of any changes which might be tied to the measured performance. 46

The use of extended control group designs have long been recommended for training evaluations by many; Campbell, et al.; Goldstein; Solomon; Entwisle; and Canter. All have outlined the many advantages this type of design. Campbell and Stanley reviewed six basic experimental designs relative to their internal and external validity in social research settings. Internal validity is concerned with correctly concluding that an independent variable is, in fact, responsible for variation in the dependent variable. Threats to internal validity include: history, maturation, testing, instrumentation, regression, selectivity, mortality, and interaction. External validity is concerned with the generalizability of research findings to and across populations of subjects or settings. Threats to external validity include the following: interaction of selection, interaction of testing, reactive effects, and

multiple treatment interference. Of the six designs, the Solomon 4-Group design, with two experimental groups and two control groups and with both a pretest and a posttest, was identified as the most feasible and complete design available.

Bunker, in a review of the pre-test literature from various attitude and learning situations, concluded the following:

1. Pretest effects have been shown to be potentially serious contaminants in certain evaluation situations.
2. The risk of confounding pretest effects is greatest in situations where the pretest involves learning or recall of previously-learned material.
3. Interactions involving pretest performance and personal and situational variables such as I.Q., sex, voluntary-involuntary participation, and the time lag between pretest and treatment have the potential to mask the main effects of pretesting "Pretest x Treatment" interactions.
4. Pretest effects have not been a problem in attitude research except in situations where the pretest constituted a learning device.
5. Extended control group designs are required to test for the presence of contaminating pretest effects.
6. Failure to control for pretest effects in the evaluation design may lead to erroneous predictions and/or conclusions regarding treatment effects.

Bunker goes on to point out several unique characteristics of the industrial training setting that suggest the potential for evaluation contamination by pretesting may be quite high. One characteristic concerns the fact that


48 Bunker, op. cit., p. 22.
the focus of much training is in imparting technical skill and knowledge to employees who may have prior, though limited, exposure to the material being presented. Trainees have often been exposed to the subject matter by working on the job for a period of time before being sent for training. Even if the training is designed to teach new skills to an experienced worker, the material is often a related extension of the tasks the worker is already performing on the job. In addition, where training is somewhat systematic and the sequence of training is well-defined, it is possible for trainees to begin "preparing" for the next level of their training through study or on-the-job observation and experience. Bunker concludes by offering several additional advantages for the practice of pretesting in industrial training:

1. Deciding whether to waive training courses for individuals who already have the requisite skills and/or knowledge;
2. Determining the optimal place to insert a given course in the training sequence (given employee skills, aptitudes, and learning potential for the subject matter);
3. Establishing individual indices of training impact as opposed to an estimate of the overall effectiveness of the program for the average trainee; and
4. Determining the mean level and variance of relevant employee behaviors prior to training, in order that courses can be established that will be appropriate for company needs and trainee potential. 49

49 Ibid., p. 25.
Summary

It seems clear that the choices facing the training professional, when it comes to evaluation, are many: from doing none at all; to resorting to the basic, but potentially useful "happiness" indexes; to a well-designed and implemented evaluation study applying the principles of appropriate research design. In conclusion, Perry summarizes well by stating:

The definition of evaluation in use determines, to a large extent, the procedures used in any evaluation study. If evaluation is equated with measurement, a program might be assessed, for example, by measurement of students on a standardized test. If evaluation is defined as synonymous with professional judgment, the worth of the program would be assessed by experts. If evaluation is defined as a comparison between performance and objectives, behaviorally-stated objectives would be established for the program and relevant student behaviors would be measured by using either standardized or evaluator-constructed instruments. Finally, where evaluation is seen as a process leading to informed decision-making, the gathering of information in order to assist in the making of a good judgment will comprise the evaluation activity. 50

50 Perry, op. cit., p.9.
TRAINING EVALUATIONS IN THE LITERATURE

As has already been seen, there is a scarcity in the training literature of well-designed training evaluation studies. Though it is likely that this may indeed indicate that there is not much serious training evaluation being done by training professionals in business and industry today, it is this author's contention that there are some "restraining influences" which might prohibit publication of the results of these studies and thus prevent some from ever reaching the training literature.

Some of these restraining influences may include:

1. The proprietary nature of the results - The unwillingness to share program specifics with a public that might include competitors, etc.

2. Corporate policy which inhibits or may outright prohibit the publication of studies completed internally, for internal use only.

3. The reluctance on the part of some training directors to publish negative, or less than "overwhelmingly positive" evaluation results which may reflect unfavorably on the department and the worth of its programs.

Unquestionably, there is a genuine lack of commitment
to the importance of evaluation. A recent survey of 103 East Coast companies found that 41 per cent evaluate programs solely by means of participant evaluations after a session. Asked why they make no further effort at evaluation, responses varied but could perhaps best be summed up by the statement: "There is a real question as to the cost-effectiveness in any organized effort to follow-up." 51

As has been previously cited, Clement reviewed twenty-six studies found in the management training literature since 1970 and compared them to seventy-three studies reviewed by Campbell, et al., covering the pre-1970 period. Both reviews focused only on studies that utilized the combination of control groups and pre and post training measures. Clement's conclusion is that overall there has been little progress in the practice of training evaluation.

A representative sample of evaluations found in the literature will be reviewed illustrating the various types of training evaluation previously identified. Campbell, et al., and Clement identified four issues of concern in their reviews of evaluations studies: 1) training outcomes measured (either internal or external), 2) comparisons of

51 Lawrence S. Munson, How to Conduct Training Seminars (New York, 1984).
training method effectiveness, 3) measurement of the influence of individual differences, and 4) measurement of the organizational environment. Studies will be reviewed in light of these criteria.

A study focusing on internal and external outcomes as well as on the comparison of effectiveness of three different training methods is one done by Smith. It also employed equivalent training and control groups and used pre and post measures of training outcomes. According to Clement, it is one of the best post-1970 studies to be found in the literature. 52

The study consisted of two parts. In the first study, modeling training was administered to eighteen branch managers of IBM that were matched with a control group of thirteen branch offices for geography, size, and employee satisfaction. The training focused on how to communicate effectively with individual employees and groups of employees, how to feedback opinion survey information, and how to prepare meaningful action plans to improve morale. The control group of thirteen managers received opinion survey data by mail, conducted branch office meetings to present the data to employees, and then submitted action plans to their superiors. The trained

52 Clement, op. cit., p. 10.
A group of eighteen managers received modeling training before and after the branch office meeting to present the survey results to employees. Four months after training, a forty-seven question "meeting effectiveness" questionnaire was given to both groups. Thirty-one per cent of the employees in the trained group rated the feedback as more effective versus 20 per cent in the control group. This difference was significant at the .01 level.

A second study, designed to pursue the issue of the effectiveness of the modeling training, utilized three trained groups: a one-day Traditional Session, a two-day Modeling session, and a two-day Modeling Session Plus Team-Building session. Training effectiveness was measured in terms of the branch managers' improved communication skills, customer satisfaction survey scores, and branch office sales performance versus quota. A control group again was used and all four groups were matched on levels of customer satisfaction, sales performance, and geography. The results on the communication measure indicated no significant change in the Traditional group, but significant improvement in communication effectiveness for both the Modeling and the Modeling Plus Team-Building groups. Customer satisfaction was measured four months after training and the level of communication effectiveness was positively correlated with the level of customer satisfaction ($r = .743$, $p < .001$). Sales performance was measured ten
months after training with only the Modeling Plus Team-Building group showing an improvement (+7.9 per cent).\(^{53}\)

This study combined the best of several evaluation concepts and, perhaps most importantly, yielded meaningful and useful data for its organization.

Hayes and Williams undertook a study to measure the influence of individual differences. The study measured the change in supervisory attitudes resulting from 12.5 hours of leadership training. They found the amount of attitude change to be inversely related to age, seniority and span of control. They concluded that:

...supervisory training programs are more effective when the participants are young, relatively new to supervisory ranks, responsible for a small number of subordinate personnel, and have a short period of total service. In order to achieve program effectiveness... among the older supervisory personnel having lengthy position and total service tenure and large span of control responsibilities, a different type of program must be undertaken by training personnel.\(^{54}\)

What this "different type of program" might be, was not specified by the researchers.

Another study focusing on individual differences was one done by Schein which measured the changes in attitudes,


interests, and personality characteristics that were the results from an eight-month management training course for college graduates. Among other things, she measured the extent to which the trainees learned a more favorable attitude toward business, a greater desire to exert a leadership role, and a more considerate leadership style. She also attempted to determine if certain individual difference variables—intelligence, personality, and background—could be identified as predictors of these changes.

The course was found to have been successful in producing the expected attitude, interest, and personality changes. More importantly, the results showed that the individual difference variables—particularly the background variables—could be identified as predictors of the attitude, interest, and/or personality changes. For example, the higher the education level of the trainee, the greater was the favorable change in the trainee's attitude toward business. The researcher concluded that individual difference variables—particularly measures of background and past experience—could be used to select those trainees who will benefit most from the training. Further, research was called for to establish the validity of these results.

House, in a rather extensive review of the research on leadership training and leadership behavior, identified
three organizational factors that influenced the transfer
of training to the job. These factors were:

1. The formal authority system within the organization. This refers to the objectives, policies, and practices established by top management, within which the trainee must work.

2. The immediate superior of the trainee. This refers to his or her right to administer rewards and punishments and possessing the ability to encourage the trainee to apply principles taught in the training course.

3. The primary work group of the trainee. This refers to the expectations of peers and immediate subordinates and how they will influence the trainee's ability to successfully apply learned concepts.

Hand, Richards, and Slocum completed a study on the impact of these organizational variables. They completed a longitudinal study of a human relations course that taught a consultative approach to managing. Two experimental groups were used; one consisted of trainees who perceived their organizational climate as favoring a consultative approach to managing, while the other group of trainees viewed their organizations as less democratic and more structured. Eighteen months after the course, both experimental groups had experienced the expected changes in attitudes toward consultative management. However, only the consultative experimental group was found to have transferred the consultative approach to their job performance. The organizational variable that apparently had

influenced the post-training behavior of the trainees was the decision-making of top management with regard to salary increases and promotions. Whereas the consultative experimental group was encouraged by means of such rewards to apply the training on the job, the group of trainees from the less democratic organization was not encouraged in this manner. The researchers concluded that:

"Training courses should be designed and conducted in light of the influences of the organizational environment and every attempt should be made to insure that influential factors within that environment are identified during the assessment of training needs."

The only study that could be found in the training evaluation literature utilizing the Solomon 4-Group research design is a study done by Bunker. The investigation was designed to identify possible "pre-test" effects, but also included the impact of individual differences and their effects on training outcomes. The study evaluated the success of a basic electricity training program for telephone installer repairmen. The subjects were 131 male installers with approximately equal company experience in terms of tenure, training, and work exposure. Each were randomly assigned to one of the four conditions in the Solomon model. The pretest and posttest consisted of

equivalent forms of an objective test of basic electricity knowledge. A numerical aptitude test was also administered to divide groups into three levels of numerical aptitude (high, medium, and low). Analysis of data indicated a substantial main effect for training, a nonsignificant main effect for pretesting, and no evidence of interactions. However, a three-way ANOVA was performed to provide a test of the hypothesis that intelligence factors can act to moderate and/or mask simple pretest effects and/or interactions. A significant Numerical Aptitude × Pretest × Training interaction was obtained. It was determined that the posttest performance of subjects low in numerical aptitude was hindered by exposure to the pretest in the training condition, but was unaffected by such exposure in the control condition. Meanwhile, persons of medium aptitude were hindered by pretest exposure in the training setting, but were helped when pretested in the control setting. Subjects scoring high on numerical aptitude were unaffected by pretest exposure in either condition. These results demonstrate the complexity of the pretest contamination problem and support Clement and Campbell’s, et al., contention that the individual difference variables have a direct effect on training outcomes—in this case, obscuring simple pretest effects even when an extended control group design is carefully designed and implemented.

In interpreting his results, Bunker states the
following:

Restriction of the training impact for pretested medium and low aptitude trainees may have resulted from: (1) a misguided attention to the limited sample of the training material stimulated by the pretest experience...thus, the trainee may have focused only on developing answers to the specific items to which he or she had been exposed during pretesting..., (2) the arousal of fear of failure due to the difficulty encountered in completing the pretest, and (3) an inability to integrate pretest extractions of the material and the training course itself, or (4) a combination of the above.

A significant conclusion to be drawn from this study is that in the absence of controls for both pretesting and the contributing influence of numerical aptitude, the self-biasing aspect of the evaluation process would not have been detected. A traditional Two-Group design may have even led to the erroneous conclusion that the training program was ineffective.

Though by no means exhaustive, the studies cited indicate the importance of a carefully designed and well-thought-out evaluation design. The majority of training evaluation being done today is limited almost solely to trainee reactions, or at best, immediate on-the-job outcomes. Though these studies do serve a useful purpose, they are limited in their ability to provide meaningful data that can benefit the organizations which underwrite their expense.

The challenge faced by the evaluator is to balance the many issues regarding appropriate evaluation and implement a design that is an acceptable compromise to all interested parties, while at the same time maintaining research/evaluation integrity. The alternative, which sadly too many companies are electing, is to do none at all. Ultimately, it makes little sense for American business to spend billions of dollars on training and development activites for its employees and almost nothing to determine their effectiveness.
CHAPTER III

DESIGN AND METHODOLOGY

INTRODUCTION

This chapter is comprised of seven sections. In the first section a statement of purpose is presented. The second section outlines the major hypotheses of the study. The third section describes the experimental setting. A detailed description of the measurement instruments is found in the fourth section, and in the fifth, the procedural chronology is detailed. The last two sections include a detailed description of the subjects and the design for statistical analysis.
STATEMENT OF PURPOSE

The purpose of this study is to address the problem of how to best design and implement an evaluation of a specific training program. The one selected for this study is the Advanced Operations Course, a ten-day course offered by McDonald's Corporation to restaurant managers at its corporate training facility, Hamburger University.

As has been seen in the review of the training literature, there are many factors to consider in designing an effective training evaluation study. Among them are:

Audience the Evaluation Should Serve - The management and faculty of Hamburger University is the target audience for this evaluation. It is hoped that the results of this evaluation will provide specific information regarding the effectiveness of the present curriculum and may indicate ways it can be improved.

Criteria of Evaluation - Kilpatrick's four levels of training evaluation: knowledge, attitude, behavior, and results, offer the most comprehensive criteria on which to base any evaluation of a training program. However, the difficulties associated with the attempt to gather meaningful data on results measures in this study influenced the researcher to limit the criteria for this evaluation to knowledge, attitude, and behavior outcomes.
Selecting the Evaluators - The need for additional evaluators was particularly indicated in attempting to gather useful data regarding behavior changes of those receiving the training. The most logical choices for these evaluators were the immediate supervisor and subordinates of each restaurant manager. The supervisor and subordinates were best able to evaluate the trainee's on-the-job behavior before and after training.

Selecting the Research Design - The design selected for this study was the Solomon 4-Group design. Campbell and Stanley cite as its chief merit the use of two experimental and two control groups, thus effectively controlling for individual differences and allowing for the determination of the actual effect of the experimental treatment. In addition, the design enables the study to identify any possible "pretest sensitization," an effect that can undermine the true strength of the treatment effect. This design also effectively addresses both the issues of internal and external validity.

Internal validity asks the question: To what extent is this design able to detect and control for alternative explanations of the evaluation results which might otherwise contaminate the genuine effects of the training program? The selected design specifically addresses this issue through the use of pretests designed to measure the pre-training levels on the relevant variables and by
randomly assigning subjects to the experimental and control
groups. Some additional benefits of pretesting include:

1. Pretesting generates greater information relative
to the individual strengths and weaknesses of
trainees.
2. If relevant individual differences on critical
variables are not distributed evenly across
comparison groups, this will become apparent
from the pretest scores.
3. Pre-training differences may be detected among
subgroups which permit examining the effects of
training in the context of pre-existing
differences. 58

The present study uses three separate pretests measuring
cognitive, affective, and behavioral course objectives.

External validity issues are concerned with the
capability to generalize the impact of the training as
measured by cognitive and affective measures (learning and
attitude) to specific behavior change on the job. Through
the use of a Management Behavior Index, the study attempts
to identify the frequency of specific on-the-job behaviors
as perceived by the trainee's supervisor and subordinates.

Subjects were assigned to one of four experimental
conditions:

PTP = Pretested, Trained, Posttested
PUP = Pretested, Untrained, Posttested
UTP = Unpretested, Trained, Posttested
UUP = Unpretested, Untrained, Posttested

58 Bunker and Cohen, op. cit., p. 6.
Since the study is investigating three separate outcome measures, three different "Pretests/Posttests" will be used:

1. Knowledge Outcomes = A.O.C. Pretest
2. Attitude Outcomes = Attitude Survey Index
3. Behavior Outcomes = Management Behavior Index

The analytic paradigm for the design is found in Table 1, where "X" represents the Advanced Operations Course (the experimental treatment) and the Pretest/Posttests consist of the learning, attitude, and behavioral measures.
Table 1  
Analytic Paradigm for the Study

<table>
<thead>
<tr>
<th>Group #</th>
<th>Pretest</th>
<th>Independent Variable</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTP (1)</td>
<td>Y1</td>
<td>X</td>
<td>Y2</td>
</tr>
<tr>
<td>PUP (3)</td>
<td>Y3</td>
<td></td>
<td>Y4</td>
</tr>
<tr>
<td>UTP (2)</td>
<td></td>
<td>X</td>
<td>Y5</td>
</tr>
<tr>
<td>UUP (4)</td>
<td></td>
<td></td>
<td>Y6</td>
</tr>
</tbody>
</table>
HYPOTHESES

The main hypothesis for this study is aimed at assessing the effectiveness of the Advanced Operations Course and is stated as follows:

1. There will be no statistically significant difference between Trained and Untrained subjects as measured by the change in Pre versus Post measures in knowledge, attitude, and behavior.

As a result of the Solomon 4-Group research design, several other hypothesized relationships will be tested.

2. There will be no statistically significant difference between Posttest versus Pretest measures of Pretested/Trained subjects (Y2 vs Y1).

3. There will be no statistically significant difference between Posttest measures of Pretested/Trained subjects versus Pretested/Untrained subjects (Y2 vs Y4).

4. There will be no statistically significant difference between Posttest measures of Unpretested/Untrained subjects versus Unpretested/Trained subjects (Y5 vs Y6).

5. There will be no statistically significant difference between Posttest measures of Unpretested/Trained subjects and Pretest measures of Pretested/
Untrained subjects (Y5 vs Y3).

6. There will be no statistically significant difference between Pretest measures of Trained subjects and Pretest measures of Untrained subjects (Y1 vs Y3).

7. There will be no statistically significant difference between Posttest measures of Pretested/Trained subjects and Unpretested/Trained subjects (Y2 vs Y5).

8. There will be no statistically significant difference between Posttest measures of Pretested/Untrained subjects and Unpretested/Untrained subjects (Y4 vs Y6).

The hypothesized relationship between "Need for Achievement" and "Academic Performance" during the training experience will be tested by the following hypothesis:

9. There will be no statistically significant relationship between the "Need for Achievement" measure and the "Total Points Earned" or "Academic Achievement" measure for all subjects.

The hypothesis that will test the relationship between individual differences among subjects and their impact on training outcomes is:

10. There will be no statistically significant relationship between the variables Education and Type and Post measures of knowledge, attitude, and behavior or the "Total Points Earned" measure for all subjects. To measure the effectiveness of training over time, a final hypothesis
There will be no significant difference in Post training measures of knowledge, attitude, and behavior taken immediately after training versus forty-five days after training.

Operational definitions are as follows:

**Independent Variables**

**EDUCATION**
Formal education expressed in number of years of formal schooling.

**TYPE**
Form of subject employment, either in a company-owned restaurant and thus an employee of McDonald's Corporation, or in a franchise restaurant and thus an employee of an Owner Operator. Also includes Owner Operators; Registered Applicants (new operators, not yet assigned a restaurant); Staff (members of McDonald's corporate and field staff); and International (students from countries outside the continental U.S.).

**NEED FOR ACHIEVEMENT**
A personality variable which refers to a subject's response to situations where some standard of excellence can be applied to his or her behavior; a motivation to be the best or
to excel. It will be measured using the Edwards Personal Preference Schedule which is scored on a range of 1 to 100.

PRETESTS:

KNOWLEDGE

Score on a fifty-item multiple-choice objective test. Items cover the material that is taught in the Advanced Operations Course.

ATTITUDE

The per cent of favorable (agree) responses on a twenty-nine item Attitude Survey designed to measure subject's attitudes about McDonald's Corporation and the job of restaurant manager.

BEHAVIOR

The composite score of three separate indexes complete by each subject's immediate superior and two subordinates. Index consists of twenty behavioral statements describing behaviors directly related to course objectives.

Dependent Variables

POSTTESTS:

KNOWLEDGE

The same fifty-item multiple-choice test as the pretest.

ATTITUDE

The per cent favorable responses on the same
twenty-nine item attitude survey as the
pretest.

BEHAVIOR

A composite score of three indexes completed
by the same three individuals on the subject
after training. The same index is used as the
pretest.

TOTAL
POINTS
EARNED

Total number of correct responses by a subject
on six separate fifty-item multiple-choice
tests and a faculty evaluation of one hundred
points (four hundred total points possible).
EXPERIMENTAL SETTING

The Advanced Operations course is a two-week course conducted by McDonald's Corporation at its management training center, Hamburger University. The course is targeted at individuals who are about to become or who have just been promoted to restaurant manager. McDonald's has a well-defined and standardized training program used to train its restaurant management which begins at the "manager trainee" level. The program consists of a combination of in-restaurant and on-the-job training as well as classroom instruction which takes place over a two to three year period and culminates with the opportunity to attend the Advanced Operations Course at Hamburger University. Attendees have already completed this prior training and include students from throughout the United States and around the world. Students also represent the two "types" of McDonald's restaurants, company-owned and franchised. The Advanced Operations Course has been offered by McDonald's Corporation since 1961 and presently numbers over twenty thousand graduates. The curriculum has evolved over the years to reflect the changing nature of the McDonald's system and presently requires a staff of twenty-three faculty and support staff to conduct the
fourteen to sixteen classes offered each year. The curriculum has been designed to reinforce the trainee's prior restaurant operations training and to provide additional management skills targeted at increasing effectiveness in his/her role as a restaurant manager. During the course, trainees are housed in McDonald's Lodge, which is located adjacent to the training facility and specifically constructed for this purpose. A typical class day consists of seven to eight hours of direct instruction. Instructional techniques include lecture, role play, case study, demonstration, workshops, and in the case of equipment training, hands-on work. Topics addressed within the curriculum include team-building, goal-setting, problem-analysis and decision-making, personnel practices, training methods for hourly employees and subordinate management, communications and motivational techniques (transactional analysis), paperwork and restaurant controls, and history (a motivational review of McDonald's corporate history). Nearly 40 per cent of the A.O.C. curriculum is devoted to training on the various pieces of McDonald's restaurant equipment including kitchen equipment (fryers, grills, ice machines, beverage systems, toasters, steamers, timers, shake and sundae machines, etc.) and rooftop equipment (exhaust fans, heating, ventilating and air-conditioning equipment, etc.). This training is facilitated in part by fully operational and complete
equipment labs where trainees individually receive the hands-on opportunity to review parts identification, basic planned maintenance procedures, equipment adjustment procedures and basic troubleshooting and minor repair procedures. Throughout the training, trainees complete six separate fifty-item multiple-choice tests designed to measure learning. They must also complete other assigned homework and workshop activities. Though all tests are scored and academic results tabulated, no one "fails." A final course evaluation is completed by the faculty on each student, which is sent back to the student's immediate supervisor. All students graduate and receive a special diploma as part of a special graduation ceremony.
Four separate instruments were used to measure Knowledge, Attitude, Behavior, and Need for Achievement. Since both a pre and post score was required, it was decided to use the same instrument for both the pre and post measure. As has been discussed, the selected design provides a strong measure for detecting the possibility of a pretest effect.

Knowledge Test

This measure was the fifty-item multiple-choice test known as the A.O.C. Pretest. Designed by the Hamburger University faculty, it tests the cognitive objectives of the course through a representative sample of test items which are based on a test specification chart. Test items cover the topics of teambuilding, goal-setting, problem analysis and decision-making, personnel practices, training methods, communication and motivational techniques and equipment - its basic operation, planned maintenance requirements, calibration and basic troubleshooting. The present fifty-items are drawn from a computerized test bank of over two thousand items which are constantly evaluated each time they are utilized on a test. Two measures are
monitored for each test item: a Discrimination Index and a Difficulty index. The Discrimination Index measures how well the test discriminates between those who "know" and those who "do not know." It is determined by taking completed tests and separating them into two categories - High Correct (students getting a high percentage of all test items correct), and Low Correct (students getting a low percentage of all items correct). A certain number of High and Low Correct are randomly selected (normally twenty) and the following formula is used to calculate a Discrimination Index for each item:

\[ \frac{HC - LC}{HC + LC} \]

HC is equal to the number of High Correct test papers getting the item correct. LC is the number of Low Correct test papers getting the item incorrect and it is subtracted from the High Correct number. The resulting difference is divided by the sum of the High and Low Correct. An acceptable range that has been established by the Hamburger University faculty for this index is 0.15 to 0.25. Utilizing a specially developed computer program to do the calculations, this index is updated each time a test is administered and scored. The Discrimination Index for the present A.O.C. Pretest is 0.20.

The Difficulty Index is defined as how well the item consistently measures the desired objective. This index is simply computed by dividing the number of correct responses
to an item by the total number of correct and incorrect responses. The acceptable range that has been established is 0.40 to 0.95. The A.O.C. Pretest has a Difficulty Index of 0.80.

Attitude Survey

This index was designed by the researcher and was intended to assess a student's attitude about him or herself, McDonald's Corporation, his/her job, and his/her training experiences. The company has been using this type of instrument for several years to assess attitudes of employees at all levels of the organization: hourly employees, store management, and middle management and corporate staff. Using these existing instruments as a guide, a new instrument was developed. It was called an "Opinion Survey" so students would not confuse it with the Attitude Surveys. Each statement on the survey describes an attitude or feeling about the company, one's own job, the work environment, level of training received, etc. Respondents are asked to make one of three responses: "agree" or yes, "?" or not sure, and "disagree" or no. The survey is scored by calculating the number of "agree" responses and expressing them in a "Per Cent Favorable Response" percentage.

Several statements found on the existing survey were incorporated into the new survey along with some new
statements resulting in a twenty-nine statement survey. Once finalized, the survey was submitted to the individual in the corporate Personnel Department who was originally responsible for developing the pre-existing Attitude Surveys. His editorial critique was solicited in finalizing each of the statements. The final instrument was then pilot-tested with two A.O.C. classes. Since fourteen of the twenty-nine Opinion Survey statements were identical to the original Attitude Survey instrument, it was possible to compare the overall Per Cent Favorable Response rates on the new instrument to the national norms for the fourteen individual statements. In the two pilot test groups, the per cent favorable response rate on each of the fourteen statements was nearly identical. The pilot test results indicated that the new instrument was strongly correlated to the pre-existing instrument used on a national level by McDonald's Corporation to measure employee attitudes and that it was a reliable indicator of attitude \( r = .91 \). Hereafter, the new instrument will be referred to as the Attitude Survey.

Behavior Index

The Management Behavior Index was co-developed by the researcher, Hamburger University faculty, and an outside consultant who works with the Hamburger University staff. The index was developed in the following manner:
1. All learner objectives for the Advanced Operations Course were assembled and carefully reviewed.

2. All learner objectives that were stated in behavioral terms (perform, demonstrate, show, complete, etc.) were selected. Those objectives remaining assessed either the cognitive or affective domain and were thus measured by the other two instruments.

3. From these remaining objectives, a list of statements was constructed with each stated in terms of observable manager behaviors.

4. This list was submitted to Hamburger University faculty and the outside consultant for their review and critique. Their respective subject-matter expertise helped to improve the wording of the statements to insure they accurately described the behavior intended by each of the course objectives.

5. After final editorial changes were made in the statements, the instrument was pilot-tested with a group of twenty A.O.C. students.

It was decided that an open-ended question would also provide an additional data source and would serve particularly well in helping to measure pre versus post results. The instrument, designed to be completed by the subject’s immediate supervisor, a peer, and a subordinate, (a total
of three per subject), is completed anonymously. The respondent is asked to read each management behavior statement and, based on his/her recent working experience with this individual, to select one of five responses relating to frequency: always, often, sometimes, occasionally, or never. A numeric scale was also listed along with these descriptors to make the assessment easier: 10, 7, 5, 3, and 1 respectively.

To assess the reliability of this instrument, two separate correlation coefficients were calculated using the twenty subjects in the original pilot test. The Split-Half method was first utilized, yielding a correlation coefficient of $r = .88$. Since three separate individuals completed instruments on each subject, it was also possible to calculate inter-rater reliability. The inter-rater reliability coefficient was $r = .83$.

Need for Achievement

After a review of the literature and consultation with a Loyola faculty expert on the subject, it was decided that the Edwards Personal Preference Schedule was the best standardized instrument available to measure the Need for Achievement variable. It consists of 225 paired statements and asks the student to select either the "a" or "b" statement, based on which is most preferred or least objectionable. Once scored, it yields results on fourteen
separate personality variables. The correlation coefficient for the Need for Achievement variable is $r = .74$.

After carefully reviewing the Index, it was felt that the questions assessing the personality variable "heterosexuality" were too controversial and potentially offensive to use in this setting. Thus the researcher modified the test booklet and answer key, and eliminated those thirty items dealing with this variable. During test administration, a careful explanation was provided to insure accuracy of completion. Since scoring is based on gender and since results were to be correlated with academic achievement, subjects were asked to record their student numbers on the answer sheet. All students were offered the option of not completing the test at all or completing it anonymously. Twenty-three students (9 percent) elected not to complete the test.
PROCEDURAL CHRONOLOGY

The original concept of this study was developed in the fall of 1983. As preliminary ideas for structuring the study were being developed, a written questionnaire was sent to twenty-two present and former members of the Hamburger University instructional staff soliciting their ideas and input on such factors as:

What are the significant individual difference factors that affect academic performance?

What on-the-job behaviors are most impacted by having attended A.O.C.?

What are the most effective ways to measure "in-restaurant" results of A.O.C.?

Three important decisions were made as a result of this initial research:

1. Several of the H.U. faculty felt that motivation brought by students to the course was the most powerful determinant of academic achievement.

2. Though it had originally been the intention of the researcher to include a "Results" measure along with the Knowledge, Attitude, and Behavior measures, the difficulties associated with trying to arrive at appropriate in-restaurant measures
that could be directly related to A.O.C. proved nearly impossible. Thus, it was decided not to include the "Results" measure in the study.

3. To establish which of the several individual difference factors were significantly related to academic performance, a preliminary study was done. Total Points Earned is a summed total representing academic performance in the course. Thus this variable was compared to sixteen separate individual difference variables. There were 115 students in this study. A separate analysis of variance was done on Total Points Earned with each of the sixteen variables. Four were found to be significant at the .01 level as seen in Table 2. They are: Entex, a fifty-item multiple-choice test designed to measure knowledge of content learned in prior McDonald's training; Pretest and Posttest - identical versions of a fifty-item multiple-choice test designed to measure the course objectives of A.O.C.; and Education - the number of years of schooling. The variable Type, whether the student was an employee of a company-owned or franchised McDonald's restaurant, came close to the .05 level, at .07.

Once the concept for the research design was finalized, the researcher met with the Dean and Assistant Dean
Table 2

Analysis of Variance of Total Points Earned by Individual Difference Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>F Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENTEX</td>
<td>40.15</td>
<td>0.0001</td>
</tr>
<tr>
<td>PRETEST</td>
<td>30.77</td>
<td>0.0001</td>
</tr>
<tr>
<td>POSTTEST</td>
<td>26.01</td>
<td>0.0001</td>
</tr>
<tr>
<td>EDUC</td>
<td>3.09</td>
<td>0.01</td>
</tr>
<tr>
<td>TYPE</td>
<td>3.34</td>
<td>0.07</td>
</tr>
<tr>
<td>IOC</td>
<td>1.53</td>
<td>0.22</td>
</tr>
<tr>
<td>MCDEXP</td>
<td>1.47</td>
<td>0.20</td>
</tr>
<tr>
<td>AGE</td>
<td>1.37</td>
<td>0.25</td>
</tr>
<tr>
<td>JOB</td>
<td>1.30</td>
<td>0.31</td>
</tr>
<tr>
<td>BOC</td>
<td>1.20</td>
<td>0.28</td>
</tr>
<tr>
<td>TIME</td>
<td>1.19</td>
<td>0.25</td>
</tr>
<tr>
<td>SEX</td>
<td>1.07</td>
<td>0.30</td>
</tr>
<tr>
<td>FFEXP</td>
<td>0.49</td>
<td>0.91</td>
</tr>
<tr>
<td>SOURCE</td>
<td>0.47</td>
<td>0.50</td>
</tr>
<tr>
<td>MDP3</td>
<td>0.45</td>
<td>0.50</td>
</tr>
<tr>
<td>AEC</td>
<td>0.06</td>
<td>0.81</td>
</tr>
</tbody>
</table>
of Hamburger University to overview plans for data collection. Two classes, class number reference 352 and 354, were selected from the sixteen classes offered by Hamburger University in the calendar year. The data collection process was complicated by the four groups called for in the Solomon 4-Group design and the need to collect pre and post-training measures on knowledge, attitude, and behavior. Behavior instruments were to be handled by having a subject select three individuals: his/her immediate supervisor and two subordinates (or a peer and a subordinate). An instrument, a cover letter explaining how to complete the instrument, and a pre-postage paid return envelope was mailed to each of the individuals the subject selected. Though both classes would ultimately complete the A.O.C. training, Class 352 was selected as the "trained" group and randomly divided into "Pretested" and "Unpretested" groups, thus creating the PTP and UTP groups. Knowledge and Attitude Pretests were administered to these two groups on the first day of the Advanced Operations Course.

Class 354 was treated as the "Untrained" group and was also randomly divided into "Pretested" and "Unpretested" groups, thus creating the PUP and UUP groups. Knowledge, Attitude, and Behavior Pretests were collected on this group by mail prior to class, and "Posttest"
measures were collected on the first day of A.O.C. In other words, Class 354 was "Untrained" because they were pretested by mail and posttested prior to any formal instructional activities on the day Class 354 was begun.

The following chronology of data collection was observed:

4/24/84 - Class 352 begins at 8:00 A.M., 138 students. Randomly selected half the class (seventy) as the PTP group and administered the Knowledge and Attitude Pretests. Subjects mailed Behavior instruments with signed cover letters to the three selected individuals requesting return within thirty days.

4/30/84 - Administered the Need for Achievement measure (the Edwards Personal Preference Test) to all students.

5/4/84 - Last day of Class 352. Administered Knowledge and Attitude Posttests to entire class—both the PTP and UTP groups.

5/7/84 - Mailed Knowledge, Attitude, and Behavior Pretests to half (sixty students) in Class 354 (the PUP group) with detailed instructions on how to complete. Subjects were requested to mail within thirty days or actually bring the completed instruments with them to A.O.C.
6/1/84 - Mailed Knowledge, Attitude, and Behavior Posttest measures to all members (both the PTP and UTP groups) of Class 352 with detailed instructions on how to complete and a return request of thirty days.

6/11/84 - Class 354 begins at 8:00 A.M.- 136 students. Administered the Knowledge and Attitude Posttest measures to both the PUP and UUP groups.

6/16/84 - Administered the Need for Achievement measure to all of Class 354. Prepared Behavior instruments (Posttests) for mailing to the three selected individuals for each student.

This chronology provided for the necessary data collection to generate the pre and post measures required for the Solomon 4-Group design. In addition, the 6/1/84 mailing of Knowledge and Attitude measures to the Trained group allowed for the collection of data on training outcomes forty-five days after training.
DESCRIPTION OF THE SAMPLE

Subjects were randomly sampled to be included in this study in the following manner. Hamburger University offers sixteen Advanced Operations Courses each calendar year. In reviewing the 1984 calendar, the researcher sought two class dates with the maximum time between them to facilitate the data collection that was to be done through the mail. Classes 352 and 354 were selected based on a five-week gap between them. Class 352 was selected as the "Trained" group, since it occurred first and all pretest data could be collected on the first day of class. Class 354 was selected to be the "Untrained" group. Although Class 354 would ultimately attend the course, pretest measures were completed by mail and posttest measures were completed on the first day of class, prior to training. Students were randomly assigned to the two experimental groups within each class through the use of an alphabetized student roster and alternate assignment to each group. Once this random assignment was completed, the resulting groups were reviewed to identify non-target students. A non-target student was anyone who was not a Manager or Assistant Manager of a McDonald's restaurant. It included middle-management staff, owner operators, and registered
applicants (new owner operators who have not yet opened their restaurant). Where possible, these individuals were reassigned, so as to balance the groups in number of non-target students. This technique produced a systematic random sample.  

As has been previously noted, the Advanced Operations Course is the final step in a highly systematic training program the McDonald's Corporation provides. The entire training system consists of a series of in-restaurant, on-the-job training experiences coupled with classroom instruction. Entrance requirements for admission to A.O.C. have been set by Hamburger University so students will meet the following criteria:

1. First Assistant by Job Title.

2. Completed all appropriate prior in-restaurant training (the first three volumes of the McDonald's Management Development Program) and all prior classroom training.

3. Achieved a passing score (80 per cent or better) on a field-administered Entrance Exam. This test measures knowledge and skills the student should already possess as a result of completing the training referred to in #2.

Occasionally, Hamburger University will waive one or more of these criteria at the request of field management. This is done normally to assure the presence of an H.U. graduate in every new McDonald's restaurant at the time it is ready to open. Because the time between when a student registers and the time a student attends A.O.C. can be as long as six months, it is frequently the case that although the student was a First Assistant when originally registered, he/she may have since been promoted to Store Manager.
DESIGN OF THE ANALYSIS

The Solomon 4-Group design selected for this study allows for several separate statistical analyses. The major analytic paradigm consists of a 2 X 2 factorial ANOVA with the posttest scores of the four groups in the cells. As can be seen below, the design enables the researcher to analyze several different comparisons and possible interactions:

<table>
<thead>
<tr>
<th></th>
<th>UNTRAINED</th>
<th>TRAINED</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRETESTED</td>
<td>Y4</td>
<td>Y2</td>
</tr>
<tr>
<td>UNPRETESTED</td>
<td>Y6</td>
<td>Y5</td>
</tr>
</tbody>
</table>

Separate comparisons will be made on each of the following hypothesized relationships among the mean posttest scores by using separate "t" tests:

\[
\begin{align*}
Y2 & > Y1  \\
Y2 & > Y4  \\
Y5 & > Y6  \\
Y5 & > Y3  \\
Y1 & = Y3  \\
Y2 & = Y5  \\
Y4 & = Y6
\end{align*}
\]
These analyses will be done on each of the three outcome measures being evaluated: Knowledge, Attitude, and Behavior.

To test the hypothesized relationship between Need for Achievement and Academic Achievement (Total Points Earned), a 3 x 2 x 2 ANOVA of posttest scores partitioned by Need for Achievement will be done. Its analytic paradigm will contain posttest scores in the cells. Possible interactions between Need for Achievement and Trained or Untrained and Pretested or Unpretested will all be assessed for significance.

<table>
<thead>
<tr>
<th></th>
<th>TRAINED</th>
<th>UNPRETESTED</th>
<th>PRETESTED</th>
<th>UNPRETESTED</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI N.A.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MED N.A.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LO N.A.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER IV

RESULTS

INTRODUCTION

This chapter analyzes the results of the study. It begins by examining the issue of comparability of groups, the overall sample with Hamburger University Group norms and then the homogeneity of the four sample subgroups. A sequential treatment of each of the eleven hypotheses is undertaken and statistical results related to each are presented.
INITIAL COMPARABILITY OF GROUPS

As previously discussed in Chapter III, the method of selecting A.O.C. Classes 352 and 354 as the sample raises the issue of comparability of groups. In addition, the validity of the analytical procedures used with the Solomon 4-Group design is, in part, dependent upon initial comparability of the subjects in the various groups relative to the content of the course. The process used by Hamburger University to register students to attend class insures that this comparability generally exists. Certain "entrance requirements" specify that students have completed levels of McDonald's field training, both classroom and on-the-job. Students must demonstrate a criterion knowledge level as measured by a passing score (80 per cent) on a field-administered Entrance Exam. Table 3 illustrates the strong similarities between the sample selected for this study (Classes 352 and 354) and existing H.U. norms, among several demographic variables. In addition, Pretest mean comparisons were done on Trained versus Untrained groups on Knowledge, Attitude and Behavior outcome measures. These comparisons attempted to detect if there were any significant differences in the Pretest scores between these groups. As indicated in Table 4, no significant
Table 3

Comparison of Demographic Data Between Hamburger University's Student Population Norms and Sample Groups

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>NORMS</th>
<th>SAMPLE (N=263)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCOPCO vs LICENSEE</td>
<td>27% VS 73%</td>
<td>33.5% VS 66.5%</td>
</tr>
<tr>
<td>MALE vs FEMALE</td>
<td>57% VS 43%</td>
<td>65% VS 35%</td>
</tr>
<tr>
<td>JOB TITLE:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STORE MANAGERS</td>
<td>51.4%</td>
<td>49%</td>
</tr>
<tr>
<td>FIRST ASSISTANTS</td>
<td>47.7%</td>
<td>38%</td>
</tr>
<tr>
<td>REGISTERED APPLICANTS</td>
<td>.9%</td>
<td>5%</td>
</tr>
<tr>
<td>OTHER</td>
<td></td>
<td>8%</td>
</tr>
<tr>
<td>EDUCATION (TOT. YEARS)</td>
<td>14.2 YEARS</td>
<td>13.8 YEARS</td>
</tr>
<tr>
<td>PRETEST SCORE</td>
<td>54.20%</td>
<td>55.38%</td>
</tr>
<tr>
<td>TOTAL POINTS EARNED</td>
<td>674</td>
<td>679</td>
</tr>
</tbody>
</table>
# Table 4

Pretest Mean Comparisons Testing For Differences In Trained and Untrained Groups on Knowledge (K), Attitude (A), and Behavior (B) Measures.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N</th>
<th>MEANS</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained vs Untrained</td>
<td>K</td>
<td>63 vs 43</td>
<td>54.63 vs 56.47</td>
<td>.88</td>
</tr>
<tr>
<td>(PTP &amp; UTP) vs (PUP &amp; UUP)</td>
<td>A</td>
<td>63 vs 42</td>
<td>85.27 vs 89.33</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>54 vs 51</td>
<td>6.82 vs 7.12</td>
<td>1.35</td>
</tr>
</tbody>
</table>
pretest differences were detected as a function of group assignment. These comparisons establish the initial comparability of the groups.
MAIN RESULTS

The asymmetrical nature of the Solomon 4-Group design precludes utilizing all six sets of scores (pretests and posttests) in a single analysis (Campbell and Stanley, 1963). The accepted analytical procedure is to disregard the pretest except as an additional treatment classification variable (pretested or not pretested) and to conduct a 2 X 2 analysis of variance on the four sets of posttest scores. This analysis was further complicated by the existence of three types of posttest scores: Knowledge, Attitude, and Behavior. Six sets of means (two pretest and four posttest) were used for each type of outcome measure (Knowledge, Attitude, and Behavior) to evaluate the effectiveness of the training and to determine the extent of evaluation contamination due to pretesting. The means, standard deviations, and sample sizes of the pretest and posttest scores of the four groups for each of the three outcome measures appear in Table 5.

Main Effect of Training. The efficacy of training on Knowledge, Attitude, and Behavior outcome measures was tested in four separate comparisons of pretest and posttest means. These mean comparisons represent the hypothesized relationships contained in Hypotheses 1-4.
Table 5

Means, Standard Deviations, and Sample Sizes of Pretest and Posttest Scores

<table>
<thead>
<tr>
<th>Group</th>
<th>Knowledge</th>
<th></th>
<th>Attitude</th>
<th></th>
<th>Behavior</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>1*</td>
<td>X=54.64</td>
<td>X=91.57</td>
<td>X=85.27</td>
<td>X=90.86</td>
<td>X=6.81</td>
<td>X=7.57</td>
</tr>
<tr>
<td></td>
<td>SD=8.87</td>
<td>SD=9.72</td>
<td>SD=10.59</td>
<td>SD=8.28</td>
<td>SD=1.19</td>
<td>SD=1.12</td>
</tr>
<tr>
<td></td>
<td>N = 63</td>
<td>N = 65</td>
<td>N = 63</td>
<td>N = 65</td>
<td>N = 53</td>
<td>N = 19</td>
</tr>
<tr>
<td>2*</td>
<td>N.A.</td>
<td>X=91.43</td>
<td>N.A.</td>
<td>X=91.77</td>
<td>N.A.</td>
<td>X=7.63</td>
</tr>
<tr>
<td></td>
<td>SD=8.84</td>
<td>N.A.</td>
<td>SD=7.23</td>
<td>N.A.</td>
<td>SD=1.21</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td>N = 67</td>
<td>N = 66</td>
<td>N = 66</td>
<td>N = 66</td>
<td>N = 33</td>
<td>N = 33</td>
</tr>
<tr>
<td>3*</td>
<td>X=56.47</td>
<td>X=58.00</td>
<td>X=89.33</td>
<td>X=89.30</td>
<td>X=7.12</td>
<td>X=7.35</td>
</tr>
<tr>
<td></td>
<td>SD=12.52</td>
<td>SD=9.94</td>
<td>SD=10.77</td>
<td>SD=10.05</td>
<td>SD=1.10</td>
<td>SD=1.34</td>
</tr>
<tr>
<td></td>
<td>N = 43</td>
<td>N = 71</td>
<td>N = 42</td>
<td>N = 71</td>
<td>N = 51</td>
<td>N = 53</td>
</tr>
<tr>
<td>4*</td>
<td>N.A.</td>
<td>X=55.12</td>
<td>N.A.</td>
<td>X=89.20</td>
<td>N.A.</td>
<td>X=6.93</td>
</tr>
<tr>
<td></td>
<td>SD=13.15</td>
<td>N.A.</td>
<td>SD=6.76</td>
<td>N.A.</td>
<td>SD=1.04</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td>N = 57</td>
<td>N = 55</td>
<td>N = 55</td>
<td>N = 55</td>
<td>N = 36</td>
<td>N = 36</td>
</tr>
</tbody>
</table>

* Group 1 = Pretested, Trained, Posttested (N=65)
Group 2 = Unpretested, Trained, Posttested (N=68)
Group 3 = Pretested, Untrained, Posttested (N=73)
Group 4 = Unpretested, Untrained, Posttested (N=57)
The results of these mean comparisons are illustrated in Table 6. The posttest means of those trained were significantly higher than the comparable pretest mean (Comparisons 1 and 4/Hypotheses 2 and 5). Additionally, the mean posttest scores of trained subjects were significantly higher than those of the untrained subjects for most of the outcome measures (Comparisons 2 and 3/Hypotheses 3 and 4). The exceptions are the Attitude and Behavior posttest means for Pretested-Trained-Posttested subjects versus the Pretested-Untrained-Posttested subjects. The overall results of these comparisons demonstrate a positive main effect due to the Advanced Operations Course (A.O.C.). Null hypotheses 2, 4, and 5 are thus rejected. Hypothesis 3 is rejected for the Knowledge outcome measure, but cannot be rejected for the Attitude and Behavior measure. At this stage of the analysis, the lack of significance could be suggesting that there is some type of pretest effect for trained versus untrained subjects. The differences in pretest and posttest performance of the experimental (Trained) and control (Untrained) groups are presented graphically in Figures 1, 2, and 3. Pretest estimates plotted for the unpretested groups (UPT and UUP) were derived by taking the average of the pretest scores of the pretested groups (PTP and PUP).

Main Effect of Pretesting. The main effect of pretesting for trained subjects was evaluated by comparing the
Table 6

Pretest-Posttest Mean Comparisons Testing for the Main Effect of Training on Knowledge (K), Attitude (A), and Behavior (B) Outcomes

<table>
<thead>
<tr>
<th>Comparison</th>
<th>K</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. PTP Posttest &gt; K-91.57 vs 54.64</td>
<td>38.06</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td>PTP Pretest</td>
<td>A-90.86 vs 85.27</td>
<td>3.08</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>(Y2 &gt; Y1)</td>
<td>B- 7.57 vs 6.81</td>
<td>2.73</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>2. PTP Posttest &gt; K-91.57 vs 58.00</td>
<td>19.89</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td>PUP Posttest</td>
<td>A-90.86 vs 89.30</td>
<td>.99</td>
<td>N.S.</td>
</tr>
<tr>
<td>(Y2 &gt; Y4)</td>
<td>B- 7.57 vs 7.35</td>
<td>.70</td>
<td>N.S.</td>
</tr>
<tr>
<td>3. UTP Posttest &gt; K-91.43 vs 55.12</td>
<td>17.72</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td>UUP Posttest</td>
<td>A-91.77 vs 89.20</td>
<td>2.02</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>(Y5 &gt; Y6)</td>
<td>B- 7.63 vs 6.93</td>
<td>2.60</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>4. UTP Posttest &gt; K-91.43 vs 56.47</td>
<td>12.49</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td>PUP Pretest</td>
<td>A-91.77 vs 89.33</td>
<td>3.08</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>(Y5 &gt; Y3)</td>
<td>B- 7.63 vs 7.12</td>
<td>1.70</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>
Figure 1. Pretest to Posttest Changes in Mean Score Performance on Knowledge Measures

* This point represents the estimated Pretest Means for the Unpretested groups obtained by averaging the scores of the Pretested groups.
Figure 2. Pretest to Posttest Changes in Mean Score Performance on Attitude Measures.

* This point represents the estimated Pretest mean for the Unpretested groups obtained by averaging the scores of Pretested groups.
Figure 3. Pretest to Posttest Changes in Mean Score Performance on Behavior Measures.

* This point represents the estimated Pretest mean for the Unpretested groups obtained by averaging the scores of Pretested groups.
posttest mean of the Pretested-Trained group (PTP) with that of the Unpretested-Trained group (UTP), Hypothesis 7. A similar comparison for Untrained subjects was made between the means of the Pretested-Untrained group (PUP) and the Unpretested-Untrained group (UUP), Hypothesis 8. The results of these two comparisons for the three outcome measures are presented in Table 7.

Failure to reject the null hypothesis of equality of posttest scores in comparison (1) for all three outcome measures indicates that pretesting had no significant impact on the posttest performance of the trained subjects. Similarly, the results of comparison (2) indicate an absence of pretest effects in the posttest performance of untrained subjects.

**Pretest X Training Interaction.** An estimate of Pretest X Training interaction effects was provided by comparing the main Pretest effect for trained subjects (PTP Posttest - UTP Posttest) with the Pretest effect for untrained subjects (UUP Posttest - PUP Posttest). This calculation was done for all three outcome measures. A difference between the two Pretest main effects would have been suggestive of a Pretest X Training interaction in that it would indicate that the pretest had differential effects on posttest performance, depending on whether or not the training had been received. Table 8 contains the results. Differences were -2.74, -0.19, and +0.36 for Knowledge,
Table 7
Pretest-Posttest Mean Comparisons Testing for the Main Effect of Pretesting on Knowledge (K), Attitude (A), and Behavior (B) Outcome Measures.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>X</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) PTP Posttest =</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UTP Posttest</td>
<td>91.57 vs 91.43</td>
<td>K = 0.08</td>
<td>N.S.</td>
</tr>
<tr>
<td></td>
<td>90.86 vs 91.77</td>
<td>A = 0.67</td>
<td>N.S.</td>
</tr>
<tr>
<td></td>
<td>7.57 vs 7.63</td>
<td>B = 0.19</td>
<td>N.S.</td>
</tr>
<tr>
<td>(2) PUP Posttest =</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UUP Posttest</td>
<td>58.00 vs 55.12</td>
<td>K = 1.37</td>
<td>N.S.</td>
</tr>
<tr>
<td></td>
<td>89.30 vs 89.20</td>
<td>A = 0.06</td>
<td>N.S.</td>
</tr>
<tr>
<td></td>
<td>7.35 vs 6.93</td>
<td>B = 1.67</td>
<td>N.S.</td>
</tr>
</tbody>
</table>
Table 8

Pretest Effect On Trained vs Untrained Subjects by Subtraction of Posttest Means for Knowledge (K), Attitude (A), and Behavior (B) Outcome Measures.

<table>
<thead>
<tr>
<th>Mean Comparisons</th>
<th>Outcome</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>((PTP - UTP) - (UUP - PUP))</td>
<td>K</td>
<td>-2.74</td>
</tr>
<tr>
<td>((91.57-91.43)-(58.00-55.12))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>((90.86-91.77)-(89.30-89.20))</td>
<td>A</td>
<td>-0.19</td>
</tr>
<tr>
<td>((7.57 - 7.63)-(7.35 - 6.93))</td>
<td>B</td>
<td>+0.36</td>
</tr>
</tbody>
</table>
Attitude, and Behavior outcomes respectively. None of these differences appear to be significant, thus indicating that an interaction between Pretesting and Training is not evident based on an analysis of the mean data. This observation will be confirmed later through the lack of a significant F value in the analysis of variance.
ANALYSIS OF VARIANCE OF POSTTEST SCORES

The recommended procedure for use in analyzing the six sets of means, given initial comparability of groups, is to perform a 2 X 2 analysis of variance on the posttest scores with "presence or absence of training" and "presence or absence of the pretest" as treatment variables.

Sample sizes in the four groups of posttest scores in the present study differed slightly due to factors unrelated to the treatment variable, primarily attrition. (PTP n=65, UTP n=67, PUP n=71, UUP n=57). Since the differences in cell frequencies were not considered to be treatment related, an analysis of variance using unweighted means was performed on the data for each of the outcome measures, allowing each treatment mean to contribute equally to the sums of squares for treatments. Summary data from this two-way ANOVA are presented in Table 9.

The results of this ANOVA supported the conclusions drawn from the results of the mean comparison analyses. There was a strong positive main effect for training, Trained subjects scoring higher than Untrained subjects on all three outcome measures of Knowledge, Attitude, and Behavior. There was no evidence of a main effect for pretesting and no Pretest X Training interaction effect.
Table 9
Analysis of Variance of Posttest Scores for Knowledge, Attitude, and Behavior Outcomes

<table>
<thead>
<tr>
<th>Source</th>
<th>ss</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knowledge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>78723.625</td>
<td>1</td>
<td>78723.625</td>
<td>725.133</td>
<td>.000*</td>
</tr>
<tr>
<td>Pretesting</td>
<td>141.066</td>
<td>1</td>
<td>141.066</td>
<td>1.299</td>
<td>.255</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretesting X Training</td>
<td>121.280</td>
<td>1</td>
<td>121.280</td>
<td>1.117</td>
<td>.292</td>
</tr>
<tr>
<td>Within</td>
<td>27792.500</td>
<td>256</td>
<td>108.564</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Attitude</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>265.615</td>
<td>1</td>
<td>265.615</td>
<td>3.878</td>
<td>.05*</td>
</tr>
<tr>
<td>Pretesting</td>
<td>11.328</td>
<td>1</td>
<td>11.328</td>
<td>0.165</td>
<td>.685</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretesting X Training</td>
<td>16.145</td>
<td>1</td>
<td>16.145</td>
<td>0.236</td>
<td>.628</td>
</tr>
<tr>
<td>Within</td>
<td>17326.512</td>
<td>253</td>
<td>68.484</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Behavior</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>0.076</td>
<td>1</td>
<td>0.076</td>
<td>5.185</td>
<td>.024*</td>
</tr>
<tr>
<td>Pretesting</td>
<td>0.021</td>
<td>1</td>
<td>0.021</td>
<td>1.423</td>
<td>.235</td>
</tr>
<tr>
<td>Interaction Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretesting X Training</td>
<td>0.017</td>
<td>1</td>
<td>0.017</td>
<td>1.135</td>
<td>.289</td>
</tr>
<tr>
<td>Within</td>
<td>2.007</td>
<td>137</td>
<td>0.015</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at p < .05
Thus, null Hypothesis 1, no statistically significant difference between Trained and Untrained subjects, is rejected.
NEED FOR ACHIEVEMENT

To explore the effect of Need for Achievement, a three-way analysis of variance was performed on the Knowledge, Attitude, and Behavior Posttest scores to assess whether Need for Achievement acts to moderate or mask simple Pretest effects and/or Pretest X Training interactions. (Bunker found such a relationship between numerical aptitude and Pretest X Training). Need for Achievement was categorized into three levels and added to the two independent variables utilized in the previous 2 X 2 ANOVA (Pretesting X Training), yielding a 3 X 2 X 2 design.

Need for Achievement scores were available on 192 subjects. Though the Need for Achievement test was administered to all subjects, 71 students (27 per cent) elected not to record their student numbers, thereby protecting their anonymity. This decision rendered their test results unusable. Table 10 indicates the raw score ranges that were utilized to categorize the scores. The ranges were determined by simply dividing the entire group into three roughly equal sub-groups.

Since assignment to treatment groups was done without knowledge of the Need for Achievement variable, it was
<table>
<thead>
<tr>
<th>Category</th>
<th>Raw Score Range</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>79-99</td>
<td>60</td>
</tr>
<tr>
<td>Medium</td>
<td>50-74</td>
<td>72</td>
</tr>
<tr>
<td>Low</td>
<td>2-47</td>
<td>60</td>
</tr>
</tbody>
</table>
impossible to balance the size of the cell frequencies. This problem was further complicated by the fact that posttest measures were collected by mail and there were missing data on subjects who failed to mail their posttests.

Tables 11-A, B, and C are a summary of the results of the three-way ANOVA for all three measures. For the Behavior outcome measure, there is a significant main effect for both Need for Achievement and Training. There is also a significant Need for Achievement X Training interaction. This finding suggests that Need for Achievement acts to moderate the impact of Training on Behavior posttest scores. Table 12 contains the sample sizes and posttest means of Behavior scores for each of the 6 cells in the 3 X 2 matrix for the Training X Need for Achievement interaction. A plot of this two-way interaction is presented in Figure 4. It is apparent from this figure that there was virtually no difference between the posttest performance on the behavior measure for Medium Need for Achievement subjects in both Trained and Untrained subjects. In fact, Untrained subjects scored slightly higher. In both the High and Low Need for Achievement conditions, however, Trained subjects scored higher than Untrained subjects.

Hypothesis 9 requires an investigation into the relationship between Need for Achievement and Academic
Table 11-A

Analysis of Variance of Posttest Scores Partitioned by Need for Achievement Level For Knowledge Outcome Measures

<table>
<thead>
<tr>
<th>Source</th>
<th>ss</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N.A.*</td>
<td>346.559</td>
<td>2</td>
<td>173.279</td>
<td>1.752</td>
<td>.176</td>
</tr>
<tr>
<td>pretesting</td>
<td>454.004</td>
<td>1</td>
<td>454.004</td>
<td>4.592</td>
<td>.033</td>
</tr>
<tr>
<td>Training</td>
<td>56282.719</td>
<td>1</td>
<td>56282.719</td>
<td>569.208</td>
<td>.000**</td>
</tr>
<tr>
<td><strong>Interaction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N.A. X Training</td>
<td>392.634</td>
<td>2</td>
<td>196.317</td>
<td>1.985</td>
<td>.195</td>
</tr>
<tr>
<td>N.A. X Pretesting</td>
<td>13.67</td>
<td>2</td>
<td>6.836</td>
<td>0.069</td>
<td>.933</td>
</tr>
<tr>
<td>Pretesting X Training</td>
<td>167.229</td>
<td>1</td>
<td>167.229</td>
<td>1.691</td>
<td>.282</td>
</tr>
<tr>
<td>N.A. X Pretesting X Training</td>
<td>337.630</td>
<td>2</td>
<td>168.815</td>
<td>1.707</td>
<td>.184</td>
</tr>
<tr>
<td><strong>Within</strong></td>
<td>17699.332</td>
<td>179</td>
<td>98.879</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Need for Achievement
** Significant at p < .01
Table 11-B

Analysis of Variance of Posttest Scores
Partitioned by Need for Achievement Level
For Attitude Outcome Measures

<table>
<thead>
<tr>
<th>Source</th>
<th>ss</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N.A.*</td>
<td>379.015</td>
<td>2</td>
<td>189.508</td>
<td>2.842</td>
<td>.061</td>
</tr>
<tr>
<td>Pretesting</td>
<td>0.364</td>
<td>1</td>
<td>0.364</td>
<td>0.005</td>
<td>.941</td>
</tr>
<tr>
<td>Training</td>
<td>55.890</td>
<td>1</td>
<td>55.890</td>
<td>0.838</td>
<td>.361</td>
</tr>
<tr>
<td><strong>Interaction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N.A. X Training</td>
<td>214.181</td>
<td>2</td>
<td>107.090</td>
<td>1.606</td>
<td>.204</td>
</tr>
<tr>
<td>N.A. X Pretesting</td>
<td>162.559</td>
<td>2</td>
<td>81.280</td>
<td>1.219</td>
<td>.298</td>
</tr>
<tr>
<td>N.A. X Pretesting X Training</td>
<td>122.420</td>
<td>2</td>
<td>61.210</td>
<td>0.918</td>
<td>.401</td>
</tr>
<tr>
<td><strong>Within Cell</strong></td>
<td>11867.941</td>
<td>178</td>
<td>66.674</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Need for Achievement
Table 11-C

Analysis of Variance of Posttest Scores Partitioned by Need for Achievement Level For Behavior Outcome Measures

<table>
<thead>
<tr>
<th>Source</th>
<th>ss</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N.A.*</td>
<td>87487.000</td>
<td>2</td>
<td>43743.500</td>
<td>3.478</td>
<td>.035**</td>
</tr>
<tr>
<td>Pretesting</td>
<td>27544.141</td>
<td>1</td>
<td>27544.141</td>
<td>2.190</td>
<td>.142</td>
</tr>
<tr>
<td>Training</td>
<td>64313.859</td>
<td>1</td>
<td>64313.859</td>
<td>5.114</td>
<td>.026**</td>
</tr>
<tr>
<td>Interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N.A. X Training</td>
<td>157760.938</td>
<td>2</td>
<td>78880.438</td>
<td>6.272</td>
<td>.003**</td>
</tr>
<tr>
<td>N.A. X Pretesting</td>
<td>46903.043</td>
<td>2</td>
<td>23451.520</td>
<td>1.865</td>
<td>.161</td>
</tr>
<tr>
<td>Pretesting X Training</td>
<td>2894.990</td>
<td>1</td>
<td>2894.990</td>
<td>0.230</td>
<td>.632</td>
</tr>
<tr>
<td>N.A. X Pretesting X Training</td>
<td>5595.250</td>
<td>2</td>
<td>2797.625</td>
<td>0.222</td>
<td>.801</td>
</tr>
<tr>
<td>Within</td>
<td>1207382.000</td>
<td>96</td>
<td>12576.895</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Need for Achievement
** Significant at p < .05
### Table 12

Cell Frequencies and Posttest Means for the 2 x 2 ANOVA on Behavior Posttest Scores

<table>
<thead>
<tr>
<th>Levels of N.A.</th>
<th>Trained</th>
<th>Untrained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low N.A.</td>
<td>n=7</td>
<td>n=22</td>
</tr>
<tr>
<td></td>
<td>X=8.30</td>
<td>X=6.40</td>
</tr>
<tr>
<td>Medium N.A.</td>
<td>n=16</td>
<td>n=33</td>
</tr>
<tr>
<td></td>
<td>X=7.34</td>
<td>X=7.61</td>
</tr>
<tr>
<td>High N.A.</td>
<td>n=14</td>
<td>n=16</td>
</tr>
<tr>
<td></td>
<td>X=7.50</td>
<td>X=6.86</td>
</tr>
</tbody>
</table>
Figure 4. Behavior Posttest Score Interactions Plotted by Training Level For Each Need for Achievement Category.
Achievement (Total Points Earned). Table 13 indicates the results of an analysis of variance between these two variables which indicates a significant main effect. In further analyzing Total Points Earned by the three Need for Achievement levels, Table 14 indicates that the mean raw scores are higher for each successive level. High N.A. subjects did significantly better than Low N.A. subjects (694.65 versus 669.37). This difference is significant at the .01 level. Therefore, the null hypothesis that there is no statistically significant relationship between Need for Achievement and Academic Achievement must be rejected.
Table 13

Analysis of Variance of Academic Achievement (TOT) by Need for Achievement (EPPT)

<table>
<thead>
<tr>
<th>Source</th>
<th>ss</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for Achievement (EPPT)</td>
<td>19320.965</td>
<td>2</td>
<td>9660.480</td>
<td>3.426</td>
<td>.035*</td>
</tr>
<tr>
<td>Within</td>
<td>532895.563</td>
<td>189</td>
<td>2819.53</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at p < .05
Table 14
Means, Standard Deviations for Academic Achievement (TOT) Classified by Need for Achievement Level (EPPT)

<table>
<thead>
<tr>
<th>Level</th>
<th>X</th>
<th>S.D.</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Need for Achievement</td>
<td>694.65</td>
<td>49.31</td>
<td>60</td>
</tr>
<tr>
<td>Medium Need for Achievement</td>
<td>680.22</td>
<td>52.95</td>
<td>72</td>
</tr>
<tr>
<td>Low Need for Achievement</td>
<td>669.37</td>
<td>55.59</td>
<td>60</td>
</tr>
</tbody>
</table>
INDIVIDUAL DIFFERENCE VARIABLES

The significance of the individual difference variables of Education (number of years of formal schooling) and Type (category of subject employment - whether in a company-owned or licensee restaurant) on outcome measures, Hypothesis 10, was investigated through an analysis of variance of posttest scores for each of the three outcome measures. The results, found in Table 15, indicate a significant relationship only between Education and posttest measures of Attitude. However, an analysis of variance of Total Points Earned (Academic Achievement) by Education and Type, indicates a highly significant (<.01) relationship between Education and Total Points Earned, as seen in Table 16. Thus, null Hypothesis 10 is rejected.

Post-training versus 45-day post-training measures.

Hypothesis 11 deals with the retention of Knowledge, Attitude, and Behavior gains over time. To test these relationships, comparisons were made between Pre, Post, and 45-Day Post measures of Knowledge and Attitude. The difficulties posed by collection of data by mail, the length of time involved in the completion of the Behavior instruments required of those who were asked to complete
Table 15

Analysis of Variance of Posttest Scores
By Education Level and Type of Employment
For Knowledge, Attitude, and Behavior Outcomes

<table>
<thead>
<tr>
<th>Source</th>
<th>ss</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knowledge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main Effects Education</td>
<td>5170.488</td>
<td>8</td>
<td>646.311</td>
<td>1.548</td>
<td>.142</td>
</tr>
<tr>
<td>Type</td>
<td>1296.331</td>
<td>2</td>
<td>648.166</td>
<td>1.553</td>
<td>.214</td>
</tr>
<tr>
<td>Interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education X Type</td>
<td>964.815</td>
<td>7</td>
<td>137.831</td>
<td>0.330</td>
<td>.940</td>
</tr>
<tr>
<td>Within</td>
<td>96836.000</td>
<td>232</td>
<td>417.396</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Attitude</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main Effects Education</td>
<td>1184.893</td>
<td>8</td>
<td>148.112</td>
<td>2.197</td>
<td>.029*</td>
</tr>
<tr>
<td>Type</td>
<td>82.759</td>
<td>2</td>
<td>41.380</td>
<td>0.614</td>
<td>.542</td>
</tr>
<tr>
<td>Interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education X Type</td>
<td>212.930</td>
<td>7</td>
<td>30.419</td>
<td>0.451</td>
<td>.869</td>
</tr>
<tr>
<td>Within</td>
<td>15505.461</td>
<td>230</td>
<td>67.415</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Behavior</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main Effects Education</td>
<td>40673.516</td>
<td>7</td>
<td>5810.500</td>
<td>0.379</td>
<td>.913</td>
</tr>
<tr>
<td>Type</td>
<td>1402.509</td>
<td>1</td>
<td>1402.509</td>
<td>0.091</td>
<td>.763</td>
</tr>
<tr>
<td>Interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education X Type</td>
<td>116588.313</td>
<td>6</td>
<td>19431.383</td>
<td>1.267</td>
<td>.277</td>
</tr>
<tr>
<td>Within</td>
<td>1885834.000</td>
<td>123</td>
<td>15331.980</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at p < .05
Table 16

Analysis of Variance of Total Points Earned By Education Level and Type of Employment

<table>
<thead>
<tr>
<th>Source</th>
<th>ss</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>86244.938</td>
<td>10</td>
<td>8624.492</td>
<td>3.042</td>
<td>.001*</td>
</tr>
<tr>
<td>Type</td>
<td>2659.582</td>
<td>1</td>
<td>2659.582</td>
<td>0.938</td>
<td>.334</td>
</tr>
<tr>
<td>Interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education X Type</td>
<td>19087.406</td>
<td>8</td>
<td>2385.926</td>
<td>0.842</td>
<td>.567</td>
</tr>
<tr>
<td>Within</td>
<td>787478.313</td>
<td>256</td>
<td>3076.087</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at p < .01
them, and the perceived difficulties in differentiating behavior changes in such a short time (six weeks) by individuals who are not normally in daily contact with the subjects, all combined to influence the researcher to omit the data collection of 45-day post behavior measures. Table 17 contains the means, standard deviations, and sample sizes of pre and posttest scores of Knowledge and Attitude measures for Trained subjects. Using this data, paired mean comparisons were made. As seen in Table 18, all post versus pre measures are significant at the .01 level. The 45-day post versus post and pre comparisons show a significant difference still exists for knowledge but not for attitude. Several factors may account for these results, each of which will be discussed in Chapter V.
Table 17


<table>
<thead>
<tr>
<th>Outcome</th>
<th>Pretest</th>
<th>Posttest</th>
<th>45-Day Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>X=54.64</td>
<td>X=91.50</td>
<td>X=86.42</td>
</tr>
<tr>
<td></td>
<td>SD=8.70</td>
<td>SD=9.38</td>
<td>SD=9.03</td>
</tr>
<tr>
<td></td>
<td>N=63</td>
<td>N=133</td>
<td>N=53</td>
</tr>
<tr>
<td>Attitude</td>
<td>X=85.27</td>
<td>X=91.32</td>
<td>X=90.42</td>
</tr>
<tr>
<td></td>
<td>SD=10.59</td>
<td>SD=7.75</td>
<td>SD=9.35</td>
</tr>
<tr>
<td></td>
<td>N=63</td>
<td>N=131</td>
<td>N=53</td>
</tr>
</tbody>
</table>
Table 18
Pretest, Posttest, 45-Day Posttest Paired Mean Comparisons Testing For Retention of Gains in Knowledge and Attitude Outcomes

<table>
<thead>
<tr>
<th>Comparison</th>
<th>n</th>
<th>Means</th>
<th>SD</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>K Post</td>
<td>63</td>
<td>92.19</td>
<td>8.62</td>
<td>38.06</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>vs K Pre</td>
<td></td>
<td>54.63</td>
<td>8.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K Post 45</td>
<td>20</td>
<td>85.27</td>
<td>9.27</td>
<td>-2.13</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>vs K Post</td>
<td></td>
<td>90.64</td>
<td>12.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K Post 45</td>
<td>20</td>
<td>85.30</td>
<td>9.30</td>
<td>14.56</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>vs K Pre</td>
<td></td>
<td>54.60</td>
<td>9.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Post</td>
<td>63</td>
<td>90.76</td>
<td>8.36</td>
<td>3.08</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>vs A Pre</td>
<td></td>
<td>85.27</td>
<td>10.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Post 45</td>
<td>24</td>
<td>90.75</td>
<td>10.33</td>
<td>0.49</td>
<td>N.S.</td>
</tr>
<tr>
<td>vs A Post</td>
<td></td>
<td>90.33</td>
<td>8.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Post 45</td>
<td>22</td>
<td>90.55</td>
<td>10.68</td>
<td>0.76</td>
<td>N.S.</td>
</tr>
<tr>
<td>vs A Pre</td>
<td></td>
<td>88.27</td>
<td>10.46</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER V

DISCUSSION

The major purpose of this study was to address the problem of how to best design and implement an evaluation of a specific training program. The selected training program is the final step in a highly systematic series of training experiences designed to prepare trainees to successfully manage a McDonald's restaurant. The main hypothesis for the study was aimed at assessing the effectiveness of the training course by comparing pre and post measures of three separate outcomes: Knowledge, Attitude, and Behavior. In addition, hypotheses testing the relationship between several individual difference variables and their impact on training outcomes and the effectiveness of training over time were investigated. It is clear from the data presented that the completeness of information derived from an evaluation of a training program can vary depending upon the level of evaluation chosen to analyze training outcomes.
VERIFICATION OF THE IMPACT OF TRAINING

The most typical form of training evaluation, when it is conducted at all, is the measurement of pretest to posttest performance change on the relevant dimension(s) for a trained group only. In the present study, an analysis of the mean data for the PTP subjects is analogous to this single-group evaluation design. The significant pre-post change reported for these Pretested-Trained-Posttested subjects (see Table 5) supports the efficacy of the training experience at this level of evaluation.

At one level higher in the evaluation hierarchy, a pretested-untrained control group (PUP) is added to the design. The results of this extra control group add greater credibility to the contention that a main effect for training was responsible for the pre-post improvement of the trained subjects (PTP). The significant difference between the posttest scores of the Pretested-Trained and Pretested-Untrained subjects (see Table 5), and the absence of a significant difference between Untrained subject's pretest and posttest means, suggests that PTP subjects benefitted from the training received and not simply from the "Hawthorne Effect."

131
Comparisons involving the remaining means generated by the 4-Group design provide a clear delineation of the training main effect. With the exception of the attitude and behavior comparisons on the Pretested-Trained versus Pretested-Untrained groups, all other comparisons indicate a main effect for training. Overall, on the basis of the mean data, one would have to conclude that the Advanced Operations Course had a significant positive impact on the performance of trainees on Knowledge, Attitude, and Behavior measures.

Similarly, the two-way ANOVA on the posttest scores (see Table 9) for all three outcome measures verifies the presence of a substantial main effect for training, and identifies no contaminations involving pretesting that would require qualification of the training impact.

The results of the three-way ANOVA on the posttest scores for each outcome measure partitioned by Need for Achievement, however, produced some mixed results. Here a main effect was observed for training on the Knowledge and Behavior outcome measures, but not the Attitude measure. In addition, it was found that the Need for Achievement variable, when partitioned into three levels, interacts with training, with the posttest performance of Medium Need for Achievement-Untrained subjects slightly higher than Trained subjects at the same Need for Achievement level.
The successive levels of analysis afforded by the Solomon 4-Group Design consistently demonstrates a strong main effect for training for all three outcome measures. It was anticipated that some pretest sensitization effects were present in the Advanced Operations Course environment, particularly in light of the fact that all three pretest/posttest measures of Knowledge, Attitude, and Behavior were each identical forms of the same test. Additionally, there are several factors in the A.O.C. environment which were felt would be significant contributors to a pretest effect: 1) the emphasis on testing throughout the Advanced Operations Course creates a high awareness of the importance of doing well, 2) the practice of posting test scores, by student number, further heightens a sense of competition and pressure to perform among all students, and 3) the specific nature of the pretest could have targeted student attention to specific areas of the course, thus causing a misguided attention to a limited sample of training content presented throughout the course. As has already been seen, however, the analysis indicates that there was no significant pretest effect, even when including the Need for Achievement variable (Tables 9 and 11-A, B, and C).
INDIVIDUAL DIFFERENCE VARIABLES

Several individual difference variables were investigated in this study to assess their relationship to the Knowledge, Attitude, and Behavior outcomes. It was originally felt by the Hamburger University faculty that the single most important variable accounting for excellent performance in the Advanced Operations Course was a strong motivation to do well, defined in this study as Need for Achievement. In addition to the factors cited above, the nature of the Hamburger University experience in McDonald's corporate culture brings added pressure on most students to perform well. Most students have already heard about the special awards given for outstanding academic achievement (the A.O.C. Archie) and the recognition received for finishing in the top 10 per cent of the class (the Dean's List). Some students are even told by their supervisors (or owner operators) that they will receive additional compensation if they win one of the awards. Finally, since Hamburger University has a twenty-year history, and all individuals in the company ultimately attend the Advanced Operations Course, there is often pressure placed on students to perform as well or better than those who have "come before."
In investigating the Need for Achievement variable, it was found that there was a strong relationship between this variable and Total Points Earned (TOT), also defined as Academic Achievement. An analysis of variance of these two variables yielded a significant F value (Table 13). To further investigate this relationship, mean scores for the Total Points Earned variable were analyzed for each of the three Need for Achievement levels: High, Medium, and Low. The raw score analysis indicates that each successively higher level of Need for Achievement had a corresponding increase in the mean TOT score (Table 14). The difference between High and Low was highly significant (.01). This type of relationship has been well documented in the general literature for students in more traditional academic settings, thus lending credence to these findings.

As has previously been indicated, the results of a three-way ANOVA on each outcome measure partitioned by three distinct levels of Need for Achievement (High, Medium, and Low) resulted in an interaction at the Medium level. It was found that the posttest performance of Trained subjects was slightly higher than Untrained subjects at the Medium Need for Achievement level. (See Table 11-C and Figure 4). This is an unexpected outcome in that it was expected that regardless of Need for Achievement level, all of the other evaluation results would suggest that Trained subjects should score better on
the posttest than Untrained subjects. An analysis of the cell frequencies for this analysis (Table 12) indicates that there were twice as many Untrained subjects (n=33) as there were Trained subjects (n=16) at the Medium Need for Achievement level. The difference in mean scores of these two groups is small, .37 (7.61 versus 7.34). It is entirely possible that had the cell frequencies for this analysis been more equivalent, this difference would not have occurred. Since only Untrained Medium Need for achievement subjects had higher Behavior Posttest scores, there appears to be no rational explanation that would explain why this is true at the Medium level and not the High level. Thus, it is the researcher's conclusion that these results are produced by the substantial differences in cell frequencies and not by other factors.

The two other individual difference variables of interest in this study were Education (number of years of formal schooling) and Type (category of subject employment, whether in a company-owned or licensee restaurant). A previous study had indicated that these two variables might be related to outcomes of the course. In an analysis of variance of Education, Attitude, and Type with each of the three outcome measures (Table 15), only Education and Attitude were found to have a significant relationship. An analysis of variance of Total Points Earned (Academic
Achievement) with the same individual difference variables indicates a significant relationship between Education and Total Points Earned. It was originally hypothesized that there would be such a relationship (Hypothesis 10). Several factors related to a subject's level of education should enhance his/her performance in the Advanced Operations Course. These would include better note-taking skills, more developed study habits, a greater familiarity with one's own learning style, and more advanced test-taking skills. That Education was strongly related to Total Points Earned and not the Knowledge Posttest may be explained by the significant difference in these two variables. The Knowledge Posttest represents a single, fifty-item multiple-choice test which tests some of the major concepts of the course. The Total Points Earned variable includes six separate multiple-choice tests, each one focused on testing specific curriculum content areas of the course. It also includes a 100-point faculty evaluation. Thus, the Total Points Earned variable allows greater latitude for the advanced learning and testing skills of more-educated A.O.C. Students to positively impact results.

The relationship between Education and Attitude Posttest measures is an interesting and unexpected one. Students with higher levels of education scored higher on the Attitude Posttest than students with lower levels of education. Though it is speculative, the researcher's
own experience as a faculty member at Hamburger University suggests the following explanation. AOC students with higher levels of education come to Hamburger University with some pre-conceived notions about what to expect based on their previous, more traditional educational experiences. Two common reactions to the A.O.C. experience which are shared by graduates regardless of their level of education are: 1) they are highly impressed by the quality of the curriculum, the sophistication of the Hamburger University learning facility, its array of "high tech" classroom audio-visual equipment and working equipment labs, the professionalism and expertise of the faculty, and 2) they are highly appreciative of the knowledge and skills that they acquire during the course which they will be able to apply back on the job to become more effective restaurant managers. The combination of these two factors creates an expressed positive attitude toward McDonald's Corporation and Hamburger University. Students with higher levels of education, and thus more exposure to traditional educational settings, are better able to appreciate the quality and value of the course. This could account for the significant relationship between Education and Attitude. These findings are consistent with a study done by Schien which found that education level was correlated
with a positive attitude change as a result of training.
Another issue of great interest in this study was to investigate whether the main effect that was hypothesized for training on the Knowledge, Attitude, and Behavior measures would persist over time. It was the original intent of the researcher to collect data on all three of these outcome measures immediately at the end of training and again forty-five days later. Several factors combined to make the data collection of the forty-five day Behavior Posttest nearly impossible:

1. The length of time required to complete these instruments coupled with the fact that they had to be completed by the subject's immediate supervisor, peer, and a subordinate made the entire data collection process more complicated and difficult.

2. In discussions with some of the Hamburger University faculty and with some of those individuals asked to complete the Behavior Index, it became obvious that forty-five days was too short a time period to differentiate behavior. Many of those asked to complete the Behavior Index are not normally in daily contact with the subjects,
thereby making the differentiation between observed behavior immediately after the course and forty-five days later extremely difficult.

Thus, it was decided to eliminate data collection of the forty-five day Behavior Index. As has been previously discussed, all Post measures of Knowledge, Attitude, and Behavior are significantly higher than their corresponding Pretest measures. Paired mean comparisons were made on both the Knowledge and Attitude measures. Comparisons were made between forty-five day Posttest scores versus Posttest scores, and forty-five day Posttest scores and Pretest scores for each. (See Table 18). For the Knowledge outcome measure, both comparisons were significantly different. However, the raw score comparisons indicate that there is some slippage in retention; K Pre = 54, K Post = 92, and K Post 45 = 85. The slight differences in raw scores in each comparison is a function of the unequal "n's." Only 20 of the 63 Pretested-Trained-Posttested group returned their forty-five day Post knowledge test. Thus, if can be concluded that the significant gain in knowledge achieved by A.O.C. graduates does persist forty-five days later, but at a lower level. These findings are consistent with the well-documented findings in more traditional academic settings that most learner's retention of learned material decreases over time (Green, Bloom, and Kimball; McGovern; and Underwood).
The raw scores for the Attitude measures for the PTP group are as follows: A Pre = 85, A Post = 90, A Post 45 = 90. There is a significant difference between the pre and posttest scores as seen in Table 18. However, the scores on the forty-five day Attitude posttest and the posttest administered at the end of the course are nearly identical, 90.75 versus 90.33. Only 24 of the 63 PTP group returned forty-five day Attitude Posttests. The results of the forty-five day Posttest versus the Pretest comparison are somewhat misleading. The 22 students on which the paired comparisons were made had a higher attitude pretest mean score, 88.27, than did the larger (63) Pretested-Trained-Posttested group, whose mean pretest score was 85.27. (See Table 18). The raw score analysis, however, demonstrates that trained students have a more favorable attitude after training, based on a favorable response rate on the Attitude measure, and it remains at this same high level forty-five days later.
IMPLICATIONS OF THE STUDY

The main finding of this study is that the Advanced Operations Course produces a significant main effect for graduates, as measured by the change in pre versus post measures of Knowledge, Attitude, and Behavior. Both the Knowledge and Attitude gains still persist forty-five days after the training course has ended. Obviously, McDonald's Corporation has believed that the course provides a benefit to the organization, as evidenced by the significant capital investment the company has made in the Hamburger University facility and the Advanced Operations Course curriculum. These findings serve to further document the measurable results of the course and thus demonstrate that a tangible return on the capital investment is being realized.

Though not included as a part of the statistical analysis for this study, both the Opinion Survey and the Behavior Index generated a significant amount of data in the form of comments by A.O.C. students themselves, and by superiors, peers, and subordinates. Two observations can be made about the general nature of these comments:

1. A.O.C. students reported that McDonald's training in general, and the A.O.C./Hamburger University
experience in particular, had a significant positive impact on their attitudes toward themselves, their jobs, and the company.

2. Superiors, peers, and subordinates' comments generally referred to a noticeable increase in the A.O.C. graduate's level of enthusiasm about the job and a strong desire to implement new ideas and techniques to improve store operations.

These findings are highly consistent with the "Reaction"-type evaluations that Hamburger University faculty have had graduates complete at the end of the course. Students consistently identify the course and the experience as the "highlight of their McDonald's career."

Peters and Waterman, in a study of America's best run companies, selected McDonald's Corporation as one of the best. Among the traits and characteristics they found common to most of the "excellent" companies was a strong corporate culture and shared values, a key to unifying the social dimensions of the organization. They cited a list of dominant beliefs among the excellent companies, which include:

1. A belief in being the best
2. A belief in the importance of the details of execution, the nuts and bolts of doing the job well
3. A belief in the importance of people as individuals
4. A belief in superior quality and service
5. Explicit belief in and recognition of the importance of economic growth and profits. 60

It seems clear in analyzing the data collected in this study that the content of the Advanced Operations Course and the experience of attending Hamburger University is a primary means by which McDonald's Corporation establishes a strong corporate culture and commitment to its values among its Restaurant Managers. The statistical evidence of this study indicates that the Advanced Operations Course has a significant effect on student's knowledge, attitude, and behavior.

The original intent of this study was to assess four levels of training outcomes: knowledge, attitude, behavior, and results, as prescribed in Kilpatrick's evaluation model. The difficulties associated with trying to isolate and measure specific in-restaurant results that could be directly attributable to the Advanced Operations Course prevented the researcher from including results measures in the study. The data collection of in-restaurant measures is further complicated by three other factors:

1. Students attending A.O.C. come from restaurants throughout the U.S. and the world, thus making in-restaurant follow-up and data collection more difficult.

2. Data collection in company-owned restaurants is much easier than in licensee restaurants because of the manner in which company restaurants are structured and controlled by the Corporation. However, A.O.C. students from company-owned restaurants represent only 27 per cent of the total student population.

3. It is possible that there may be more than one A.O.C. graduate working in the same McDonald's
restaurant. This is more likely in licensee restaurants, thereby confounding the impact of A.O.C. training.

These factors can be overcome, however, and it is recommended that further research be done to investigate what specific in-restaurant benefits are derived. Potentially, the multi-baseline design would yield meaningful data if targeted at specific behaviors. Additionally, a longitudinal study which attempts to correlate A.O.C. training with sales, profit trends, and levels of customer service is strongly recommended.

A major area for further research as indicated by the literature review is the efficacy of various training methods. The Hamburger University Faculty presently utilizes a variety of training methods throughout the course; however, the dominant method is lecture. Further research needs to be done to investigate each of the alternate methods and their impact on cognitive and affective outcomes. Clearly, some topics may lend themselves more appropriately to alternate training methods. The sophistication and capacities of the present H.U. facility also argue for the testing of computer-based training (CBT) techniques as an alternate method of instruction.
SUMMARY

The purpose of the study was to design and implement an evaluation of an organizational training program and to assess multiple outcomes.

The program, the Advanced Operations Course, is offered by McDonald's Corporation at its corporate training center, Hamburger University. Utilizing the Solomon 4-Group design, 263 subjects were divided into four groups based on the Trained/Untrained and Pretested/Unpretested conditions. Instruments were designed to test knowledge, attitude, and behavior outcomes. The study also investigated the impact of the individual difference variables - Need for Achievement and Education level - and retention of the three outcomes over time.

The results of the study demonstrated a positive main effect for training on knowledge, attitude, and behavior outcomes. Forty-five days later, the knowledge effect was still present, though at a slightly lower level, while the attitude effect remained at the same post-training level. Need for Achievement and Education were found to be strongly correlated with overall academic achievement in the course, as were Education and Attitude.
Further research is indicated to investigate more fully the benefits of the course, in terms of improved restaurant operating results and in the efficacy of the various training methods utilized throughout the training program.
REFERENCES


APPENDIX A
A.O.C. KNOWLEDGE TEST

This 50-item multiple choice test is designed to measure your general knowledge level of some of the subject matter included in the Advanced Operations Course taught at Hamburger University.

Please observe the following instructions:

1) Working alone, and with no other resources (Operations & Training Manual, Hamburger University Notebook, etc.) complete the test. Allow yourself 25-30 minutes.

2) Use the enclosed IBM Answer Sheet and a No. 2 lead pencil (no ink pens). Record one response for each of the 50 questions. Erasures should be as complete as possible. Be sure to fill in your name, social security number and A.O.C. class number on the answer sheet.

3) Use the enclosed postage-paid envelope to return both the test and the completed answer sheet to Hamburger University as soon as possible (within the next 3-5 days).

The validity of this test is based upon how well it measures your individual knowledge. Please observe the above instructions and then return both the test and the completed answer sheet to Hamburger University.
1. "The combined efforts of everyone is greater than the sum of the individual efforts" is called:
   A. Alliance  
   B. Energy  
   C. Group Dynamics  
   D. Synergy

2. If a team was unable to reach agreement, chances are the objective was not clear; or the team failed to:
   A. Establish a list of rights  
   B. Establish methods  
   C. Identify a leader  
   D. Place a time limit on the meeting

3. The time posts on the defrost clock indicate:
   A. Beginning of the defrost cycle  
   B. Correct time of day  
   C. Failsafe settings  
   D. All of the above

4. People who are successful in time management all have one thing in common:
   A. They make a daily ‘to do’ list  
   B. They set time and motion goals  
   C. They spend time away from the office  
   D. They work less hours

5. The evaporator is the component that allows the refrigerant to change:
   A. From a gas  
   B. To a gas  
   C. To a high pressure gas  
   D. To a liquid

6. Which term best describes time spent on those activities you want to do?
   A. Boss imposed  
   B. Discretionary  
   C. System imposed  
   D. Subordinate imposed

7. What is the McDonald’s recommended temperature for our restaurants in the summer months?
   A. 50° F (10° C)  
   B. 60° F (15.5° C)  
   C. 68° F (20° C)  
   D. 78° F (25.5° C)

8. One of the four steps of training is "try out". This must include a demonstration of the task by the trainee. Why?
   A. To give basic instructions  
   B. To provide practice  
   C. To put trainee at ease  
   D. To verify skills transfer

9. A team’s "objective" is:
   A. Formal agenda or task  
   B. How the team interacts  
   C. Its process  
   D. The act of processing

10. Ownership of feelings and expressing perceptions are key components in what part of team activity?
    A. Content  
    B. Method  
    C. Processing  
    D. Task Accomplishment

11. White males that normally just help other white males to succeed in McDonald’s are usually:
    A. Different in appearance from other employees  
    B. Ingroup members  
    C. Outgroup members  
    D. Sensitive to outgroup members

12. The way you honestly feel about an outgroup member contributes to:
    A. Their attitude toward you  
    B. Your behavior  
    C. Your sensitivity  
    D. All of the above

13. Your second assistant, although thoroughly trained, submits a crew schedule with errors. What form of training should be used?
    A. Advanced  
    B. Corrective  
    C. Enrichment  
    D. Example

14. A question not to ask during the interview is:
    A. Do you own a car?  
    B. How long will you work?  
    C. Why did you leave your last job?  
    D. Why work at McDonald’s?
<table>
<thead>
<tr>
<th>Question</th>
<th>Choices</th>
<th></th>
<th>Question</th>
<th>Choices</th>
</tr>
</thead>
</table>
| 15. The major key to retention of our crew people is: | A. A designated crew recruiter  
B. A good crew incentive program  
C. An organized crew training program  
D. Good communications | 22. The goal of the nurturing parent is: | A. To be accepted  
B. To be competent  
C. To be loved  
D. To be superior |
| 16. In situations that require equalizing baffles, their purpose is to: | A. Extract grease from the air  
B. Prevent negative air flow  
C. Properly move air across the grill surface  
D. Provide positive air flow | 23. For a stroke to be recorded it must be: | A. Accepted  
B. Parallel  
C. Positive  
D. All of the above |
| 17. An objective has four elements of measurability. Which one identifies how well it will be accomplished? | A. Conditions  
B. Quality  
C. Quantity  
D. Time | 24. At times, the adapted child can be: | A. Competent  
B. Nurturing  
C. Rebellious  
D. Withdrawn |
| 18. The first step in the setting objectives process is setting the objective. What is the third step? | A. Action Plan  
B. Goal Setting  
C. Performance Review  
D. Work Review | 25. You're experiencing an overrun percentage of 40% in your sundaes, which of the following could be the cause? | A. Air orifice too large  
B. Draw temperature too low  
C. The condition of the check bands  
D. The condition of the topping pumps |
| 19. Which ego state (or part of one) is creative: | A. Adapted parent  
B. Adult  
C. Natural child  
D. Nurturing child | 26. Self-contained thermostats are located in the: | A. Blower compartment  
B. Return air duct  
C. Supply air duct  
D. Lobby |
| 20. Active listening is an indicator of which ego state? | A. Adapted child  
B. Adult  
C. Critical parent  
D. Nurturing parent | 27. How often should the exhaust stacks in the grill area be cleaned? | A. Every month  
B. Every 2 months  
C. Every 3 months  
D. Every 6 months |
| 21. When we say one thing, but mean something else, what type of transaction occurs? | A. Crossed  
B. Diagram  
C. Parallel  
D. Ulterior | 28. An exhaust fan v-belt that can be turned 270 degrees indicates: | A. A loose belt  
B. An exhaust fan that runs slow  
C. An inefficient exhaust system  
D. All of the above |
<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
</table>
| 29. To achieve a 68° F (20° C) temperature in the store, the 68° F (20° C) temperature must be measured: | A. At head level standing in the kitchen  
B. In the stockroom next to the perishables  
C. Next to the rooftop condenser  
D. Outside of the drive-thru window |
| 30. The T-9 filter is designed to: | A. Filter undissolved impurities  
B. Remove all minerals  
C. Soften hard water  
D. All of the above |
| 31. Monthly turnover for August is 20%. I will need 75 people to handle the volume. How many people will I need to replace because of turnover? | A. 4  
B. 12  
C. 15  
D. 60 |
| 32. Which of the following is not one of the four steps in the "how to" of delegation? | A. Communicate how and how well  
B. Communicate what  
C. Follow-up  
D. Set priorities |
| 33. Which of the following helps prevent lime build-up in the ice machine? | A. Booster Pump  
B. Micromet Feeder  
C. Solvent Remover  
D. T-9 Filter |
| 34. A store invested $800 in a promotion. The sales generated were $2700. According to L.S.M. guidelines regarding breakeven, this store has: | A. Achieved good public relations  
B. Exceeded the guidelines  
C. Met the guidelines  
D. Not met the guidelines |
| 35. How many parts are there to the adult ego state? | A. Four  
B. Three  
C. Two  
D. One |
| 36. On the large walk-in freezer the failsafe should be set for ________ from the beginning of the defrost cycle. | A. 30 minutes  
B. 50 minutes  
C. 55 minutes  
D. 60 minutes |
| 37. "We will conduct our team proceedings my way," is an example of a statement that would be made by a: | A. Topic Jumper  
B. Non-Team Player  
C. Rotating Leader  
D. Team Member |
| 38. Which three team roles are always negative? | A. Blocker Clown Dominator  
B. Blocker Clown Informer  
C. Blocker Dominator Topic Jumper  
D. Blocker Judge Topic Jumper |
| 39. If we receive a cold prickly, we should: | A. Accept the valid information  
B. Accept the feeling  
C. Reply with a cold prickly  
D. Zilch the person |
| 40. A Ronald McDonald appearance is a promotion that is designed to: | A. Reinforce an advertising promise  
B. Solve an operational problem  
C. Take advantage of kid's spendable income  
D. Promote a new product |
| 41. Overweight shakes can be caused by improper: | A. Checkbands  
B. Priming  
C. Syrup Calibration  
D. All of the above |
| 42. How should we lubricate the pressure switch diaphragm? | A. Lightly  
B. Moderately  
C. Heavily  
D. Not at all |
<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>43. A person who functions predominantly from the parent ego state has a tendency to be a (an):</td>
<td></td>
</tr>
<tr>
<td>A. Almost Winner</td>
<td>50. An Outgroup Member:</td>
</tr>
<tr>
<td>B. Almost Loser</td>
<td>A. Is excluded from Ingroup</td>
</tr>
<tr>
<td>C. Loser</td>
<td>B. Is visibly different</td>
</tr>
<tr>
<td>D. Winner</td>
<td>C. Wants to belong</td>
</tr>
<tr>
<td></td>
<td>D. All of the above</td>
</tr>
<tr>
<td>44. To best manage the almost loser, you should use:</td>
<td></td>
</tr>
<tr>
<td>A. Adapted Child</td>
<td></td>
</tr>
<tr>
<td>B. Critical Parent</td>
<td></td>
</tr>
<tr>
<td>C. Natural Child</td>
<td></td>
</tr>
<tr>
<td>D. Nurturing Parent</td>
<td></td>
</tr>
<tr>
<td>45. Which of the following is not an attribute of the “I’m OK · You’re OK” person?</td>
<td></td>
</tr>
<tr>
<td>A. Accepts delegation</td>
<td></td>
</tr>
<tr>
<td>B. Feels equal</td>
<td></td>
</tr>
<tr>
<td>C. Makes snap decision</td>
<td></td>
</tr>
<tr>
<td>D. Never disagrees</td>
<td></td>
</tr>
<tr>
<td>46. Which of the following will not result from a clogged condenser?</td>
<td></td>
</tr>
<tr>
<td>A. Electricity usage</td>
<td></td>
</tr>
<tr>
<td>B. Equipment Life Shortage</td>
<td></td>
</tr>
<tr>
<td>C. Evaporator freeze-up</td>
<td></td>
</tr>
<tr>
<td>D. Loss of efficiency</td>
<td></td>
</tr>
<tr>
<td>47. A basic concept that applies to all refrigeration is:</td>
<td></td>
</tr>
<tr>
<td>A. Cold items contract</td>
<td></td>
</tr>
<tr>
<td>B. Heated items expand</td>
<td></td>
</tr>
<tr>
<td>C. Heat travels to cold</td>
<td></td>
</tr>
<tr>
<td>D. Pressure and heat have an inverse relationship</td>
<td></td>
</tr>
<tr>
<td>48. In the sundae machine, what controls the mix pump?</td>
<td></td>
</tr>
<tr>
<td>A. Beater Assembly</td>
<td></td>
</tr>
<tr>
<td>B. Draw Gate Valve</td>
<td></td>
</tr>
<tr>
<td>C. Poppet Valve Assembly</td>
<td></td>
</tr>
<tr>
<td>D. Pressure Switch</td>
<td></td>
</tr>
<tr>
<td>49. The primary benefit in using a crew recruiter is that it:</td>
<td></td>
</tr>
<tr>
<td>A. Improves morale</td>
<td></td>
</tr>
<tr>
<td>B. Increases QSC</td>
<td></td>
</tr>
<tr>
<td>C. Increases turnover</td>
<td></td>
</tr>
<tr>
<td>D. Saves money</td>
<td></td>
</tr>
</tbody>
</table>
INSTRUCTIONS: Opinion surveys are used to help us understand how you feel about McDonald's and your job. Unlike the Attitude Surveys, which was generally used by McDonald's Corporation and its licensees as a basis for identifying areas which need improvements, this Opinion Survey is intended only to measure your opinions. No subsequent action on the results should be expected. To help us please complete this questionnaire.

HOW TO RESPOND: You will receive a separate answer sheet on which to record your responses. After reading each statement, select the response that best describes how you feel about the statement. Selecting the "?" indicates that:

-Sometimes you agree, sometimes you don't, or
-That you do not understand the question, or
-You are too new in McDonald's to have an opinion.

If your answer is Yes mark "A" on your answer sheet.
If your answer is ? mark "B" on your answer sheet.
If your answer is No mark "C" on your answer sheet.

1. The job I have allows me to improve my skills ..............................................................

2. Our pay system is competitive .................................................................

3. I can have a satisfactory career in McDonald's ........................................

4. I would recommend working for McDonald's to my friends ....................

5. I am interested in my next possible job promotion ............................

6. McDonald's promotes from within ............................................................... 

7. My training is helping to develop me for future positions ........................

8. McDonald's is the best restaurant chain in the world ..............................

9. The people I work with are professional and competent and I enjoy working with them

10. I understand McDonald's policies and procedures ........................................

11. Evaluation of my learning would be more effective if it were done less frequently

12. I feel proud to work for McDonald's.................................................................

13. I feel secure in my job at McDonald's ............................................................

14. I look forward to coming to work .................................................................

15. McDonald's has adequately trained me to do my job ............................

16. I prefer working for McDonald's over any other company I know ........

17. McDonald's treats its employees as well or better than any company I know.

18. I see myself working for McDonald's 3 years from now ...........................

19. I enjoy talking to family and friends about McDonald's and about my job

20. I get angry when I see anti-McDonald's television commercials, newspaper stories, etc.

21. It is important to me that I work for a "Number 1" company in it's industry

22. McDonald's training activities have a positive affect on my attitude

23. My attitude about McDonald's affects how well I do my job

24. McDonald's cares a great deal about it's employees

25. Acceptable job performance insures my security

26. McDonald's is an excellent company to work for and I would recommend it to my friends

27. My work gives me a sense of achievement

28. I feel McDonald's menu items are superior in both quality and value to the competition's

29. I enjoy being identified as an employee of McDonald's to family, friends or acquaintances

NOTE: The following demographic information will be used to interpret and analyze results. Completion is optional. Select the appropriate letter on your answer sheet for each item.

30. My sex is:  A. Male  B. Female

31. My job is:  A. Store Manager  B. First Assistant  C. Registered Applicant  D. Staff  E. Other

32. My attitude about McDonald's is most affected by

...........................................................................................................................................
APPENDIX C
MANAGEMENT BEHAVIOR INDEX

Manager's Name

This index asks you to assess the behavior of the manager whose name appears above. It is to be completed anonymously, so please don't sign your name.

Your tendency may be to respond to these statements by subjectively evaluating the person. Instead, this instrument is intended to measure BEHAVIOR and its frequency. As you read each statement, ask yourself:

A. Does this manager BEHAVE this way?
B. How frequently have I observed this BEHAVIOR?

Your responses should be based on your direct work experience with this manager.

It may help to think of the five choices as frequency levels on a numerical scale from 10 (always) to 1 (never):

<table>
<thead>
<tr>
<th>10</th>
<th>9</th>
<th>8</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always</td>
<td>Often</td>
<td>Occasionally</td>
<td>Seldom</td>
<td>Never</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you have never had the opportunity to observe the BEHAVIOR identified in a particular statement, use your judgment based on your knowledge of the manager and his/her normal behavior. Then circle the appropriate response.

Before you turn the page, please check the title in the upper right-hand corner of the page that describes your relationship to this manager.
### BEHAVIOR

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Frequency</th>
</tr>
</thead>
</table>
| 1. Treats all subordinates equally in applying store personnel practices. | A. Always 10  
B. Often 7  
C. Occasionally 5  
D. Seldom 3  
E. Never 1 |
| 2. Effectively sets personal and store objectives.                       | A. Always 10  
B. Often 7  
C. Occasionally 5  
D. Seldom 3  
E. Never 1 |
| 3. Effectively assists subordinates in preparing their personal and store-related objectives. | A. Always 10  
B. Often 7  
C. Occasionally 5  
D. Seldom 3  
E. Never 1 |
| 4. Demonstrates the ability to apply a consistent and effective method for making decisions. | A. Always 10  
B. Often 7  
C. Occasionally 5  
D. Seldom 3  
E. Never 1 |
| 5. Effectively sets store priorities.                                    | A. Always 10  
B. Often 7  
C. Occasionally 5  
D. Seldom 3  
E. Never 1 |
| 6. Uses appropriate group techniques to establish and maintain productive team process within the management team. | A. Always 10  
B. Often 7  
C. Occasionally 5  
D. Seldom 3  
E. Never 1 |
| 7. Uses participative management, when appropriate, with the management team to identify and resolve store problems. | A. Always 10  
B. Often 7  
C. Occasionally 5  
D. Seldom 3  
E. Never 1 |
| 8. Effectively identifies store training needs.                          | A. Always 10  
B. Often 7  
C. Occasionally 5  
D. Seldom 3  
E. Never 1 |
| 9. Effectively manages all store training.                               | A. Always 10  
B. Often 7  
C. Occasionally 5  
D. Seldom 3  
E. Never 1 |
<table>
<thead>
<tr>
<th>BEHAVIOR</th>
<th>FREQUENCY</th>
</tr>
</thead>
</table>
| 10. Demonstrates personal time management skills in setting priorities, organizing time, and applying successful time management techniques. ("To Do" lists, planning calendar, etc.) | A. Always ...................................... 10  
B. Often .......................................... 7  
C. Occasionally .............................. 5  
D. Seldom ....................................... 3  
E. Never ......................................... 1  |
| 11. Effectively delegates tasks to subordinates, including appropriate follow-up. | A. Always ...................................... 10  
B. Often .......................................... 7  
C. Occasionally .............................. 5  
D. Seldom ....................................... 3  
E. Never ......................................... 1  |
| 12. Develops and successfully implements local store marketing programs. | A. Always ...................................... 10  
B. Often .......................................... 7  
C. Occasionally .............................. 5  
D. Seldom ....................................... 3  
E. Never ......................................... 1  |
| 13. Demonstrates the ability to effectively communicate with subordinates (verbally). | A. Always ...................................... 10  
B. Often .......................................... 7  
C. Occasionally .............................. 5  
D. Seldom ....................................... 3  
E. Never ......................................... 1  |
| 14. Effectively provides positive verbal recognition when appropriate to create a motivational environment for each subordinate. | A. Always ...................................... 10  
B. Often .......................................... 7  
C. Occasionally .............................. 5  
D. Seldom ....................................... 3  
E. Never ......................................... 1  |
| 15. Effectively administers verbal reprimands and redirects subordinates' activities when necessary. | A. Always ...................................... 10  
B. Often .......................................... 7  
C. Occasionally .............................. 5  
D. Seldom ....................................... 3  
E. Never ......................................... 1  |
| 16. Effectively evaluates store crew staffing needs. | A. Always ...................................... 10  
B. Often .......................................... 7  
C. Occasionally .............................. 5  
D. Seldom ....................................... 3  
E. Never ......................................... 1  |
| 17. Maintains an on-going crew recruitment program. | A. Always ...................................... 10  
B. Often .......................................... 7  
C. Occasionally .............................. 5  
D. Seldom ....................................... 3  
E. Never ......................................... 1  |
| 18. In dealing with all pieces of store equipment: | A. Always ...................................... 10  
B. Often .......................................... 7  
C. Occasionally .............................. 5  
D. Seldom ....................................... 3  
E. Never ......................................... 1  |
| a. Is familiar with basic component parts of each piece of equipment and how they function. | A. Always ...................................... 10  
B. Often .......................................... 7  
C. Occasionally .............................. 5  
D. Seldom ....................................... 3  
E. Never ......................................... 1  |
18. BEHAVIOR

<table>
<thead>
<tr>
<th>Behavior</th>
<th>FREQUENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>b. Can successfully troubleshoot equipment malfunctions and complete</td>
<td>A. Always</td>
</tr>
<tr>
<td>minor adjustments, small parts replacement, or repair.</td>
<td>B. Often</td>
</tr>
<tr>
<td></td>
<td>C. Occasionally</td>
</tr>
<tr>
<td></td>
<td>D. Seldom</td>
</tr>
<tr>
<td></td>
<td>E. Never</td>
</tr>
<tr>
<td>c. Calls in service agencies for equipment repair only when truly</td>
<td>A. Always</td>
</tr>
<tr>
<td>needed.</td>
<td>B. Often</td>
</tr>
<tr>
<td></td>
<td>C. Occasionally</td>
</tr>
<tr>
<td></td>
<td>D. Seldom</td>
</tr>
<tr>
<td></td>
<td>E. Never</td>
</tr>
<tr>
<td>d. Can correctly perform all planned maintenance activities on Planned</td>
<td>A. Always</td>
</tr>
<tr>
<td>Maintenance System and can properly demonstrate them to others.</td>
<td>B. Often</td>
</tr>
<tr>
<td></td>
<td>C. Occasionally</td>
</tr>
<tr>
<td></td>
<td>D. Seldom</td>
</tr>
<tr>
<td></td>
<td>E. Never</td>
</tr>
</tbody>
</table>

19. Projects a positive attitude about McDonald's and the job of a Manager.

<table>
<thead>
<tr>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Always</td>
</tr>
<tr>
<td>B. Often</td>
</tr>
<tr>
<td>C. Occasionally</td>
</tr>
<tr>
<td>D. Seldom</td>
</tr>
<tr>
<td>E. Never</td>
</tr>
</tbody>
</table>

20. Use the space below to include any additional observations you may have regarding this manager's recent management behavior.

__________________________________________________________________
__________________________________________________________________
__________________________________________________________________
__________________________________________________________________
__________________________________________________________________
__________________________________________________________________
__________________________________________________________________
__________________________________________________________________
__________________________________________________________________
APPENDIX D
Edwards Personal Preference Schedule
Allen L. Edwards, University of Washington

DIRECTIONS

This schedule consists of a number of pairs of statements about things that you may or may not like; about ways in which you may or may not feel. Look at the example below.

A I like to talk about myself to others.
B I like to work toward some goal that I have set for myself.

Which of these two statements is more characteristic of what you like? If you like "talking about yourself to others" more than you like "working toward some goal that you have set for yourself," then you should choose A over B. If you like "working toward some goal that you have set for yourself" more than you like "talking about yourself to others," then you should choose B over A.

You may like both A and B. In this case, you would have to choose between the two and you should choose the one that you like better. If you dislike both A and B, then you should choose the one that you dislike less.

Some of the pairs of statements in the schedule have to do with your likes, such as A and B above. Other pairs of statements have to do with how you feel. Look at the example below.

A I feel depressed when I fail at something.
B I feel nervous when giving a talk before a group.

Which of these two statements is more characteristic of how you feel? If "being depressed when you fail at something" is more characteristic of you than "being nervous when giving a talk before a group," then you should choose A over B. If B is more characteristic of you than A, then you should choose B over A.

If both statements describe how you feel, then you should choose the one which you think is more characteristic. If neither statement accurately describes how you feel, then you should choose the one which you consider to be less inaccurate.

Your choice, in each instance, should be in terms of what you like and how you feel at the present time, and not in terms of what you think you should like or how you think you should feel. This is not a test. There are no right or wrong answers. Your choices should be a description of your own personal likes and feelings. Make a choice for every pair of statements; do not skip any.

The pairs of statements on the following pages are similar to the examples given above. Read each pair of statements and pick out the one statement that better describes what you like or how you feel. Make no marks in the booklet. On the separate answer sheet are numbers corresponding to the numbers of the pairs of statements. Check to be sure you are marking for the same item number as the item you are reading in the booklet.

If your answer sheet is printed in BLACK ink:
For each numbered item draw a circle around the A or B to indicate the statement you have chosen.

If your answer sheet is printed in OTHER THAN BLACK ink:
For each numbered item fill in the space for A or B as shown in the Directions on the answer sheet.

Do not turn this page until the examiner tells you to start.
1  A I like to help my friends when they are in trouble.
    B I like to do my very best in whatever I undertake.
2  A I like to find out what great men and women have
    B I would like to accomplish something of great signifi-
3  A Any written work that I do I like to have precise, neat,
    B I would like to be a recognized authority in some job,
4  A I like to tell amusing stories and jokes at parties.
    B I would like to write a great novel or play.
5  A I like to be able to come and go as I want to.
    B I like to be able to say that I have done a difficult
6  A I like to solve puzzles and problems that other people
    B I like to follow instructions and to do what is expected
7  A I like to experience novelty and change in my daily
    B I like to tell my superiors that they have done a good
8  A I like to plan and organize the details of any work
    B I like to follow instructions and to do what is expected
9  A I like people to notice and to comment upon my ap-
    B I like to read about the lives of great men and women.
10 A I like to avoid situations where I am expected to do
    B I like to read about the lives of great men and women.
11 A I would like to be a recognized authority in some job,
    B I like to have my work organized and planned before
12 A I like to find out what great men and women have
    B If I have to take a trip, I like to have things planned
13 A I like to finish any job or task that I begin.
    B I like to keep my things neat and orderly on my desk
14 A I like to tell other people about adventures and strange
    B I like to have my meals organized and a definite time
15 A I like to be independent of others in deciding what I
    B I like to keep my things neat and orderly on my desk
16 A I like to be able to do things better than other people
    B I like to tell amusing stories and jokes at parties.
17 A I like to conform to custom and to avoid doing things
    B I like to talk about my achievements.
18 A I like to have my life so arranged that it runs smoothly
    B I like to tell other people about adventures and strange
19 A I like to criticize people who are in a position of au-
    B I like to use words which other people often do not
20 A I like to accomplish tasks that others recognize as re-
    B I like to be able to come and go as I want to.
21 A I like to praise someone I admire.
    B I like to feel free to do what I want to do.
22 A I like to keep my letters, bills, and other papers neatly
    B I like to be independent of others in deciding what I
23 A I like to ask questions which I know no one will be
    B I like to criticize people who are in a position of au-
24 A I get so angry that I feel like throwing and breaking
    B I like to avoid responsibilities and obligations.
25 A I like to be successful in things undertaken.
    B I like to form new friendships.
26 A I like to follow instructions and to do what is expected
    B I like to have strong attachments with my friends.
27 A I like to have my life so arranged that it runs smoothly
    B I like to make as many friends as I can.
28 A I like to tell amusing stories and jokes at parties.
    B I like to write letters to my friends.
29 A I like to solve puzzles and problems that other people
    B I like to share things with my friends.
30 A I like to be independent of others in deciding what I
    B I like to understand how my friends feel about various
31 A I like to keep my things neat and orderly on my desk
    B I like to judge people by why they do something—not
32 A I like to accept the leadership of people I admire.
    B I like to tell amusing stories and jokes at parties.
33 A I like to tell amusing stories and jokes at parties.
    B I like to study and to analyze the behavior of others.
65 A I like to do things that other people regard as un-
B conventional.

When I am in a group, I like to accept the leadership 
A of someone I admire.
B

81 A I like to do things for my friends.
B

A I like to put in long hours of work without being 
B distracted.

When planning something, I like to get suggestions 
A from other people whose opinions I respect.
B

82 A I like to put myself in someone else's place and to 
B imagine how I would feel in the same situation.

A I like to tell my superiors that they have done a good 
B job on something, when I think they have.

83 A I like my friends to be sympathetic and understanding 
B when I have problems.

A I like to accept the leadership of people I admire.
B

84 A When serving on a committee, I like to be appointed 
B or elected chairperson.

A When I am in a group, I like to accept the leadership 
B of someone else in deciding what the group is go-
A ing to do.
B

85 A If I do something that is wrong, I feel that I should 
B be punished for it.

A I like to conform to custom and to avoid doing things 
B that people I respect might consider unconventional.
B

86 A I like to share things with my friends.
B I like to make a plan before starting in to do some-
A thing difficult.
B

87 A I like to understand how my friends feel about vari-
B ous problems they have to face.

A If I have to take a trip, I like to have things planned 
B in advance.
B

88 A 9 I like my friends to treat me kindly.
B I like to have my work organized and planned before 
A beginning it.
B

89 A I like to be regarded by others as a leader.
B I like to keep my letters, bills, and other papers neatly 
A arranged and filed according to some system.
B

90 A I feel that the pain and misery that I have suffered has 
B done me more good than harm.

A I like to have my life so arranged that it runs smoothly 
B and without much change in my plans.
B

91 A I like to have strong attachments with my friends.
B I like to say things that are regarded as witty and 
A clever by other people.
B

92 A I like to think about the personalities of my friends 
B and to try to figure out what makes them as they are.

A I sometimes like to do things just to see what effect 
B it will have on others.
B

93 A I like my friends to make a fuss over me when I am 
B hurt or sick.

A I like to talk about my achievements.
B

94 A I like to tell other people how to do their jobs.
B I like to be the center of attention in a group.
B

95 A I feel timid in the presence of other people I regard 
B as my superiors.
B I like to use words which other people often do not 
A know the meaning of.
B

96 A I like to do things with my friends rather than by 
B myself.
B I like to say what I think about things.
B
I like to study and to analyze the behavior of others.  
I like to do things that other people regard as unconventional.  
I like my friends to feel sorry for me when I am sick.  
I like to avoid situations where I am expected to do things in a conventional way.  
I like to supervise and to direct the actions of other people whenever I can.  
I like to do things in my own way without regard to what others may think.  
I feel that I am inferior to others in most respects.  
I like to be successful in things undertaken.  
I like to analyze my own motives and feelings.  
I like my friends to help me when I am in trouble.  
I like to do things for my friends.  
I like to argue for my point of view when it is attacked by others.  
I like to write letters to my friends.  
I feel guilty whenever I have done something I know is wrong.  
I like to have strong attachments with my friends.  
I like to share things with my friends.  
I like to accept the leadership of people I admire.  
I like to analyze my own motives and feelings.  
I like my friends to do many small favors for me cheerfully.  
I feel better when I give in and avoid a fight, than I would if I tried to have my own way.  
I like to judge people by why they do something—not by what they actually do.  
I like to judge people by why they do something—not by what they actually do.  
I like my friends to feel sorry for me when I am sick.  
I like my friends to do many small favors for me cheerfully.  
I feel that I should confess the things that I have done that I regard as wrong.  
I like my friends to sympathize with me and to cheer me up when I am depressed.  
I like to do things with my friends rather than by myself.  
I like to argue for my point of view when it is attacked by others.  
I like to think about the personalities of my friends and to try to figure out what makes them as they are.  
I like to be able to persuade and influence others to do what I want to do.  
I like my friends to sympathize with me and to cheer me up when I am depressed.  
When with a group of people, I like to make the decisions about what we are going to do.  
I like to ask questions which I know no one will be able to answer.  
I like to tell other people how to do their jobs.  
I feel timid in the presence of other people I regard as my superiors.  
I like to supervise and to direct the actions of other people whenever I can.  
I like to participate in groups in which the members have warm and friendly feelings toward one another.  
I feel guilty whenever I have done something I know is wrong.  
I like to analyze the feelings and motives of others.  
I feel depressed by my own inability to handle various situations.  
I like my friends to feel sorry for me when I am sick.  
I feel better when I give in and avoid a fight, than I would if I tried to have my own way.  
I like to be able to persuade and influence others to do what I want.  
I feel depressed by my own inability to handle various situations.  
I like to criticize people who are in a position of authority.  
I feel timid in the presence of other people I regard as my superiors.  
I like to participate in groups in which the members have warm and friendly feelings toward one another.  
I like to help my friends when they are in trouble.  
I like to analyze my own motives and feelings.  
I like to sympathize with my friends when they are hurt or sick.  
I like my friends to help me when I am in trouble.  
I like to treat other people with kindness and sympathy.  
I like to be one of the leaders in the organizations and groups to which I belong.  
I like to sympathize with my friends when they are hurt or sick.
130 A I feel that the pain and misery that I have suffered has done me more good than harm.
   B I like to show a great deal of affection toward my friends.

131 A I feel that I like to do things with my friends rather than by myself.
   B I like to experiment and to try new things.

132 A I feel that I like to think about the personalities of my friends and to try to figure out what makes them as they are.
   B I like to try new and different jobs—rather than to continue doing the same old things.

133 A I feel that I like my friends to be sympathetic and understanding when I have problems.
   B I like to meet new people.

134 A I feel that I like to argue for my point of view when it is attacked by others.
   B I like to experience novelty and change in my daily routine.

135 A I like to do things for my friends.
   B When I have some assignment to do, I like to start in and keep working on it until it is completed.

136 A I like to do things for my friends.
   B I like to move about the country and to live in different places.

137 A I like to analyze the feelings and motives of others.
   B I like to avoid being interrupted while at my work.

138 A I like to do things that are small favors for me cheerfully.
   B I like to stay up late working in order to get a job done.

139 A I like to be regarded by others as a leader.
   B I like to put in long hours of work without being distracted.

140 A I like to do something that is wrong, I feel that I should be punished for it.
   B I like to stick at a job or problem even when it may seem as if I am not getting anywhere with it.

141 A I like to write letters to my friends.
   B I like to read newspaper accounts of murders and other forms of violence.

142 A I like to predict how my friends will act in various situations.
   B I like to attack points of view that are contrary to mine.

143 A I like to make a fuss over me when I am hurt or sick.
   B I like to blame others when things go wrong for me.

144 A I like to tell other people how to do their jobs.
   B I feel like getting revenge when someone has insulted me.

145 A I feel that I am inferior to others in most respects.
   B I feel like telling other people off when I disagree with them.

146 A I like to help my friends when they are in trouble.
   B I like to do my very best in whatever I undertake.

147 A I like to travel and to see the country.
   B I like to accomplish tasks that others recognize as requiring skill and effort.

148 A I like to work hard at any job I undertake.
   B I would like to accomplish something of great significance.

149 A I like to read newspaper accounts of murders and other forms of violence.
   B I would like to write a great novel or play.

150 A I like to do small favors for my friends.
   B When planning something, I like to get suggestions from other people whose opinions I respect.

151 A I like to experience novelty and change in my daily routine.
   B I like to tell my superiors that they have done a good job on something, when I think they have.

152 A I like to stay up late working in order to get a job done.
   B I like to praise someone I admire.

153 A I feel like getting revenge when someone has insulted me.
   B When I am in a group, I like to accept the leadership of someone else in deciding what the group is going to do.

154 A I feel like being interrupted while at my work.
   B I like to make a plan before starting in to do something difficult.
172 A I like to keep my friends when they are in trouble.
B I like to be loyal to my friends.

173 A I like to do new and different things.
B I like to form new friendships.

174 A I like to finish any job or task that I begin.
B I like to keep my things neat and orderly on my desk or workspace.

175 A I like to meet new people.
B Any written work that I do I like to have precise, neat, and well organized.

176 A I like to finish any job or task that I begin.
B I like to keep my things neat and orderly on my desk or workspace.

177 A I like to do new and different things.
B I like to form new friendships.

178 A When I have some assignment to do, I like to start in and keep working on it until it is completed.
B I like to participate in groups in which the members have warm and friendly feelings toward one another.

179 A I like to attack points of view that are contrary to mine.
B I like to write letters to my friends.

180 A I like to attack points of view that are contrary to mine.
B I like to write letters to my friends.

181 A I like to be generous with my friends.
B I like to observe how another individual feels in a given situation.

182 A I like to eat in new and strange restaurants.
B I like to put myself in someone else's place and to imagine how I would feel in the same situation.

183 A I like to stay up late working in order to get a job done.
B I like to understand how my friends feel about various problems they have to face.

184 A I like to forgive my friends who may sometimes hurt me.
B I like my friends to encourage me when I meet with failure.

185 A I like to forgive my friends who may sometimes hurt me.
B I like my friends to encourage me when I meet with failure.

186 A I like to show new and different jobs—rather than to continue doing the same old things.
B I sometimes like to do things just to see what effect it will have on others.

187 A I like to stick at a job or problem even when it may seem as if I am not getting anywhere with it.
B I like people to notice and to comment upon my appearance when I am out in public.

188 A I like to experiment and to try new things.
B I like to keep working at a puzzle or problem until it is solved.

189 A I like to experiment and to try new things.
B I like to keep working at a puzzle or problem until it is solved.

190 A I like to show a great deal of affection toward my friends.
B I like to say things that are regarded as witty and clever by other people.

191 A I like to try new and different jobs—rather than to continue doing the same old things.
B I like to be regarded by others as a leader.

192 A I like to do new and different things.
B I like to be loyal to my friends.

193 A I like to do new and different things.
B I like to form new friendships.
| 195 | A | I get so angry that I feel like throwing and breaking things.  
    | B | I like to tell other people how to do their jobs. |
| 196 | A | I like to show a great deal of affection toward my friends.  
    | B | When things go wrong for me, I feel that I am more to blame than anyone else. |
| 197 | A | I like to move about the country and to live in different places.  
    | B | If I do something that is wrong, I feel that I should be punished for it. |
| 198 | A | I like to stick at a job or problem even when it may seem as if I am not getting anywhere with it.  
    | B | I feel that the pain and misery that I have suffered has done me more good than harm. |
| 200 | A | I feel like blaming others when things go wrong for me.  
    | B | I feel that I am inferior to others in most respects. |
| 201 | A | I like to do my very best in whatever I undertake.  
    | B | I like to help other people who are less fortunate than I am. |
| 202 | A | I like to do new and different things.  
    | B | I like to treat other people with kindness and sympathy. |
| 203 | A | When I have some assignment to do, I like to start in and keep working on it until it is completed.  
    | B | I like to help other people who are less fortunate than I am. |
| 205 | A | I like to attack points of view that are contrary to mine.  
    | B | I like my friends to confide in me and to tell me their troubles. |
| 206 | A | I like to treat other people with kindness and sympathy.  
    | B | I like to travel and to see the country. |
| 207 | A | I like to conform to custom and to avoid doing things that people I expect might consider unconventional.  
    | B | I like to participate in new fads and fashions. |
| 208 | A | I like to work hard at any job I undertake.  
    | B | I like to experience novelty and change in my daily routine. |
| 210 | A | I feel like telling other people off when I disagree with them.  
    | B | I like to participate in new fads and fashions. |
| 211 | A | I like to help other people who are less fortunate than I am.  
    | B | I like to finish any job or task that I begin. |
| 212 | A | I like to move about the country and to live in different places.  
    | B | I like to put in long hours of work without being distracted. |
| 213 | A | If I have to take a trip, I like to have things planned in advance.  
    | B | I like to keep working at a puzzle or problem until it is solved. |
| 215 | A | I like to tell other people what I think of them.  
    | B | I like to avoid being interrupted while at my work. |
EDWARDS PERSONAL PREFERENCE SCHEDULE

<table>
<thead>
<tr>
<th>s.s.</th>
<th>class</th>
<th>test book</th>
</tr>
</thead>
</table>

For each numbered item in the booklet select either statement A or B and encircle either A or B alongside the corresponding number on this sheet.

For example:

<table>
<thead>
<tr>
<th>item</th>
<th>A/B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 A</td>
<td>B</td>
</tr>
</tbody>
</table>

DO NOT WRITE IN THESE COLUMNS

n r c s  
such --- --- --- |
def --- --- --- |
ord --- --- --- |
exh --- --- --- |
aut --- --- --- |
aff --- --- --- |
int --- --- --- |
suc --- --- --- |
dom --- --- --- |
aba --- --- --- |
nur --- --- --- |
chg --- --- --- |
end --- --- --- |
het --- --- --- |
agg --- --- --- |
con --- --- --- |
APPROVAL SHEET

The dissertation submitted by Thomas L. Watson has been read and approved by the following committee:

Dr. Todd Hoover, Director
Associate Professor, Curriculum and Instruction
Loyola

Dr. Barney Berlin
Associate Professor, Curriculum and Instruction
Loyola

Dr. Diane Schiller
Assistant Professor, Curriculum and Instruction
Loyola

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the dissertation is now given final approval by the Committee with reference to content and form.

The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

4/17/85
Date

Todd Hoover
Director's Signature