This is to certify that the

thesis entitled

AN EMPIRICAL TEST OF THE HIERARCHY

THEORY OF TRAINING EVALUATION

presented by

Ronald W. Clement

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Management

Frederick R. Wickert

Major professor

Date June 21, 1978

AN EMPIRICAL TEST OF THE HIERARCHY THEORY OF TRAINING EVALUATION

By

Ronald W. Clement

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Management

1978

ABSTRACT

AN EMPIRICAL TEST OF THE HIERARCHY THEORY
OF TRAINING EVALUATION

By

Ronald W. Clement

With regard to the evaluation of training programs, the world of training in recent years has tended to accept a hierarchical model. The hierarchy is seen as being composed of several levels of criteria: favorable training effects at the lowest criterion level are seen to be necessary but not sufficient for favorable training effects at the next criterion level, and so on to the highest criterion level.

One goal of the current study was to evaluate a training course in Michigan state government by collecting data relevant to each level of an abbreviated version of one such hierarchical model. Therefore, it was possible to use the same data to test the hierarchical model. The linkages of the model (i.e., the hypotheses) tested were the following:

--An improvement in reactions (level 1) will be a necessary but not sufficient condition for an improvement in learning (level 2) to occur.

--An improvement in learning (level 2) will be a necessary but not sufficient condition for improvements in job behavior (level 3) to occur.

--Improvements in job behavior (level 3) will be necessary but not sufficient conditions for improvements in organizational variables (level 4) to occur.

It was assumed that the above statements regarding the hierarchy referred to general training effects because I was not sensitized to the importance of differentiating between general and specific training effects. Therefore, a partial multicontent-multimethod matrix approach was used to test the hierarchy in the current study.

Considering the content dimension of the matrix, a training effect may fall anywhere along a continuum from very general to very specific content; further, to get specific training effects at a higher level of the hierarchy, it is probably necessary to get content-related specific training effects at the next lower level of the hierarchy. Considering the method dimension, the instruments used to measure the training effects at each level of the hierarchy may fall anywhere along a continuum from very subjective to very objective.

The current study measured general training effects at the first two levels of the hierarchy (improvements in reactions and learning) and specific effects at the next two highest levels (improvements in the job behaviors of managing, communicating, and job training, and improvement in the organizational variable of subordinate satisfaction with supervisory style). The instruments used to measure these training effects included self-reports (rather subjective--

used to measure improvements in reactions, managing behavior, communicating behavior, and job training behavior), a knowledge test (rather objective--used to measure the improvement in learning), and subordinate reports (relatively objective--used to measure the improvement in subordinate satisfaction with supervisory style).

Partial support was found for the first two hypotheses specified above. Perhaps more important for those who intend to test the hierarchy theory and those who intend to use the hierarchy theory to evaluate a training course are the following implications:

--Attempt to include training effects at each level of the hierarchy from the full range of points along the content dimension of the multicontent-multimethod matrix. Also attempt to identify--through assessment of training needs--the relatively specific training effects at one level of the hierarchy that might be expected to be content-related to certain relatively specific training effects at the next highest level of the hierarchy.

--Attempt to use several measurement methods falling along the full range of the method continuum of the multicontent-multimethod matrix.

To my family with love

for their support and understanding.

ACKNOWLEDGMENTS

To my committee:

Frederic R. Wickert, Chairman

Eugene E. Jennings

Lawrence W. Foster

and to my colleague:

James M. McFillan

many thanks for their advice and support.

TABLE OF CONTENTS

Chapter | Page

LIST OF TABLES

LIST OF FIGURES

Chapter I

Introduction

BACKGROUND AND PURPOSE

The evaluation of a training program usually requires the examination of several criteria of success. "There are few, if any, single measures that can adequately reflect the complexity of most training programs..." (Goldstein, 1974, p. 58). This is probably true whether the training is directed at supervisors or nonsupervisors, in public or in private employment.

The world of training, in addressing itself to the problems of the complexity of criteria applicable to evaluating training programs, in recent years has tended to accept a hierarchical model. The hierarchy is seen as being composed of several levels of criteria: favorable training effects at the lowest criterion level are seen to be necessary but not sufficient conditions for favorable training effects at the next highest criterion level, and so on to the highest criterion level.

Both Kirkpatrick (1967) and Hamblin (1974) have proposed hierarchical models of training evaluation; the latter's proposal is an extension of the former's. The purpose of the current study was to test an abbreviated yet adequately representative version of the Hamblin model. Before

introducing the abbreviated version tested, both the Kirkpatrick model and the Hamblin model are described below. Since the Kirkpatrick model preceded and formed a foundation for the Hamblin model, the former is described first.

THE KIRKPATRICK MODEL

Figure I-1 shows the four levels of training effects in the Kirkpatrick model of training evaluation: reactions (level 1), learning (level 2), job behavior (level 3), and results (level 4). Each level is described below.

Levels of Training Effects	Explanation
1. Reactions	Satisfaction with the training program.
2. Learning	Principles, facts, and skills learned in the training program.
3. Job Behavior	Improvements in job behavior arising from the training program.
4. Results	Tangible outcomes for the trainee's organization due to the training program.

Figure I-1

Levels of Training Effects in the Kirkpatrick Model

Reactions (Level 1)

Kirkpatrick defined reactions (level 1) "...as how well the trainees liked a particular training program. Evaluating in terms of reactions is the same as measuring the feelings of the conferees" (Kirkpatrick, p. 88). He claimed that measurement of reactions was important because the reactions of a few key trainees might influence the decisions of top management regarding the continuance of a program. It should

be added that management probably expects reactions to be measured, if only because they have been measured so frequently in the past: in a survey of the evaluation practices of 110 organizations, Catalanello and Kirkpatrick (1968) found that 77 percent had measured reactions.

Learning (Level 2)

Kirkpatrick defined learning as "...the principles, facts, and skills which were understood and absorbed by the conferees. In other words, it does not include the on-the-job use of these principles, facts, and skills" (p. 96). He said that evidence of learning was important because this allowed training directors to better sell management on future training courses.

Job Behavior (Level 3)

This level in the Kirkpatrick model refers to the improvements in job behavior arising from a training program. "Evaluation of training programs in terms of on-the-job behavior is more difficult than the reactions [level 1] and learning [level 2] evaluations described in the two preceding sections. A more scientific approach is needed, and many factors must be considered" (Kirkpatrick, p. 101). Kirkpatrick said that evaluation of job behavior improvements was necessary to increase the effectiveness of training programs and to show the benefits of training programs to top management.

Results (Level 4)

"The objectives of most training programs can be stated in terms of results such as reduced turnover, reduced costs, improved efficiency, reduction in grievances, increase in quality and quantity of production, or improved morale... From an evaluation standpoint, it would be best to evaluate training programs directly in terms of results desired" (Kirkpatrick, p. 105). Kirkpatrick goes on to explain, however, that it is often difficult to evaluate in terms of results because one cannot measure how much of an improvement in results is due to training as compared to other factors. Because of this difficulty, he suggested that training directors focus primarily on reactions (level 1), learning (level 2), and job behavior improvements (level 3) when evaluating training programs.

Summary

The four levels of the Kirkpatrick model are the reactions of trainees toward a training program (level 1), their learning (level 2), their improvements in job behavior (level 3), and the improvements in results for the organization (level 4). This model seems to have been the foundation for the development of the Hamblin model.

THE HAMBLIN MODEL

Hamblin said that prior models of training evaluation did "...not adequately explain the interconnections between the different types of training evaluation and the ways in which they can be combined" (p. 13). He extended the

Kirkpatrick model in two ways. First, he divided the results level (level 4) into two parts: organizational variables (level 4) and ultimate value variables (level 5). Second, he proposed that the five levels of his model formed a hierarchy of training effects. Each of these extensions will now be described in greater detail.

Division of Results Level

The first three levels of the Hamblin model are the same as those in the Kirkpatrick model: reactions (level 1), learning (level 2), and improvements in job behavior (level 3). But Hamblin divided Kirkpatrick's results level into two parts because he believed it was

> ...useful to distinguish between, on the one hand, changes in the way in which the organization works (organizational variables)--[level 4], and, on the other hand, changes in the extent to which the organization achieves its ultimate goals (ultimate value variables)--[level 5] (p. 14).

Examples of organization variables (level 4) are improvements in productivity, quality, and employee morale; ultimate value variables are likely to be economic (e.g., improvements in revenues, costs, and profits).

The Hierarchy of Training Effects

Hamblin suggested that the five levels of training effects formed a hierarchy such that:

--Favorable reactions (level 1) will be a necessary but not sufficient condition for learning (level 2) to occur.

--Learning (level 2) will be a necessary but not sufficient condition for improvements in job behavior

(level 3) to occur.

--Improvements in job behavior (level 3) will be necessary but not sufficient conditions for improvements in organzational variables (level 4) to occur.

--Improvements in organizational variables (level 4) will be necessary but not sufficient conditions for improvements in ultimate value variables (level 5) to occur.

These linkages are shown in Figure I-2. The vertical arrows represent the hierarchical linkages among the five levels of training effects. The horizontal arrows represent certain other events external to the hierarchy that tend to weaken the hierarchical linkages among the five levels of training effects. In other words,

> When reactions [level 1] are unfavorable, relevant learning [level 2] will not occur. When reactions [level 1] are favorable, relevant learning [level 2] may occur (depending on other factors).
>
> When relevant learning [level 2] does not occur, job behavior improvements [level 3] will not occur. When relevant learning [level 2] occurs, job behavior improvements [level 3] may occur (depending on other factors).
>
> And so forth at the other levels (Hamblin, Note 1).

To summarize: favorable training outcomes at one level do not guarantee favorable outcomes at the next higher level because of events external to the hierarchy that also affect the success of training.

Furthermore, the number of such external events increases greatly as one proceeds in evaluation from reactions (level 1) to the ultimate value level (level 5). For example, learning

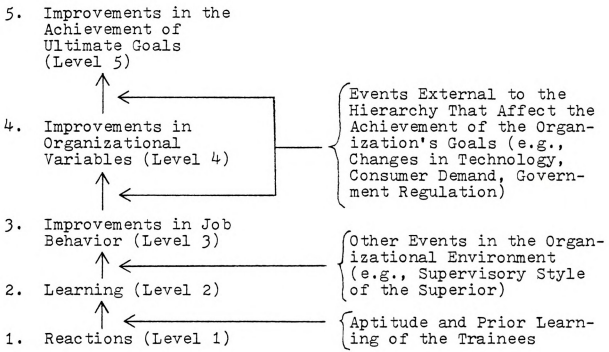


Figure I-2

The Hamblin Model, Showing the Linkages Within the Hierarchy of Training Effects and the Influence of External Events

(level 2) may be dependent upon the aptitude and prior learning of the trainees as well as their reaction to a training course (level 1); the next level, improvements in job behavior (level 3), may be dependent upon many events in the organizational environment (e.g., supervisory style of the superior, expectations of subordinates, and organization procedures) as well as the learning (level 2) that has occurred. Going to the next two levels, the events influencing the success of training at the organization (level 4) and ultimate value (level 5) levels include all those events that affect the achievement of the organization's goals.

This means that evaluation of the higher levels of the

hierarchy is more difficult than evaluation of the lower levels:

> It should be clear by now that there are many occasions when it is impossible to evaluate training at levels 4 and 5. It is too difficult to follow through all the links in the chain and see their interconnections. Provided that we have evaluated at all the levels, it may be possible to make intelligent guesses about the extent to which organizational changes have been caused by training. But usually it is very difficult, because we do not have adequate information about, or control over, the non-training activities of the firm (Hamblin, 1974, p. 24).

Hamblin's theory is that evaluation at each of the levels of the hierarchy would help to identify the reasons why successful training outcomes might occur at one level (e.g., improvements in job behavior--level 3) but not at the next highest level (e.g., improvements in organizational variables--level 4). But he said that for evaluation at a given level to be meaningful, evaluation of the next lower level in the hierarchy had to be performed.

He cited two studies to illustrate this point. Handyside (1956) measured organizational variables (level 4) after a training program and found the training program had been ineffective at this level. He was unable to say why, according to Hamblin, because he failed to measure the three previous levels of the hierarchy: reactions (level 1), learning (level 2), and improvements in job behavior (level 3). Fleishman (1953) found the reactions (level 1) and learning (level 2) of trainees in a human relations course to be favorable; but the supervisory style of the trainee's supervisor, and not the training, seemed to predict the

after-training job behavior (level 3). Since he had found that learning (level 2) had occurred, Fleishman was able to focus on the organizational environment to discover this problem. More will be said about this study in Chapter II.

Summary

The five levels of the Hamblin (1974) model are the reactions of trainees toward a training program (level 1), their learning (level 2), their improvements in job behavior (level 3), the improvements in organizational variables (level 4), and the improvements in ultimate value variables (level 5). Hamblin proposed that these levels of training effects form a hierarchy, but he said that favorable outcomes at one level of the hierarchy do not guarantee favorable outcomes at the next highest level because of the influence of certain events external to the hierarchy. The Hamblin model may be summarized as follows:

--Favorable reactions (level 1) will be a necessary but not sufficient condition for learning (level 2) to occur.

--Learning (level 2) will be a necessary but not sufficient condition for improvements in job behavior (level 3) to occur.

--Improvements in job behavior (level 3) will be necessary but not sufficient conditions for improvements in organizational variables (level 4) to occur.

--Improvements in organizational variables (level 4) will be necessary but not sufficient conditions

for improvements in ultimate value variables (level 5) to occur.

ABBREVIATIONS OF THE MODEL TESTED

The version of the Hamblin model that was tested in the current study was abbreviated in two ways. First, the study focused only on the first four levels of the hierarchy of training effects: reactions (level 1), learning (level 2), improvements in job behavior (level 3), and improvements in organizational variables (level 4); I could not identify ultimate value variables (level 5) that could be meaningfully measured at the research site under study. More will be said about this problem in Chapter III. Second, of the many events external to the hierarchy that might tend to influence the linkages within the hierarchy, the current study focused only on those in the organizational environment (e.g., supervisory style of the superior, expectations of subordinates, and organization procedures). These are the events that Hamblin said might affect the linkage between learning (level 2) and improvements in job behavior (level 3). More will also be said about this in Chapter III.

THEORETICAL HYPOTHESES

The prior description of the Hamblin model provides the rationale for the theoretical hypotheses specified below. These are the hypotheses relevant to the abbreviated version of the model tested in the current study.

--Favorable reactions (level 1) will be a necessary but not sufficient condition for learning (level 2) to occur.

--Learning (level 2) will be a necessary but not sufficient condition for improvements in job behavior (level 3) to occur.

--Improvements in job behavior (level 3) will be necessary but not sufficient conditions for improvements in organizational variables (level 4) to occur.

Chapter II

Literature Review

INTRODUCTION

No studies testing the entire Hamblin hierarchy were found in the literature. However, several studies were found that evaluated training courses at two or more levels of his five-level model. Although these studies were not intended specifically as tests of the Hamblin hierarchy, as a group they are relevant with respect to the extent to which the causal linkages suggested by Hamblin might be empirically supported.

Several preliminary comments regarding these studies are necessary. First, they do not represent an exhaustive search of the literature on training evaluation. They are merely what is probably a representative sample of the evaluation studies reported in the literature. Undoubtedly, further search would reveal additional studies. For example, Campbell, Dunnette, Lawler, and Weick (1970) reviewed 73 evaluation studies. They used

> H.O. Martin's (1957) notion of internal and external criteria...to initially divide the studies into two groups, those which attempt to demonstrate some change in behavior relevant to the training itself and those which are directly concerned with changes in job behavior. In the former category are criteria such as attitude measures, tests of decision-making ability, and general opinions concerning whether or not the

> training was 'successful' [i.e., both levels 1 and 2 in the abbreviated Hamblin hierarchy]. External criteria include objective measures of unit and manager performance in the job situation, turnover or grievances in a manager's unit, and ratings of performance by superiors, peers, or subordinates [i.e., levels 3 and 4 in the abbreviated Hamblin hierarchy] (p. 287-288).

Campbell et al. found that "...the preponderance of studies employed internal criterion measures..." (p. 321). That is, they found that most studies had evaluated reactions (level 1) and learning (level 2). Although they mentioned Kirkpatrick's levels, they used Martin's cruder levels instead; Hamblin's book was not yet available, so they could not have used its concepts.

The second preliminary comment regarding the studies reviewed in this chapter is that, like Campbell et al., the authors of the studies did not describe their work in terms of Hamblin's hierarchical structure. For this reason the present author found it necessary to interpret each study to determine which level or levels of the hierarchy of training effects were involved.

Third, no single study could be expected to evaluate an entire level of the proposed Hamblin hierarchy. Each level of the hierarchy (in effect a general construct) inescapably is made up of a large number of specific constructs. As a group, therefore, the studies reviewed here have had to take into account no more than the very small sample of these specific constructs available at the "levels" of special concern to each study.

Finally, several of the studies reviewed in this chapter have, in addition to effects that belong within the hierarchy proper, examined certain of the type of variable that Hamblin, in his adumbration of the hierarchy, mentioned could influence the causal linkages within the hierarchy. For example, Hamblin said the external variables of the aptitude and prior learning (level 2) of trainees, as well as their reactions (level 1) within the hierarchy, might influence the amount of learning (level 2) that occurs in a training course. Furthermore, the number of such externally influencing variables increases greatly as one proceeds in evaluation of training from the reactions level to the ultimate value level. However, the studies reviewed here, it so happens, for the most part focus on the variables in the organizational environment that influence the transfer of learning to the job situation, or, in Hamblin's terms, the causal linkage between learning (level 2) and improvements in job behavior (level 3).

The following review of the literature describes these studies in two separate parts: (1) those studies that evaluated training at no more than two levels (e.g., learning (level 2) and improvements in job behavior (level 3)), and, (2) those studies that evaluated training at more than two levels.

STUDIES THAT EVALUATED TRAINING AT TWO LEVELS

Studies described in this section include those that evaluated learning (level 2) and improvements in job behavior (level 3).

Latham, Wexley, and Purcell (1975) reported results that showed indirect support for the causal linkage proposed by Hamblin between learning (level 2) and improvements in job behavior (level 3). These three researchers were concerned not with the hierarchy but with comparing the effectiveness of a workshop approach with a group-discussion approach in training managers to minimize errors in rating job applicants. The study measured four such errors: contrast effects, halo effect, similarity, and first impressions. Data on both learning and the improvements in job behavior, as perceived by the trainees, happened to have been collected.

The workshop approach included presentation of videotapes of hypothetical job candidates being appraised by a manager. Trainees then discussed how they would have rated the candidate and how they thought the manager in the videotape would have rated him. The videotapes showed the four different rating errors being made, and the discussion was to help the trainees identify how such errors were made and how they could be avoided.

In the group-discussion approach, the trainer first defined each rating error. A series of group discussions followed in which trainees generated personal examples of each error and discussed ways to avoid them. A control group, which received no training, was also included in the study.

Subjects from the workshop, the group-discussion approach, and the control group were tested six months after the training. Together they viewed videotapes again showing

hypothetical job candidates being interviewed. Subjects rated the candidate shown in each of eight such videotapes. The accuracy of this rating was the measure of learning.

The measure of improvements in job behavior consisted of "...a 9-point scale on which [trainees] rated the extent to which they felt the program had been beneficial to them on their jobs during the [six months since the training]" (p. 553). I categorized these ratings as, in effect, measures of improvements in job behavior (level 3) rather than a measure of reactions (level 1) because the instrument focused on picking up perceived improvements in job behavior rather than affective reactions to the training course itself.

Trainees from the workshop, the results showed, learned more than the group-discussion trainees; the former committed no rating errors, while the latter committed impressions errors. Trainees from the workshop also perceived greater improvements in job behavior (level 3) than did the group-discussion trainees. "The mean rating on the...measure of job behavior improvements for the workshop trainees was 8.08...while that for the group-discussion trainees was 6.25... A rating of 9 indicated the greatest benefit. The difference in the ratings between the two training programs was highly significant..." (p. 533).

Indirectly, the results provided support for one of the types of causal linkages specified in the Hamblin model: the group that learned more (level 2) reported greater transfer of the learning to the job (level 3).

The control group committed similarity, halo and contrast errors; however, this result is unrelated to testing the Hamblin hierarchy because unfortunately only the learning (level 2) but not the control group's perceptions of improvements in job behavior (level 3) was measured.

Hand, Richards, and Slocum (1973) reported the results of 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, while the other group of trainees viewed their organization as less democratic and more structured. Eighteen months after the course, both groups displayed favorable changes in attitudes (level 2) and improvements in job behavior (level 3). The results supported Hamblin's contention that learning (level 2) and improvements in job behavior (level 3) are linked.

Further analysis of the results also showed support for Hamblin's observation that improvements in job behavior (level 3) are influenced by variables external to the training as well as by the learning that occurs within the training course: the group from "consultative" organizations had received significantly higher performance ratings than the other training group. Hand et al. explained that the "consultative" group had returned to an organizational environment that supported the attitudes and behaviors learned in the training while the other group returned to a less

reinforcing organizational environment.

Further support of the enormous influence of such external variables appeared in House's (1968) review of empirical evidence in leadership training regarding the influence of variables in the organizational environment on the transfer of learning to the job situation (level 3). He concluded that the variables that influence the transfer of learning were less the training than the formal authority system, the exercise of formal authority by the superior of the trainee, and the primary work group of the trainee.

STUDIES THAT EVALUATED TRAINING AT MORE THAN TWO LEVELS

Studies described in this section include those that evaluated reactions (level 1), learning (level 2), and improvements in job behavior (level 3) and those that evaluated learning (level 2), improvements in job behavior (level 3), and improvements in organizational variables (level 4). The former are described first.

Studies That Evaluated Reactions (Level 1), Learning (Level 2), and Improvements in Job Behavior (Level 3)

Fleishman (1953), whose study was mentioned in Chapter I, found that both favorable reactions (level 1) and resultant learning (level 2) occurred after a human relations course that taught foremen to be more considerate to their subordinates. Reactions toward the course were measured by means of verbal reports from the trainees. "In general, interest in the course among foremen is very high. Enthusiasm at the verbal level is almost universal" (p. 222).

Learning of new supervisory attitudes was measured by means of the Leadership Opinion Questionnaire. "By giving our attitude questionnaires to foremen the first day and again the last day of training we could get some indication of changes produced during the training course. The results of this before-and-after evaluation indicated a general increase in 'consideration' attitudes...during the course" (p. 212).

However, the extent to which the newly-learned attitudes (level 2) were reflected in the trainees' job behavior (level 3) depended on the supervisory style of the trainees' superiors and the leadership expectations of the trainees' subordinates. "...[T]hose foremen who operated under [superiors] higher in 'consideration' tended themselves to be more 'considerate' with their workers...It was also found that the behavior of foremen who returned to 'climates' consistent with what was taught in training conformed more closely to the leadership expectations of their work groups. No such improvement was found among foremen who returned to 'climates' at variance with the training course" (p. 220).

Hariton (1951) also investigated the reactions (level 1), learning (level 2), and improvements in job behavior (level 3) of trainees who attended a course in human relations principles.

The experimental group included foremen from two divisions of a large public utility who received the training. The control group included foremen from the same two divisions

who did not receive the training.

The reactions toward the course on the part of the foremen in the experimental group were found to be quite favorable (level 1), and it was found that they believed they had learned the principles taught in the course (level 2). Measurement of the reactions (level 1) and the perceived learning (level 2) of the foremen in the control group, however, showed that the course had had no significant effect on them at either level.

However, as reported by both House and Fleishman, the extent to which the learning of the experimental group was reflected in improvements in job behavior (level 3) depended more upon variables in the organizational environment than on the amount of learning (level 2) that occurred during the course. The improvements in job behavior (level 3) of the foremen in the experimental and control groups were determined by measuring the changes in the perceptions of their subordinates toward the foremen. Neither the results for the total experimental group nor those for the total control group showed significant improvements in the perceptions of the subordinates (level 3). But when Hariton examined the data for each division separately, he found more revealing results. There was a significant increase in subordinate satisfaction in one experimental division (level 3) and a significant decrease in the other (level 3). Foremen from the experimental division in which subordinate satisfaction improved (level 3) perceived their organizational environment

to be more supportive of the training principles than did the foremen from the experimental division where subordinate satisfaction decreased (level 3). For example, the former perceived greater opportunity to try out their ideas on the job, received more encouragement from their superiors to use the principles of the course, and expected that their own needs would be met if they used the principles presented in the course. Hariton concluded: "Training foremen in new human relations techniques will be most effective...when the foremen are motivated to change[,] when the climate within which the foremen operate is conducive to change...[and when] the attitudes and practices of higher levels of supervision [are] consistent with the course content" (p. 95).

Fromkin, Brandt, King, Sherwood, and Fisher (1974) evaluated reactions (level 1), learning (level 2), and improvements in job behavior (level 3) in a human relations course intended to teach police officers new interpersonal skills. An experimental group of 23 officers and a control group of 27 officers were chosen from eight Midwestern police departments. Instruments included "...the Problem Analysis Questionnaire, Rotter's I-E Scale, Rokeach's Dogmatism Scale, Rokeach's Value Questionnaire, a Training Description Questionnaire, the Community Attitude Survey, and the PIE and Behavioral Questionnaire [all level 2] which were constructed specifically for the program. In addition, 131 citizens who had recent encounters with the officers were contacted by phone and asked to complete a questionnaire"

(p. 206-207) (level 3).

The reactions (level 1) of only the experimental officers were measured, and it was found that they strongly supported the program. Regarding learning (level 2), it was found that "Experimental officers seemed more attentive to salient cues and could list a slightly greater number of alternative police behaviors when viewing videotapes of typical encounters with the public than could controls. Trained officers also showed a decrease in negative stereotyping of the poor and recognized a greater need to understand this group than did controls [level 2]...Experimental officers [also] significantly decreased their valuation of an 'exciting life' and increased their ratings for 'inner harmony' [more] than controls [level 2] " (p. 207).

Regarding improvements in job behavior (level 3), it was found that "...experimental officers assumed greater responsibility for the outcomes of their interactions with citizens and recognized an increased need for remaining flexible in these encounters when compared to control officers... Citizens perceived the experimental officers as more explanatory of their actions than control officers [all level 3]" (p. 207).

In summary, the Fleishman (1953), Hariton (1951), and Fromkin et al. (1974) studies supported Hamblin's contention that the reactions of trainees (level 1) are linked to their learning (level 2) and that their learning (level 2) is linked to their improvements in job behavior (level 3). The

Fleishman and Hariton studies also showed evidence that improvements in job behavior (level 3) are influenced by variables external to the training as well as by the learning that occurs within the training course (level 2).

Studies That Evaluated Learning (Level 2), Improvements in Job Behavior (Level 3), and Improvements in Organizational Variables (Level 4)

Goldstein and Sorcher (1974) evaluated learning (level 2), improvements in job behavior (level 3), and changes in organizational variables (level 4) in two applications of a training technique called Applied Learning. Briefly, this technique consisted of four stages: modeling, in which the desired job behavior was demonstrated to the trainee; role-playing, in which each trainee in turn attempted to act out the desired job behavior himself; social reinforcement, in which the trainee was rewarded with praise and recognition for successful role-playing; and, especially important for testing the hierarchy, transfer of training, in which certain steps were taken to provide that the behavior change would transfer to the job situation.

In one industrial application of Applied Learning, Goldstein and Sorcher attempted to reduce early turnover among young, black, male employees from disadvantaged backgrounds. The researchers gave two reasons for high early turnover in this group: "First, employees from disadvantaged backgrounds often have poor models to imitate in their own environments, since frequently few individuals known to such new employees work at steady jobs. Second, many foremen

hold beliefs that prevent them from doing a good job at helping new employees (especially those from a disadvantaged background) adapt successfully to the discipline and rigor of an industrial operation" (p. 71).

To overcome these two reasons for early turnover, the researchers trained both the new employees and their foremen in separate Applied Learning programs. "As presented by filmed incidents, model behavior for the supervisor, for example, showed concrete behavioral illustrations of tact, coolness, patience, thoroughness, and control. Model behavior for new employees...emphasized specific behaviors depicting that it takes some courage to succeed, the values of working at a job, and job success" (p. 71).

Both groups learned the intended job behaviors (level 2). Furthermore, comments from the trainees showed that foremen and employees were reinforcing each other's intended behaviors on the job (level 3).

The organizational variable (level 4) that was measured was turnover: "Six months after completion of training, the voluntary quit rate was almost three times higher for employees who were oriented with the usual programs than for employees who were oriented with Applied Learning" (p. 73).

In the second industrial application of Applied Learning, Goldstein and Sorcher attempted to teach a wide variety of desired job behaviors to manufacturing foremen. "As the first step, all second and third level managers in production and support operations participated in an Applied Learning

program. For purposes of this program, the managers were asked to assume a foreman's role so that they might have a better idea of what the foremen would later be taught" (p. 74). That is, the assumption underlying this program was that, by training the second and third level managers as well as the foremen, the latter would more easily be able to transfer the learning to the job (level 3). Therefore, this evaluation study too had built into it not only the learning stimuli to lead to change but also an organizational variable (level 4) that Hamblin had assumed would be as important as training.

The organizational variable that was measured in this study was the productive efficiency (level 4) of the employees who worked for the foremen who received Applied Learning. Their productive efficiency (level 4) was compared with that of a group of employees who worked for a control group of foremen who did not receive Applied Learning and whose superiors had also not been trained. Once again, the learning (level 2) that occurred in the Applied Learning program transferred to the job situation (level 3). Moreover, the level of productive efficiency (level 4) was significantly higher for workers supervised by trained foremen than for workers supervised by foremen in the control group.

Both of the above Applied Learning studies revealed indirect support for two linkages in the Hamblin hierarchy: the linkage between learning (level 2) and improvements in job behavior (level 3) and the linkage between improvements

in job behavior (level 3) and improvements in organizational variables (level 4). In addition, both showed the benefit of controlling for a variable in the organizational environment that Hamblin had judged could influence the transfer of learning to the job (level 3): the leadership style of the trainee's superior. In the first study above, this control was accomplished by training both the employees and their superiors (the foremen) in separate Applied Learning programs; in the second study above, control was achieved by having the second and third level superiors of the foremen-trainees themselves participate in the Applied Learning program.

In an evaluation study performed by Goodacre (1955), 800 supervisors and managers were randomly divided into an experimental group and a control group. Learning (level 2), improvements in job behavior (level 3), and improvements in organizational variables (level 4) were evaluated. Included were measures of job satisfaction (level 4), attitudes toward the company (level 4), and self-rated self-confidence (level 2); achievement tests based on the course content (level 2), and ratings of actual job performance (level 3), provided by the subject's immediate superior, were also used.

"In general, the control group did not change on any of the variables, but the experimental group did show improvement in the achievement tests (level 2) and in rated self-confidence (level 2). The experimental group was also given somewhat higher post-training (on-the-job) performance ratings [level 3]..." (Campbell et al., 1970, p. 289). Indirectly,

therefore, the results supported Hamblin's contention that learning (level 2) was linked to improvements in job behavior (level 3) and that improvements in job behavior (level 3) were linked to improvements in organizational variables (level 4).

SUMMARY

As a group, the above studies lent a measure of credibility to the Hamblin model. Overall, they showed indirect support for the proposed causal linkages between reactions (level 1) and learning (level 2) (Fleishman, 1953; Hariton, 1951; and Fromkin et al., 1974), between learning (level 2) and improvements in job behavior (level 3) (Latham et al., 1975; Hand et al., 1973; Fleishman, 1953; Hariton, 1951; Fromkin et al., 1974; and Goldstein and Sorcher, 1974), and between improvements in job behavior (level 3) and improvements in organizational variables (level 4) (the two studies in Goldstein and Sorcher, 1974; and Goodacre, 1955).

Several of these studies also showed that improvements in job behavior (level 3) were influenced by organizational variables external to the training as well as by the learning (level 2) that occurs within the training course (Hand et al., 1973; House, 1968; Fleishman, 1953; Hariton, 1951; and Goldstein and Sorcher, 1974).[1] The influencing organizational variables included the supervisory style of the trainee's

[1] The other three studies (Latham et al., 1975; Fromkin et al., 1974; and Goodacre, 1955) did not provide measures of external influencing variables.

superior, the leadership expectations of the trainee's primary work group, the formal structural properties of the organization, and the motivation of the trainee regarding the intended improvements in job behavior (level 3).

As could have been anticipated, no studies were found that examined the proposed causal linkage between the improvements in organizational variables (level 4) and the improvements in ultimate value variables (level 5). Furthermore, no studies were found that examined other variables (besides those in the organizational environment) that might influence the causal linkages within the proposed Hamblin hierarchy; no doubt additional empirical research would uncover many such variables. Also, no studies were located that measured variables on two or more levels and which failed to support the hierarchy. The precaution must be repeated, however, that the studies reviewed above are merely what is probably a representative sample of the studies reported in the literature.

Chapter III

Design and Methodology

INTRODUCTION

This chapter logically divides itself into two major parts. The first part lays out the research design. The second part provides the detailed methodology for implementing the research design.

The first part, then, develops the research design, step by step. Included as part of the research design is an explanation of the extent to which this particular research design controlled certain threats to internal and external validity.

The second part of the chapter describes the methodology used in the current study. Included are a description of the sample on whom data were gathered, an explanation of how the theoretical variables were operationalized, a statement of the operational hypotheses, a description of the instruments used to gather the data, and a description of the procedure for administering the instruments.

RESEARCH DESIGN

A pretest-posttest-control group design was used in the current study.

The experimental group included 75 first-level supervisors

from various departments of Michigan state government who were exposed to a training treatment. The control group included 21 other first-level supervisors who were peers of the experimental subjects. These 21 were chosen as controls because they were just as likely as the experimental subjects to be exposed to the training treatment, but in fact they were scheduled to be exposed to the training treatment at some time after the conclusion of the current study.

The subjects for both the experimental group and the control group, in effect, were randomly selected from the population of first-level supervisors, and measurements were performed on both groups before and after the experimental (training) treatment. The sample and the procedure for data collection are described in greater detail in the "Methodology" section of this chapter.

The pretest-posttest-control group design typically controls most of the threats to internal validity but not to external validity. The remainder of the current section explains (1) how the pretest-posttest-control group design largely controlled the threats to internal validity, and (2) why the threats to external validity were not matters of great concern in the current study. The control of threats to internal validity is described first.

Control of Threats to Internal Validity

The pretest-posttest-control group design typically controls the following threats to internal validity: differential selection, history, testing, maturation, statistical

regression, instrumentation, and experimental mortality. The following is a definition of each of these and an explanation of the extent to which each was controlled in the current study.

Differential selection. "This effect stems from biases in choosing comparison groups...[D]ifferences could occur between the two groups simply because each was different before the program began. This variable is best controlled by random selection of all participants..." (Goldstein, 1974, p. 76). Since the subjects for both the experimental group and the control group in the current study were randomly selected from the same target population, it was assumed that both groups were essentially equivalent. Results regarding this issue are presented in Chapter IV.

History. This variable refers to events other than the experimental treatment that occur between the pretest and the posttest and that could account for changes in the subjects. Except for intrasession effects (i.e., events occurring during instruction that are not part of the treatment), history should have affected the experimental group and the control group approximately equally.

Testing. "This variable refers to the influence of the pretest on the scores of the posttest. This is an especially serious problem for instructional programs in which the pretest can sensitize the participant to search for material... that provides correct answers on the posttest" (Goldstein, p. 75). In the current study, a pretest on reactions toward

the course (level 1) was administered only to the experimental group. Pretests regarding learning (level 2) and job behavior (level 3) were administered to both the experimental and control groups. Therefore, it was assumed that pretesting affected both groups approximately equally.

It would have been helpful to include a second experimental group and a second control group in the study, each of which would have been exposed, not to the pretests, but only to the experimental treatment and the posttests (i.e., a Solomon four-group design). This would have allowed determination of the extent to which the pretests, and not the treatment, were responsible for the results. Unfortunately, the sample size was not large enough to allow for the inclusion of these two extra groups.

Statistical regression. This variable occurs "...where groups have been selected on the basis of their extreme scores" (Campbell and Stanley, 1963, p. 5). In other words, an experimental or control group which scores extremely high or extremely low on a pretest will tend toward the normal mean on a posttest, even without being exposed to an experimental treatment. Again, since the experimental group and the control group of the current study were essentially equivalent, there was no reason to assume that statistical regression would have differentially affected the two groups.

Instrumentation. This threat to internal validity "...results from changes in the instruments that might result in differences between [pretest and posttest] scores"

(Goldstein, p. 75). This was not a valid threat in the current study because the same fixed instruments (e.g., printed scales and tests) were used in both pre- and posttests.

Experimental mortality. This variable refers to a "...differential loss of respondents from the comparison groups" (Campbell and Stanley, p. 5). For example, experimental subjects in a training course may have to drop out to return to work for an unexpected reason; or control subjects who perform poorly on a pretest may refuse to take a posttest.

Experimental mortality occurred in the current study. Seventy-five experimental subjects were originally assigned to the courses which were evaluated. Five of these failed to complete certain pretraining instruments, so these subjects' data were discarded. For the remaining 70 experimental subjects on whom complete pretraining data were gathered, complete posttraining data were received for only 50; only the data on these 50 are reported in this study. The pretraining data on the 20 who "dropped out" were compared to the pretraining data on the 50 on whom complete data were gathered. This comparison, the results of which are described in Chapter IV, showed that mortality probably did not seriously bias the findings.

Control of Threats to External Validity

The pretest-posttest-control group design "...is affected by external threats to validity, which are not as easily specified as the threats to internal validity" (Goldstein, p. 82). Campbell and Stanley identify these external threats

as the interaction of testing and the experimental treatment, the interaction of selection and the experimental treatment, and reactive arrangements. The following is a definition of each of these and an explanation of the extent to which each was a concern in the current study.

Interaction of testing and experimental treatment. This refers to a situation in which the subject is sensitized to the experimental treatment by the pretest such that the pretest, and not the experimental treatment, is responsible for the change in the subject. This is frequently a problem in attitude-change studies but may not have been a factor in the current study. Campbell and Stanley explain: "In the area of mass communications the researcher's interview and attitude-test procedures are quite atypical. But in research on teaching, one is interested in generalizing to a setting in which testing is a regular phenomenon, and no undesirable interaction of testing and experimental treatment would be present" (p. 18). In this study, teaching and its evaluation were taken for granted phenomena.

Interaction of selection and experimental treatment. This threat to external validity refers to the possibility that the experimental treatment effects validly demonstrated "...hold only for that unique population from which the experimental and control groups were jointly selected" (Campbell and Stanley, p. 19). This threat to external validity also was probably not a problem in the current study: the experimental and control groups were representative of the

population for which the study was intended.

Reactive arrangements. "The procedures employed in the experimental setting may limit the generalizability of the study. Observers and experimental equipment often make the participants aware of their participation in an experiment, which can lead to changes in behavior that cannot be generalized to those individuals who will participate in the instructional treatment when it is nonexperimental" (Goldstein, p. 77). The current study of Management II was the first evaluation of that course to use both pretraining and posttraining measures, and it examined variables on which prior evaluations of Management II had not focused (e.g., learning and improvements in job behavior). Consequently, reactive arrangements may possibly have limited the generlizability of the results, but, again, untoward effects were unlikely in view of the fact that some evaluation of training is routinely done in these programs and participants accept its presence.

Summary. The pretest-posttest-control group design of the current study controlled all the threats to internal validity except experimental mortality, which probably did not seriously affect the findings. Regarding the threats to external validity, only the somewhat reactive arrangements (e.g., the measurement of variables on which prior evaluations of Management II had not focused) may possibly have limited the generalizability of the results.

METHODOLOGY

The following details of the methodology are presented in this section: a description of the sample on whom data were gathered; the method used to operationalize the variables; a statement of the operational hypotheses; a description of the instruments used to gather the data; and a description of the procedure for administering the instruments.

Sample

The research site, the experimental group, and the control group are described in that order.

Research site. The research site was a five-day general supervisory training course called Management II, which is offered frequently on a rather routine basis to recently-appointed first-level supervisors in Michigan state government. The course was developed and conducted by a consulting organization associated with Lansing Community College. Management II covered the following topic areas:

Role of the supervisor	Communication
Planning	Job training
Human relations	Leadership
Motivation	

The behavioral objectives of Management II had not been identified prior to this proposal to evaluate the effectiveness of the course. Through discussion with the consultant and certain personnel from Michigan state government, it was determined that the three primary behavioral objectives of Management II were the following:

--The supervisor (experimental subject) will increase the percentage of his job time spent on managing as opposed

to operating activities.

--The supervisor will increase the degree to which he exhibits proper communicating behavior.

--The supervisor will increase the degree to which he uses the six steps of job training.

In other words, Management II was intended to cause, primarily, three improvements in job behavior (level 3). First, experimental subjects were expected to increase the percentage of time they spent getting the work done through their subordinates; they were to spend less time performing the technical duties themselves. Many of the experimental subjects were former operatives, and management wanted them to begin emphasizing managing versus operating activities. In the current study, this job behavior improvement was referred to as the improvement in managing behavior (level 3).

Second, experimental subjects were expected to communicate more effectively, especially with their subordinates. Management II attempted to teach them how to overcome barriers to effective communication. Examples of ways to do this, according to the course, were putting the subordinate at ease, showing an understanding of the subordinate's feelings, and eliminating distractions. In the current study, this job behavior improvement was referred to as the improvement in communicating behavior (level 3).

Third, experimental subjects were expected to train their subordinates more effectively. Specifically, they were expected to follow a six-step "job training" sequence

suggested by Management II. These steps were (1) determining the need for training, (2) defining the purpose of the skill to the subordinate, (3) explaining the new skill in detail, (4) demonstrating the skill, (5) allowing the subordinate to practice the skill, and (6) discussing with the subordinate any questions or difficulties. In the current study, this job behavior improvement was referred to as the improvement in job training behavior (level 3).

Two course offerings were evaluated in the study. Course I was held in Lansing, Michigan, from January 10 through January 14, 1977, and Course II was held in Detroit, Michigan, from February 28 through March 4, 1977. Forty experimental subjects attended Course I, and 35 attended Course II. During each course, the group of experimental subjects was split into two smaller groups, each assigned to a different room, to allow for more effective participation and group discussion. Five instructors taught each course: two taught about half the experimental subjects in one room, and three taught about half the experimental subjects in a second room. Each instructor was considered capable of teaching any part of the course. A wide variety of training techniques was used, including lectures, roleplays, discussions, buzz groups, and films.

Experimental group. The experimental subjects were 75 first-level supervisors from various departments of Michigan state government. Demographic data on these subjects were not available. Supervisors nominated for Management II

typically have at least six months of supervisory experience before taking the course, but there is no upper limit on the length of supervisory experience to bar entry to the course. Furthermore, other criteria such as performance, future potential, and so forth are not considered in selecting supervisors to attend the course. Consequently, the experimental group was assumed to be a random sample of the population of first-level supervisors in Michigan state government.

Control group. The subjects in the control group were 21 first-level supervisors who had not attended Management II in the past, and who were not assigned to the courses evaluated in the current study. The superiors of the experimental subjects were to nominate at least one of each experimental subject's peers to serve as a control subject. These peers were first-level supervisors who were just as likely as the experimental subjects to be chosen to be exposed to the experimental treatment. Since there were no other criteria for the selection of control subjects, it was assumed that the resulting control group was also a random sample of the population of first-level supervisors in Michigan state government. Usable data were gathered for seven control subjects during the evaluation of Course I and for 14 during the evaluation of Course II. The data on these were combined to yeild one control group.

Operationalization of Variables

The current study tested an abbreviated version of the Hamblin model; specifically, the study focused on the

variables which Hamblin referred to as reactions, learning, improvements in job behavior, and improvements in organizational variables (levels 1, 2, 3, and 4). This section reviews the Hamblin definition of each of these and explains how each was operationalized in the current study. The instruments used to measure these variables are described in a later section.

Improvement in reactions (level 1). According to Hamblin, the reactions variable includes any attitudinal response of the experimental subject toward a training course. However, he emphasized only the measurement of posttraining reactions. The improvement in reactions from pretraining to posttraining might be a more meaningful measure. The pretraining reactions would be developed from two sources: (1) The assessment process, however informal it may be; that is, when selected for a training course, trainees typically are told how the course is expected to meet their needs as well as the needs of the organization. (2) Feedback from employees who previously attended the course or had occasion to learn something about it. The posttraining reaction, of course, would develop from the trainee's own course experience. An experimental subject's posttraining reaction toward a course could be somewhat favorable, but still less favorable than his pretraining reaction; this kind of finding would imply that the course did not quite meet his expectations. Measuring only posttraining reactions would not reveal this problem, whereas measurement of the improvement in reactions would.

Therefore, in the current study the reactions level of the Hamblin model (level 1) was operationalized as the improvement in reactions from pretraining to posttraining.

Improvement in learning (level 2). The learning level in the Hamblin model refers to the improvements in knowledge, skills, and attitudes of the experimental subjects. In the current study, this level was referred to as the "improvement in learning", rather than just "learning", because all the other levels also deal with "improvements"; this provided for consistency in terminology across the four levels tested in the current study.

Improvements in job behavior (level 3). According to Hamblin, evaluation at this level attempts to discover the extent to which the experimental subjects "...have applied their behavior on the job...The problem of whether learning [level 2] acquired during training is applied on the job is the problem of the transfer of learning [level 3]" (p. 20-21).

The improvements in job behavior intended by Management II were described earlier in this chapter. These were the improvement in managing behavior, the improvement in communicating behavior, and the improvement in job training behavior. Therefore, the level of job behavior improvements (level 3) in the Hamblin model was operationalized as the improvements in three job behaviors--managing, communicating, and job training--from pretraining to posttraining.

Improvements in organizational variables (level 4). At this level, Hamblin referred to "...changes in the way in which the organization works..." (p. 14). Examples of such improvements, particularly for supervisory employees, are improvements in the morale of the experimental subjects' subordinates, improvements in the quality of the work of those subordinates, and improvements in their level of absenteeism. The organizational variable on which the current study focused was the satisfaction of the experimental subjects' subordinates with the supervisory style of the experimental subject. It was expected that the experimental subjects' subordinates would be more satisfied with the supervisory style of each experimental subject to the extent that the latter was able to transfer the intended job behavior improvements to the job situation. Therefore, level 4 of the Hamblin model was operationalized as the change, from pretraining to posttraining, in subordinate satisfaction with the supervisory style of the experimental subject (level 4).[2]

Influence of variables external to the hierarchy of training effects. Hamblin said that favorable training effects at one level of the hierarchy do not guarantee favorable outcomes at the next highest level generally because of

---

[2] According to the empirical evidence (Brayfield and Crockett, 1955; Vroom, 1964) increased job satisfaction may lead to reduced absenteeism. The original design of the study called for the measurement of the improvement in absenteeism of the experimental subjects' subordinates (level 4). Unfortunately, the data gathered on absenteeism were unreliable because the instructions to the respondents did not specify what types of absenteeism to consider (e.g., leaves, vacations). Therefore, these data had to be discarded.

the influence of variables external to the hierarchy. Furthermore, the number of such variables increases greatly as one proceeds in evaluation from improvement in reactions (level 1) through the whole hierarchy of improvements to improvements in ultimate value variables (level 5). Of the many external variables that might influence the interfaces between the levels of the hierarchy, the current study focused only on those in the organizational environment that have frequently in other research been found to influence transfer of learning (level 3). Research on these particular external variables may be especially practical because these are the influencing variables over which the management of the organization may have some control. That is, by identifying such external variables, it may be possible for management to better control their influence on the outcomes of a training course.

Several of the studies reviewed in Chapter II identified some of the organizational variables that influenced the transfer of learning to the job situation (level 3) (Hand et al., 1973; House, 1968; Fleishman, 1953; Hariton, 1951; and Goldstein and Sorcher, 1974). The influencing organizational variables included the supervisory style of the experimental subject's superior, the leadership expectations of the experimental subject's primary work group, the formal structural properties of the organization, and the motivation of the experimental subject regarding the intended improvements in job behavior (level 3).

Operational Hypotheses

The above description of the operational variables provides the framework for the specification of operational hypotheses. These are the hypotheses that were tested in the current study.

The linkages between the improvement in reactions (level 1) and the improvement in learning (level 2). The abbreviated version of the Hamblin model tested in the current study states that an improvement in reactions (level 1) is a necessary but not sufficient condition for an improvement in learning (level 2) to occur. That is, a training course may or may not produce a measurable improvement in learning (level 2) but can do so only if there is an improvement in reactions (level 1). This assumption yielded the following operational hypothesis:

> Hypothesis 1: An improvement in reactions (level 1) will be a necessary but not sufficient condition for an improvement in learning (level 2) to occur.

The linkage between the improvement in learning (level 2) and the improvements in job behavior (level 3). Of interest here is that the Hamblin model tested in the current study states that an improvement in learning (level 2) is a necessary but not sufficient condition for improvements in job behavior (level 3) to occur. That is, a training course may or may not produce measurable improvements in job behavior (level 3) but will only do so if an improvement in learning (level 2) occurs. This yields hypothesis 2 with its three

parts:

Hypothesis 2a: An improvement in learning (level 2) will be a necessary but not sufficient condition for an improvement in managing behavior (level 3) to occur.

Hypothesis 2b: An improvement in learning (level 2) will be a necessary but not sufficient condition for an improvement in communicating behavior (level 3) to occur.

Hypothesis 2c: An improvement in learning (level 2) will be a necessary but not sufficient condition for an improvement in job training behavior (level 3) to occur.

The linkage between the improvements in job behavior (level 3) and the improvement in organizational variables (level 4). Of interest here is that the Hamblin model states that improvements in job behavior (level 3) are necessary but not sufficient conditions for improvements in organizational variables (level 4) to occur. That is, a training course may or may not produce a measurable improvement in an organizational variable (level 4) but can only do so if the intended improvements in job behavior (level 3) occur. This yields hypothesis 3 with its three parts:

Hypothesis 3a: An improvement in managing behavior (level 3) will be a necessary but not sufficient condition for an improvement in subordinate satisfaction with supervisory style (level 4) to occur (but see hypotheses 3b and 3c below).

Hypothesis 3b: An improvement in communicating behavior (level 3) will be a necessary but not sufficient

condition for an improvement in subordinate satisfaction with supervisory style (level 4) to occur (but see hypotheses 3a and 3c).

Hypothesis 3c: An improvement in job training behavior (level 3) will be a necessary but not sufficient condition for an improvement in subordinate satisfaction with supervisory style (level 4) to occur (but see hypotheses 3a and 3b).

Instruments

For the experimental subjects, data were gathered on the improvement in reactions toward the course (level 1), the improvement in learning (level 2), the improvements in the three job behaviors (managing, communicating, and job training--all level 3), the improvement in subordinate satisfaction with supervisory style (level 4), and the influence of variables in the organizational environment on the transfer of learning (level 3). Except for the improvement in reactions toward the course (level 1) and the influence of variables in the organizational environment on the transfer of learning (level 3)--it was not logically possible to collect these data--the same data were gathered on the control subjects.

Exhibit I of the Appendix shows samples of the instruments administered to the experimental and control subjects. Each of these instruments is described below.

Improvement in reactions toward the course (level 1). Data regarding experimental subjects' pretraining and

posttraining reactions toward the course (level 1) were gathered on the seven-point Likert scales shown in Exhibit I of the Appendix. These pretraining and posttraining scales were found to be highly reliable, with split-half reliabilities of .83 and .88, respectively. The improvement in reactions was computed as the difference between the sums of these two scales.

Improvement in learning (level 2). Data regarding knowledge of supervision were gathered by means of a 36-item multiple-choice test (Exhibit I, Appendix). Eighteen of the items were taken from a test which had been used at the end of prior Management II courses, but that test did not thoroughly sample the topic areas covered in the course. To remedy this defect, I wrote 18 new questions that were approved by one of the course instructors for inclusion in the test. The improvement in learning (level 2) was computed as the difference between the pretraining and posttraining scores on this test.

Improvements in the three job behaviors (level 3). The behavioral objectives of Management II focused on three job behaviors: managing, communicating, and job training--all level 3. Regarding the first of these, the objective was for the experimental subject to increase the percentage of time spent on managing as opposed to operating activities. The percentage of time spent managing was estimated by the experimental subject and specified on the pretraining and posttraining questionnaires shown in Exhibit I of the

Appendix. The improvement in managing behavior (level 3) was computed as the difference between the pretraining and posttraining estimates.

Regarding the improvement in communicating behavior (level 3), the objective of Management II was for the experimental subject to increase his use of the principles of effective communication taught in the course. Pretraining and posttraining data on communicating behavior were collected on seven-point Likert scales (Exhibit I, Appendix), the items of which reflected the material presented in the course. These scales were found to be highly reliable, with split-half reliabilities of .73 and .82, respectively. The improvement in communicating behavior (level 3) was computed as the difference between the sums of the pretraining and posttraining scales.

Regarding the improvement in job training behavior (level 3), the objective of Management II was for the experimental subject to increase his use of the six steps of job training taught in the course. Pretraining and posttraining data on job training behavior were collected on seven-point Likert scales (Exhibit I, Appendix), the items of which reflected the six steps of job training. These scales were found to be highly reliable, with split-half reliabilities of .97 and .98, respectively. The improvement in job training behavior (level 3) was computed as the difference between the sums of the pretraining and posttraining scales.

The instruments used to gather the data regarding the improvements in reactions (level 1), managing behavior,

communicating behavior, and job training behavior (all level 3) obviously were based on self-ratings. Wexley and Yukl (1977) report that such ratings have certain weaknesses: "Self-appraisals are typically inflated compared with those of superiors. Consequently, there is every reason to believe that they are subject to self-interests which seriously lower their validity" (p. 223).

To estimate the validity of the experimental or control subject's responses regarding his managing behavior (level 3), his superior completed the pretraining and posttraining questionnaires shown in the Appendix, Exhibit II; these questionnaires are essentially equivalent to the questionnaires on managing behavior completed by the experimental or control subject. Table III-1 shows that there were no significant differences between the responses of the experimental subjects and the responses of their superiors.

Table III-1

Comparison of Experimental Subjects' Responses with Their Superiors' Responses to the Questionnaires on Managing Behavior

	Response from Experimental Subject	Response from Superior	Difference
Pretraining	57.6	58.5	.9
Posttraining	63.9	61.9	-2.0

Table III-2 shows that, on the pretraining questionnaires, the estimates of the control subjects did not differ significantly from the estimates of their superiors. On the

posttraining questionnaires, however, the estimates of the control subjects were significantly lower than the estimates of their superiors ($p < .05$). There is evidence that superiors and subordinates do not agree on how the latter spends his job time (O'Reilly, 1975). The superior's estimate is likely to be the less accurate since he probably cannot observe 100 percent of the subordinate's job time. Therefore, I decided to use the estimates of the control subjects.

Table III-2

Comparison of Control Subjects' Responses with Their Superiors' Responses to the Questionnaire on Managing Behavior

	Response from Control Subject	Response from Superior	Difference
Pretraining	63.8%	62.6%	-1.2%
Posttraining	60.0	72.4	12.4*

* $p < .05$

To estimate the validity of the experimental or control subject's responses regarding his communicating behavior and his job training behavior (both level 3), three of his subordinates completed the pretraining and posttraining scales shown in the Appendix, Exhibit III; these are essentially equivalent to those completed by the experimental or control subject.

Regarding the communicating behavior of the experimental or control subjects, the average of the three subordinates' responses was compared with the experimental or control subject's responses to estimate the validity of the latter.

Table III-3 shows that the comparison of experimental subjects' responses with those of their subordinates yielded no significant differences.

Table III-3

Comparison of Experimental Subjects' Responses with Their Subordinates' Responses to the Scales on Communicating Behavior

	Response from Experimental Subject	Average Response from Subordinates	Difference
Pretraining	39.0	35.9	-3.1
Posttraining	40.2	38.0	-2.2

Table III-4 shows that the comparison of control subjects' responses with those of their subordinates also yielded no significant differences.

Table III-4

Comparison of Control Subjects' Responses with Their Subordinates' Responses to the Scales on Communicating Behavior

	Response from Control Subject	Average Response from Subordinates	Difference
Pretraining	38.7	38.3	-0.4
Posttraining	39.2	36.8	-2.4

However, such a comparison was not possible for the job training behavior of the experimental or control subjects. Comments on the job training scales returned by subordinates showed that they did not understand what job training was. Therefore, the job training data from the subordinates were

not included in the analysis.

Improvement in subordinate satisfaction with supervisory style. The organizational variable (level 4) that was measured in this study was the satisfaction of the experimental subject's subordinates with the supervisory style of the experimental subject. The supervisory scale of the Job Description Index was used for this purpose. Vroom (1964) said this may be the most carefully constructed measure of job satisfaction.

This instrument was administered to three of the experimental subject's subordinates before and after training. It was assumed that obtaining three measures would yield a more valid response than having the rating of only one subordinate. The improvement in subordinate satisfaction with supervisory style (level 4) was computed as the difference between the sum of the three pretraining scores and the sum of the three posttraining scores.

Influence of variables in the organizational environment on the transfer of learning (level 3). The influence of variables in the organizational environment on the transfer of learning (level 3) was measured by means of the open-ended questionnaire shown in Exhibit I of the Appendix. This instrument focused on the three improvements in job behavior (level 3) intended by Management II. Open-ended questions were used because specific variables that might affect the transfer of learning (level 3) could not be identified a priori. It was expected that content analysis of the experimental subjects' responses would help to identify these variables.

Unfortunately, only 30 percent of the experimental subjects responded to this questionnaire in a meaningful way, so these data had to be discarded.

Procedure

Figure III-1 summarizes the instruments used in the current study. It also shows when and to whom each instrument was administered--the topic of this section of this chapter. The procedures for collecting data from experimental subjects, control subjects, their superiors, and their subordinates are presented in that order.

Procedure for experimental subjects. Figure III-1 shows that the scales on reactions toward the course (level 1) were administered to experimental subjects at the beginning of the first day of instruction and at the end of the last day. The test of supervisory knowledge (level 2) was administered to experimental subjects at these same times, but it was also administered to them two months after the course. There were two reasons for administering the third measure: (1) The overriding reason was to provide that both the experimental and the control subjects would complete a posttraining knowledge test at about the same time (I was unable to measure the posttraining knowledge of the control subjects on the last day of the training course since they could not be present at the course); (2) the administration of the knowledge tests to experimental subjects two months after the course also provided for a measure of their knowledge retention.

The experimental subjects completed the questionnaires

Instrument	When Administered	Respondents: Experimental Subject	Control Subject	Superior	Subordinates
Scale on Reaction Toward Course (Level 1)	Before Course	X			
	After Course	X			
Test of Supervisory Knowledge (Level 2)	Before Course	X	X		
	After Course	X			
	Two mos. after Course	X	X		
Questionnaire on Managing Behavior (Level 3)	Before Course	X	X	X	
	Two mos. after Course	X	X	X	
Scale on Communicating Behavior (Level 3)	Before Course	X	X		X
	Two mos. after Course	X	X		X
Scale on Job Training Behavior (Level 3)	Before Course	X	X		X
	Two mos. after Course	X	X		X
Supervisory Scale of JDI (Level 4)	Before Course				X
	Two mos. after Course				X

Figure III-1

Instruments, Respondents, and Time of Administration

on managing behavior (level 3) and the scales on communicating behavior and job training behavior (both level 3) on the first day of the course and again two months after the course. It was estimated that two months should have been sufficient time for improvements in job behavior (level 3) to manifest themselves. The experimental subjects completed the open-ended questionnaire regarding the influence of variables in the organizational environment on the transfer of learning (level 3) two months after the course.

Procedure for control subjects. All instruments administered to the control subjects were administered to them at their place of work just before and two months after the training course. These instruments included the test of supervisory knowledge (level 2), the questionnaires on managing behavior (level 3), and the scales on communicating behavior and job training behavior (level 3).

Procedure for superiors. Just before and two months after the course, the superior of each experimental subject and each control subject completed the questionnaire on managing behavior (level 3) to estimate the percentage of time that each experimental or control subject spent managing as opposed to operating.

Procedure for subordinates. Three subordinates of each experimental subject and each control subject completed the scales on communicating behavior and job training behavior (level 3) and the supervisory scale of the Job Description Index (level 4) just before and two months after the course.

SUMMARY

A pretest-posttest-control group design, which largely but not completely controlled certain threats to internal and external validity, was used in the current study. The experimental group, consisting of 75 first-level supervisors randomly selected from the population of first-level supervisors hitherto not having had the Management II course, were exposed to an experimental treatment--a five-day general supervisory training course called Management II. The control group included 21 randomly selected first-level supervisors who also had not received Management II, and who were not assigned to the current course offering.

An abbreviated version of the Hamblin model was tested in the current study. The variables measured included the improvement in reactions (level 1), the improvement in learning (level 2), the improvements in three job behaviors (managing, communicating, and job training--all level 3), and the improvement in an organizational variable (subordinate satisfaction with supervisory style--level 4).

For the experimental group, data were gathered on all these variables. For the control group, data were gathered on all but the improvement in reactions toward the course (level 1). Data regarding the managing behavior (level 3) of each experimental and control subject were gathered from the superior of each; data regarding the communicating behavior and job training behavior (both level 3) of each experimental and control subject were gathered from the

subordinates of each. These data from superiors and subordinates were used to estimate the validity of the responses from the experimental and control subjects.

Chapter IV

Results

INTRODUCTION (SECTION 1)

This chapter presents the data gathered in the study. The basic data are obviously the means and standard deviations for each of the hierarchical conditions for both the experimental and control groups. The data of central concern, however, are not so much the means and standard deviations as the measured improvements between "levels" within the experimental group as compared with the measured improvements between "levels" within the control group. These measured improvements are, by way of reminder, those for reactions toward the course (level 1) and learning (level 2), learning (level 2) and each of the three level 3 job behaviors (managing, communicating, and job training), and each of the three job behaviors as above (level 3) and subordinate satisfaction with supervisory style (level 4). Except for measuring the improvement in reactions toward the course (level 1), impossible in the case of the controls since they didn't take the course and so couldn't react to it, the set of measured improvements was the same for experimentals and controls.

For the experimental group, the measured improvements between designated levels must be significantly favorable to

provide support for the hypotheses. Obviously, non-significant measured improvements between the same levels were expected for the control group.

The following is a summary of the hypotheses of the study:

Hypothesis 1: An improvement in reactions (level 1) will be a necessary but not sufficient condition for an improvement in learning (level 2) to occur.

Hypothesis 2a: An improvement in learning (level 2) will be a necessary but not sufficient condition for an improvement in managing behavior (level 3) to occur.

Hypothesis 2b: An improvement in learning (level 2) will be a necessary but not sufficient condition for an improvement in communicating behavior (level 3) to occur.

Hypothesis 2c: An improvement in learning (level 2) will be a necessary but not sufficient condition for an improvement in job training behavior (level 3) to occur.

Hypothesis 3a: An improvement in managing behavior (level 3) will be a necessary but not sufficient condition for an improvement in subordinate satisfaction with supervisory style (level 4) to occur (but see hypotheses 3b and 3c below).

Hypothesis 3b: An improvement in communicating behavior (level 3) will be a necessary but not sufficient condition for an improvement in subordinate

satisfaction with supervisory style (level 4) to occur (but see hypotheses 3a and 3c).

Hypothesis 3c: An improvement in job training behavior (level 3) will be a necessary but not sufficient condition for an improvement in subordinate satisfaction with supervisory style (level 4) to occur (but see hypotheses 3a and 3b).

Before the data with reference to the hypotheses are presented, it is necessary to consider data regarding several preliminary conditions that could have a bearing on the testing of the hypotheses. First, the effects of experimental mortality are reported; this analysis shows the extent to which the loss of 20 subjects from the experimental group may have changed the composition of that group. Second comes a comparison of the experimental group and the control group on the pretraining dependent variables[3], the posttraining dependent variables, and the measured improvements for each of the dependent variables; these results show the extent to which the training treatment had an effect on the experimental group. Finally are presented the results of testing the hierarchy for both the experimental group and the control group, or, in other words, testing the hypotheses. The same data have been processed with three different statistical procedures in an effort to obtain confirmation of findings

---

[3] The dependent variables were the reactions toward the course (level 1), learning (level 2), three job behaviors -- managing, communicating, and job training (level 3), and subordinate satisfaction with supervisory style (level 4).

across procedures.

At this point, the reader might wonder why the results for both the experimental group and the control group were used to test the hierarchy of training effects. Consider the statement of the Hamblin model initially presented in Chapter I. Hamblin suggested that the five levels of training effects formed a hierarchy, such that:

--An improvement in reactions (level 1) will be a necessary but not sufficient condition for an improvement in learning (level 2) to occur.

--An improvement in learning (level 2) will be a necessary but not sufficient condition for improvements in job behavior (level 3) to occur.

--Improvements in job behavior (level 3) will be a necessary but not sufficient condition for improvements in organizational variables (level 4) to occur.

The experimental group of the current study was exposed to a training treatment (Management II). Hopefully they would experience an improvement in reactions toward the course (level 1). In terms of the Hamblin model, the improvement in reactions (level 1) was expected to lead to an improvement in learning; the improvement in learning (level 2) was expected to lead to improvements in three job behaviors (level 3--managing, communicating, and job training); and the three job behavior improvements (level 3) were expected to lead to an improvement in an organizational

variable (level 4--subordinate satisfaction with supervisory style). The results for the experimental group were used to test for these causal linkages.

The control group of the current study was not exposed to the training treatment, and their reactions toward the training course (level 1), as I have already explained, could not be measured. Since they were not exposed to Management II, it was expected that they would not achieve improvements in learning (level 2). Since the Hamblin model says an improvement in learning (level 2) is necessary but not sufficient for the intended job behavior improvements (level 3) to occur, the three job behavior improvements (level 3) were not expected for the control group. Since the Hamblin model says the intended improvements in job behavior (level 3) are necessary but not sufficient for the intended improvements in organizational variables (level 4) to occur, an improvement in the organizational variable (level 4--subordinate satisfaction with supervisory style) was not expected for the control group. The results for the control group were used to test for these linkages.

The effects of experimental mortality are reported next. This analysis shows the extent to which the loss of 20 subjects from the experimental group may have changed the composition of that group.

THE EFFECTS OF EXPERIMENTAL MORTALITY (SECTION 2)

Seventy-five experimental subjects were originally assigned to the courses which were evaluated in the current

study. Five of the subjects failed to complete certain pretraining instruments, so their data were discarded. For the remaining 70 experimental subjects on whom complete pretraining data were gathered, complete posttraining data were received for 50; although the other 20 subjects completed the training course, they failed to provide posttraining data (i.e., they withdrew from the study).

Table IV-1 shows that the pretraining reactions (level 1) of the 20 subjects who withdrew from the study were significantly less favorable than the 50 subjects who completed the study ($p < .05$). However, these two groups did not differ significantly on their levels of pretraining learning (level 2).

Table IV-1

Comparison of Mean Scores for Subjects Completing Study Versus Subjects Withdrawing from Study on Measures of Pretraining Reactions Toward Course (Level 1) and Pretraining Knowledge of Supervision (Level 2)

	Reactions Toward Course (Level 1)	Knowledge of Supervision (Level 2)
Subjects Who Withdrew from Study (N=20)	72.5	23.6
Subjects Who Completed Study (N=50)	78.5	24.1
Differences	6.0	0.5
t's	2.10*	0.43

* $p < .05$

The score for pretraining reactions (level 1) is a measure of the subject's expectations about the course. Therefore, the 20 subjects who withdrew from the study had significantly less favorable expectations about the course than the 50 subjects who completed the study. How this result affected the study could not be determined.

There was no mortality in the control group.

COMPARISON OF THE EXPERIMENTAL GROUP AND THE CONTROL GROUP ON PRETRAINING DEPENDENT VARIABLES, POSTTRAINING DEPENDENT VARIABLES, AND MEASURED IMPROVEMENT IN EACH DEPENDENT VARIABLE (SECTION 3)

This section presents a comparison of the experimental group and the control group on each of the following measures: the pretraining dependent variables, the posttraining dependent variables, and the measured improvements in each dependent variable. These results show the extent to which the training treatment had an effect on the experimental group.

Comparison of the Experimental Group and the Control Group on the Pretraining Dependent Variables

Typically one may assume that an experimental group and a control group are equivalent when the subjects for each have been randomly selected from a target population, as in the present study. To confirm this equivalence here, the pretraining data for each were compared. The results of this comparison are shown in Table IV-2. None of the differences between the two groups was statistically significant; therefore, it was assumed that the experimental group and the

Table IV-2

Comparison of Experimental Group and Control Group on Pretraining Dependent Variables

	Knowledge of Supervision (Level 2)	Job Behaviors (Level 3)			Subordinate Satisfaction with Supervisory Style (Level 4)
		Managing	Communicating	Job Training	
Experimental Group (N=50)	24.1	57.7	38.6	35.0	69.2
Control Group (N=21)	25.2	59.8	38.9	33.9	71.1
Differences	-1.1	-2.1	-0.3	1.1	-1.9
t's	1.15	0.33	0.19	0.71	0.51

* $p < .05$

control group before the training treatment were essentially equivalent.

Comparison of the Experimental Group and the Control Group on the Posttraining Dependent Variables

Table IV-3 shows a comparison of the experimental group and the control group on the posttraining dependent variables. The experimental group scored significantly higher than the control group on the test of knowledge of supervision (level 2) ($p < .01$) and on the scale of job training behavior (level 3) ($p < .05$) but not on the remaining dependent variables.

Comparison of the Experimental Group and the Control Group on the Measured Improvements in Each Dependent Variable

Table IV-4 first shows the measured improvement for each dependent variable for the experimental group, then shows the measured improvement for each dependent variable for the control group, and finally permits a comparison of the measured improvements for the experimental and control groups. For the experimental group, significant measured improvements were found for reactions (level 1) ($p < .001$), learning (level 2) ($p < .001$), managing behavior (level 3) ($p < .05$), and job training behavior (level 3) ($p < .05$). The measured improvement for communicating behavior (level 3) reached the .10 level of significance ($p < .07$). For the experimental group, all of these measured improvements were in the predicted direction. Table IV-4 also shows that there were no significant measured improvements for any of the dependent variables for the control group.

Table IV-3

Comparison of Experimental Group and Control Group on Posttraining Dependent Variables

	Knowledge of Supervision (Level 2)	Job Behaviors (Level 3)			Subordinate Satisfaction with Supervisory Style (Level 4)
		Managing	Communicating	Job Training	
Experimental Group (N=50)	27.2	63.8	40.0	36.3	71.8
Control Group (N=21)	24.5	56.4	39.9	32.4	69.7
Differences	2.7	7.4	0.1	3.9	2.1
t's	2.77**	1.02	0.04	2.31*	0.54

* $p < .05$

** $p < .01$

Table IV-4

Measured Improvements for Each Dependent Variable for Experimental Group and Control Group

	Reactions (Level 1)	Learning (Level 2)	Job Behaviors (Level 3			Subordinate Satisfaction with Supervisory Style (Level 4)
			Managing	Communicating	Job Training	
Experimental Group (N=50)						
Pretraining	78.5	24.1	57.7	38.6	35.0	69.2
Posttraining	88.1	27.2	63.8	40.0	36.3	71.8
Differences	9.6	3.1	6.1	1.4	1.3	2.6
t's	4.54***	5.71***	2.20*	1.86	2.14*	1.53
Control Group (N=21)						
Pretraining	--[a]	25.2	59.8	38.9	33.9	71.1
Posttraining	--	24.5	56.4	39.9	32.4	69.7
Differences	--	- 0.7	- 3.4	1.0	- 1.5	- 1.4
t's	--	1.39	1.31	1.12	1.25	0.47

[a] Reactions data were not collected for the control group.

* $p < .05$

** $p < .01$

*** $p < .001$

Summary

The comparison of the pretraining data for the experimental group and the control group showed that the two groups were essentially equivalent before the training treatment was administered to the experimental group. While no significant measured improvements were found for the control group, four significant measured improvements were found for the experimental group; the four significant measured improvements were those for reactions (level 1), learning (level 2), managing behavior (level 3), and job training behavior (level 3). However, the comparison of the posttraining data for the experimental group and the control group showed that the former had scored significantly higher than the latter only on two dependent variables: learning (level 2) and job training behavior (level 3).

THE TESTS OF THE ASSUMPTION OF A HIERARCHY OF TRAINING EFFECTS (SECTION 4)

This section presents the tests of the hierarchy of training effects. First, the intercorrelation matrices of measured improvements for both the experimental group and the control group are presented. Second, the multiple regression analyses of measured improvements for the experimental group and the control group are presented. Third, the path analyses of measured improvements for the experimental group and the control group are presented. In each of these three parts of this section the data for the experimental group are reported before the data for the control group. The section closes with a summary of the results regarding the

tests of the assumption of a hierarchy of training effects.

Intercorrelation Matrices for the Measured Improvements for the Experimental Group and the Control Group

This section presents the intercorrelation matrices for the measured improvements for the experimental group and the control group. The intercorrelation matrix for the experimental group is presented first.

Intercorrelation matrix for the experimental group. The intercorrelation matrix for the measured improvements for the experimental group (Table IV-5) showed correlations significantly different from zero between the measured improvement in reactions (level 1) and the measured improvement in learning (level 2) ($p < .01$), between the measured improvement in learning (level 2) and the measured improvement in communicating behavior (level 3) ($p < .01$), and between the measured improvement in communicating behavior and the measured improvement in job training behavior (both level 3) ($p < .01$).

The first two of these correlations were predicted by the hypotheses: The measured improvement in reactions (level 1) was expected to be linked to the measured improvement in learning (level 2), and the latter was expected to be linked to the measured improvement in communicating behavior (level 3). The correlation between the measured improvement in communicating behavior and the measured improvement in job training behavior (both level 3) is unrelated to the hypotheses of the current study.

Intercorrelation matrix for the control group. The intercorrelation matrix for the measured improvements for

Table IV-5

Intercorrelation Matrix for Measured Improvements for Experimental Group (N=50)

Measured Improvements in:	1	2	3	4	5	6
1. Reactions (Level 1)	--					
2. Learning (Level 2)	.50**[a]	--				
3. Managing Behavior (Level 3)	.03	-.02	--			
4. Communicating Behavior (Level 3)	.22	.41**	.22	--		
5. Job Training Behavior (Level 3)	.04	.01	.15	.45**	--	
6. Subordinate Satisfaction with Supervisory Style (Level 4)	.01	-.14	-.05	.11	.00	--

[a] Correlations hypothesized to be significantly different from zero are encircled.

* $p < .05$

** $p < .01$

the control group (Table IV-6) showed correlations significantly different from zero between the measured improvement in learning (level 2) and the measured improvement in communicating behavior (level 3) ($p < .01$), between the measured improvement in managing behavior (level 3) and the measured improvement in subordinate satisfaction with supervisory style (level 4) ($p < .001$), and between the measured improvement in communicating behavior and the measured improvement in job training behavior (both level 3) ($p < .05$).

Further interpretation of the correlations significantly different from zero found for both the experimental group and the control group is presented in the last section of this chapter.

Multiple Regression Analyses on the Measured Improvements for the Experimental Group and the Control Group

This section presents the multiple regression analyses on the measured improvements for the experimental group and the control group. The multiple regression analysis for the experimental group is presented first.

Multiple regression analysis for the experimental group.

The multiple regression analysis for the experimental group (Table IV-7) provided partial support for two hypotheses. The measured improvement in reactions (level 1) accounted for 25 percent of the variance in the measured improvement in learning (level 2) ($p < .01$) (Table IV-7-a). The measured improvement in learning (level 2) accounted for 17 percent of the measured improvement in communicating behavior (level 3) ($p < .05$, Table IV-7-c).

Table IV-6

Intercorrelation Matrix for Measured Improvements for Control Group (N=21)

	1	2	3	4	5	6
Measured Improvements in:						
1. Reactions[a] (Level 1)	--					
2. Learning (Level 2)	--	--				
3. Managing Behavior (Level 3)	--	.18[b]	--			
4. Communicating Behavior (Level 3)	--	.59**	.35	--		
5. Job Training Behavior (Level 3)	--	.28	.27	.45*	--	
6. Subordinate Satisfaction with Supervisory Style (Level 4)	--	.25	.69***	.38	.13	--

[a] Reactions data were not collected for the control group.

[b] Correlations hypothesized to be significantly different from zero for the experimental group are encircled.

* $p < .05$ ** $p < .01$ *** $p < .001$

Table IV-7a

Regression of Measured Improvement in Learning (Level 2) Upon Measured Improvement in Reactions (Level 1) for Experimental Group

	Simple r	F (to add)	Significance	R^2	Beta Weight
Measured Improvement in:					
Reactions (Level 1)	.50	16.0	$p < .01$	.25	.50

Table IV-7b

Regression of Measured Improvement in Managing Behavior (Level 3) Upon Measured Improvements in Learning (Level 2) and Reactions (Level 1) for Experimental Group

	Simple r	F (to add)	Significance	R^2	Beta Weight
Measured Improvement in:					
Learning (Level 2)	-.02	0.1	--	.00	-.05
Reactions (Level 1)	.03	0.1	--	.00	.05

Table IV-7c

Regression of Measured Improvement in Communicating Behavior (Level 3) Upon Measured Improvements in Learning (Level 2) and Reactions (Level 1) for Experimental Group

	Simple r	F (to add)	Significance	R^2	Beta Weight
Measured Improvements in:					
Learning (Level 2)	.41	6.9	$p < .05$	.17	.40
Reactions (Level 1)	.22	0.0	--	.17	.02

Table IV-7d

Regression of Measured Improvement in Job Training Behavior (Level 3) Upon Measured Improvements in Learning (Level 2) and Reactions (Level 1) for Experimental Group

	Simple r	F (to add)	Significance	R^2	Beta Weight
Measured Improvements in:					
Learning (Level 2)	.01	0.0	--	.00	-.02
Reactions (Level 1)	.04	0.1	--	.00	.04

Table IV-7e

Regression of Measured Improvement in Subordinate Satisfaction (Level 4) Upon Measured Improvements in Each Job Behavior (Level 3), Learning (Level 2), and Reactions (Level 1) for Experimental Group

	Simple r	F (to add)	Significance	R^2	Beta Weight
Measured Improvements in:					
Managing Behavior (Level 3)	-.06	0.1	--	.00	-.07
Learning (Level 2)	-.14	1.1	--	.02	.20
Reactions (Level 1)	.01	0.3	--	.03	.12

Table IV-7e (cont'd.)

	Simple r	F (to add)	Significance	R^2	Beta Weight
Measured Improvements in:					
Communicating Behavior (Level 3)	.11	1.2	--	.01	.20
Learning (Level 2)	-.14	1.8	--	.05	-.28
Reactions (Level 1)	.01	0.3	--	.06	.11
Measured Improvements in:					
Job Training Behavior (Level 3)	.00	0.0	--	.00	.00
Learning (Level 2)	-.14	1.0	--	.02	-.20
Reactions (Level 1)	.01	0.3	--	.03	.11

None of the other predicted linkages was found to be significant. That is, the measured improvement in learning (level 2) was not significantly linked to either the measured improvement in managing behavior (level 3) or the measured improvement in job training behavior (level 3); further, none of the measured improvements in the three job behaviors (level 3) was significantly linked to the measured improvement in subordinate satisfaction with supervisory style (level 4).

Multiple regression analysis for the control group.

The multiple regression analysis for the control group (Table IV-8) showed that the measured improvement in learning (level 2) accounted for 35 percent of the variance in the measured improvement in communicating behavior (level 3) (p .01, Table IV-8-c). The measured improvement in managing behavior (level 3) accounted for 48 percent of the variance in the measured improvement in subordinate satisfaction with supervisory style (level 4) ($p < .01$, Table IV-8-e).

Table IV-8a

Regression of Measured Improvements in Learning (Level 2) Upon Measured Improvement in Reactions (Level 1) for Control Group

	Simple r	F (to add)	Significance	R^2	Beta Weight
Measured Improvement in:					
Reactions (Level 1)[a]	--	--	--	--	--

Table IV-8b

Regression of Measured Improvement in Managing Behavior (Level 3) Upon Measured Improvements in Learning (Level 2) and Reactions (Level 1) for Control Group

	Simple r	F (to add)	Significance	R^2	Beta Weight
Measured Improvements in:					
Learning (Level 2)	.18	0.6	--	.03	.18
Reactions (Level 1)[a]	--	--	--	--	--

[a] Reactions data could not be collected from the control group.

Table IV-8c

Regression of Measured Improvement in Communicating Behavior (Level 3) Upon Measured Improvements in Learning (Level 2) and Reactions (Level 1) for Control Group

	Simple r	F (to add)	Significance	R^2	Beta Weight
Measured Improvements in:					
Learning (Level 2)	.59	10.2	$p < .01$	.35	.59
Reactions (Level 1)[a]	--	--	--	--	--

Table IV-8d

Regression of Measured Improvements in Job Training Behavior (Level 3) Upon Measured Improvements in Learning (Level 2) and Reactions (Level 1) for Control Group

	Simple r	F (to add)	Significance	R^2	Beta Weight
Measured Improvements in:					
Learning (Level 2)	.28	1.6	--	.08	.28
Reactions (Level 1)[a]	--	--	--	--	--

[a] Reactions data could not be collected from the control group.

Table IV-8e

Regression of Measured Improvement in Subordinate Satisfaction (Level 4) Upon Measured Improvements in Each Job Behavior (Level 3), Learning (Level 2), and Reactions (Level 1) for Control Group

	Simple r	F (to add)	Significance	R^2	Beta Weight
Measured Improvements in:					
Managing Behavior (Level 3)	.69	12.7	$p < .01$	.48	.69
Learning (Level 2)	.25	0.5	--	.49	.13
Reactions[a] (Level 1)	--	--	--	--	--
Measured Improvements in:					
Communicating Behavior (Level 3)	.38	1.4	--	.14	.35
Learning (Level 2)	.25	0.0	--	.14	.04
Reactions[a] (Level 1)	--	--	--	--	--
Measured Improvements in:					
Job Training Behavior (Level 3)	.13	0.1	--	.02	.07
Learning (Level 2)	.25	0.8	--	.07	.23
Reactions[a] (Level 1)	--	--	--	--	--

[a] Reactions data could not be collected from the control group.

None of the other linkages was found to be significant. That is, the measured improvement in learning (level 2) was not significantly linked to either the measured improvement in managing behavior (level 3) or the measured improvement in job training behavior (level 3); further, neither the measured improvement in communicating behavior (level 3) nor the measured improvement in job training behavior (level 3) was significantly linked to the measured improvement in subordinate satisfaction with supervisory style (level 4).

Path Analyses on the Measured Improvements for the Experimental Group and the Control Group

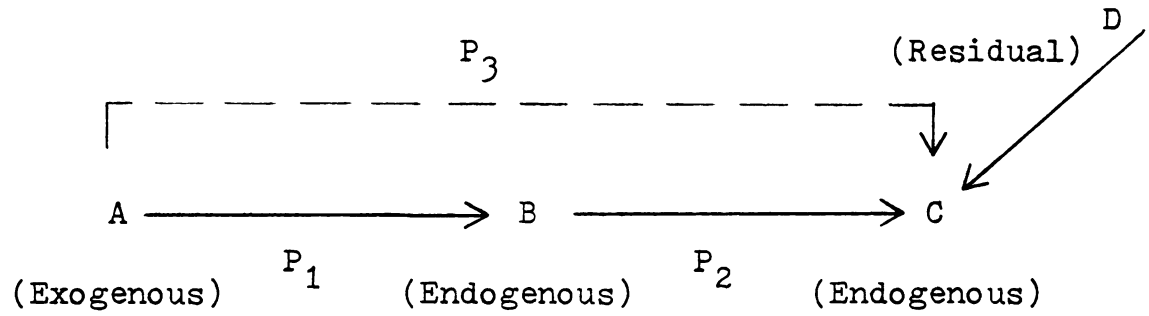
The statistical technique used to test the assumption of a hierarchy of training effects was path analysis. Following a description of the path analysis technique, the results of the path analyses for the experimental group and the control group are presented.

Path analysis. This "...is primarily a method of decomposing and interpreting linear relationships among a set of variables by assuming that (1) a (weak) causal order among these variables is known and (2) the [linkages] among these variables are causally closed" (Nie, Hull, Jenkins, Steinbrenner, and Best, 1975, p. 383). The requirement of a weak causal order is met if it is assumed or known that the causal linkage between two variables is unidirectional. The requirement for causal closure is met if it can be assumed that variables external to the model (called residuals) do not correlate with other variables.

Path analysis "...is not a method for discovering causes, but a method applied to a causal model formulated by the researcher on the basis of knowledge and theoretical considerations...In other words, path analysis is useful in testing theory rather than in generating it" (Kerlinger and Pedhazer, 1973, p. 305).

A path diagram graphically shows the assumed pattern of causal linkages among a set of variables. Three types of variables are typically included: exogenous, endogenous, and residual. Exogenous variables are those for which the variability is assumed to be determined by causes outside the model. The variation of endogenous variables is assumed to be explained by exogenous variables or other endogenous variables within the system. Residual variables are those external to the model, and they are assumed to account for any unexplained variation in the dependent variable. For example, in Figure IV-1 variable A is exogenous, variables B and C are endogenous, and variable D is residual.

A variable may have a direct effect on a second variable, or it may have an indirect effect through a third variable. For example, in Figure IV-1 variable A has a direct effect on variable B, and variable B has a direct effect on variable C. Each of these direct effects is shown with a solid arrow. Variable A also has an indirect effect on variable C through variable B; this indirect effect is shown by means of a dashed arrow from variable A through variable B to variable C.



$$B = P_1 A$$

$$C = P_2 B + P_3 A$$

Figure IV-1

An Example of Direct and Indirect Effects Within a Path Diagram

Kerlinger and Pedhazer explain that:

> Each endogenous (dependent) variable in a causal model may be represented by an equation consisting of the variables upon which it is assumed to be dependent, and a term representing residuals... For each independent variable in the equation there is a path coefficient indicating the amount of expected change in the dependent variable as a result of a unit change in the independent variable (p. 310).

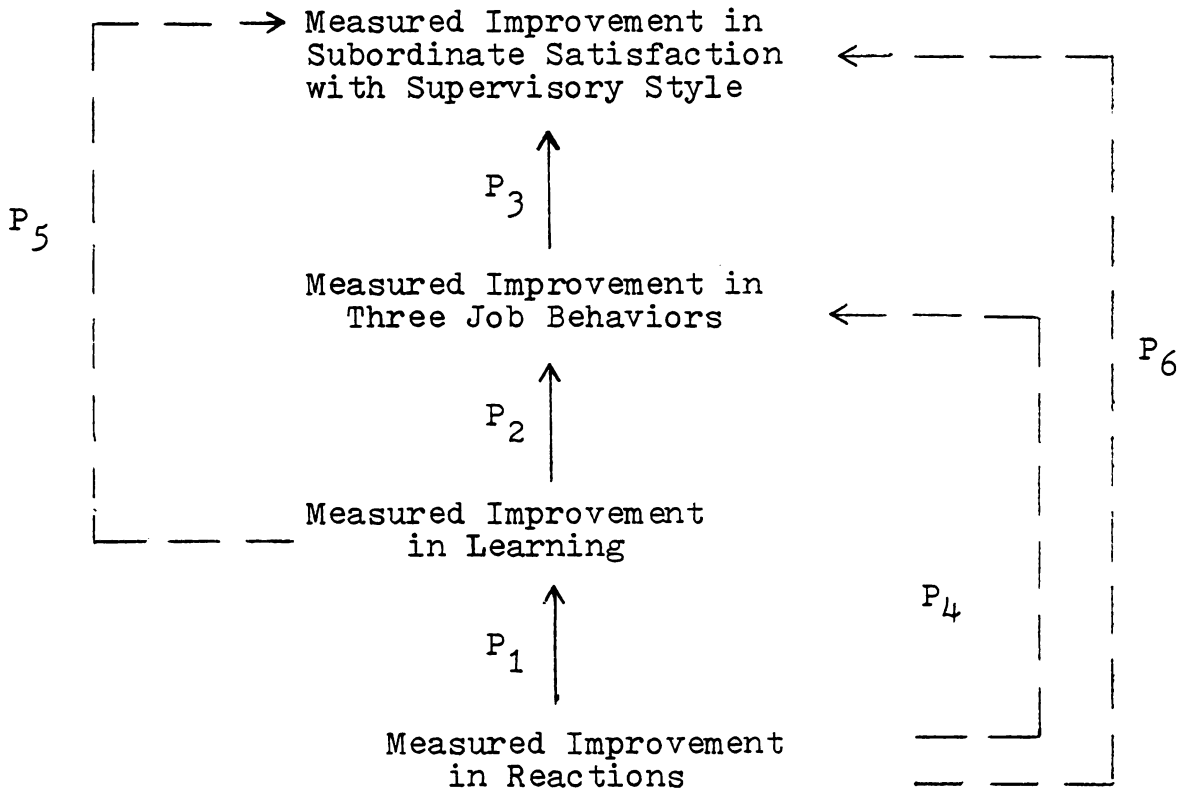
The path coefficients may be estimated by regressing each dependent variable on the independent variables in the path equations. Path coefficients are represented by the standardized beta coefficients resulting from the regression analysis. The path equations for the path diagram of

Figure IV-1 are shown below the diagram; the path coefficients are P_1, P_2, and P_3.

The path diagram and the path equations for the limited version of the Hamblin model tested in this study are shown in Figure IV-2. The improvement in reactions (level 1) is exogenous, and the other three variables are endogenous. In this path diagram, the solid arrows represent the direct effects among the variables; these are the linkages proposed by Hamblin. The dashed arrows represent the indirect effects among the variables; the Hamblin model may be assumed to predict non-significant indirect effects.

Path analysis for the experimental group. Table IV-9 shows that, for the experimental group, the path coefficient from the measured improvement in reactions (level 1) to the measured improvement in learning (level 2) was .50 ($p < .01$), and the path coefficient from the measured improvement in learning (level 2) to the measured improvement in communicating behavior (level 3) was .40 ($p < .05$). None of the other path coefficients achieved significance.

Path analysis for the control group. Table IV-10 shows that, for the control group, the path coefficient from the measured improvement in learning (level 2) to the measured improvement in communicating behavior (level 3) was .59 ($p < .01$), and the path coefficient from the measured improvement in managing behavior (level 3) to the measured improvement in subordinate satisfaction with supervisory style (level 4) was .69 ($p < .01$). None of the other path coefficients achieved significance.



Measured Improvement in Learning = P_1 Measured Improvement in Reactions

Measured Improvement in Job Behavior = P_2 Measured Improvement in Learning + P_4 Measured Improvement in Reactions

Measured Improvement in Subordinate Satisfaction with Supervisory Style = P_3 Measured Improvement in Job Behavior + P_5 Measured Improvement in Learning + P_6 Measured Improvement in Reactions

Figure IV-2

Assumed Pattern of Causal Linkages
Among the Dependent Variables

Table IV-9

Path Coefficients for Experimental Group

	Measured Improvements in:				
	Learning (Level 2)	Managing Behavior (Level 3)	Communicating Behavior (Level 3)	Job Training Behavior (Level 3)	Subordinate Satisfaction with Supervisory Style (Level 4)
Measured Improvements in:					
Reactions (Level 1)	.50**				
Learning (Level 2)		-.05	.40*	-.02	
Managing Behavior					-.07
Communicating Behavior					.20
Job Training Behavior					.00
Subordinate Satisfaction with Supervisory Style					

* $p < .05$

** $p < .01$

Table IV-10

Path Coefficients for Control Group

	Measured Improvements in:				
	Learning (Level 2)	Managing Behavior (Level 3)	Communicating Behavior (Level 3)	Job Training Behavior (Level 3)	Subordinate Satisfaction with Supervisory Style (Level 4)
Measured Improvements in:					
Reactions (Level 1)[a]	--				
Learning (Level 2)		.18	.59**	.28	
Managing Behavior					.69**
Communicating Behavior					.35
Job Training Behavior					.07
Subordinate Satisfaction with Supervisory Style					

[a] Reactions data could not be collected from the control group.

* $p < .05$

** $p < .01$

Summary

In testing the assumption of a hierarchy of training effects, the results for the experimental group provided support for two hypotheses: The measured improvement in reactions (level 1) was found to be linked to the measured improvement in learning (level 2), and the measured improvement in learning (level 2) was found to be linked to the measured improvement in communicating behavior (level 3).

The results for the control group showed that the measured improvement in learning (level 2) was found to be linked to the measured improvement in communicating behavior (level 3), and the measured improvement in managing behavior (level 3) was found to be linked to the measured improvement in subordinate satisfaction with supervisory style (level 4).

The last section of this report presents an integration of these results with the results reported in the prior section of this chapter. This integration provides an interpretation of the results from testing the assumed hierarchy of training effects.

INTEGRATION OF THE RESULTS (SECTION 5)

For the experimental group of the current study, significant measured improvements were found in reactions (level 1), learning (level 2), managing behavior (level 3), and job training behavior (level 3); the measured improvement in communicating behavior for the experimental group achieved the .10 level of significance.

Using the data of the experimental group, the test of

the assumed hierarchy of training effects showed only two out of seven linkages: one between the measured improvement in reactions (level 1) and the measured improvement in learning (level 2) and a second between the measured improvement in learning (level 2) and the measured improvement in communicating behavior (level 3). However, linkages were not found between the measured improvement in learning (level 2) and the measured improvement in managing behavior (level 3) or between the measured improvement in learning (level 2) and the measured improvement in job training behavior (level 3). That is, even though the measured improvements in learning (level 2), managing behavior (level 3), and job training behavior (level 3) were all significant, the hypothesized linkages (H2a and H2c) were not found. Obviously, the reason for this is that the degree of linkage between two variables (e.g., learning and managing behavior) depends not upon the significance of their measured improvements but upon the extent to which the two variables are correlated.

For the control group of the current study, none of the measured improvements for any of the dependent variables was significant. However, using the data of the control group, two linkages related to the hypotheses were found: The measured improvement in learning (level 2) was found to be linked to the measured improvement in communicating behavior (level 3), and the measured improvement in managing behavior (level 3) was found to be linked to the measured improvement in subordinate satisfaction with supervisory style (level 4).

Scattergrams of the data for the control group provided a different explanation for each of these two linkages. The scattergram comparing the measured improvement in managing behavior (level 3) with the measured improvement in subordinate satisfaction with supervisory style (level 4) showed that the "linkage" (i.e., high correlation) found between these two variables was due to meaningless variation in the inherent measures of the two variables.

The scattergram comparing the measured improvement in learning (level 2) with the measured improvement in communicating behavior (level 3) showed that the linkage between these two variables was in fact due to the consistency of those two measured improvements for individual subjects. That is, for individual control subjects, when the measured improvement in learning (level 2) was positive, the measured improvement in communicating behavior (level 3) was positive; similarly, when the measured improvement in learning was negative, the measured improvement in communicating behavior was negative.

For the control group, the linkage found between the measured improvement in learning and the measured improvement in communicating behavior supported hypothesis H2b. That is, even though a training treatment was not administered to the control group, many subjects in that group performed differently (better or worse) on the posttraining test of knowledge than on the pretraining test of knowledge. Those who improved their performance on the test of knowledge also improved their

communicating behavior; those who diminished their performance on the test of knowledge also experienced a detrimental effect on their communicating behavior.

In summary, the results for the experimental group supported two hypotheses: The measured improvement in reactions (level 1) was found to be linked to the measured improvement in learning (level 2), and the measured improvement in learning (level 2) was found to be linked to the measured improvement in communicating behavior (level 3). For the control group, the measured improvement in learning (level 2) was found to be linked to the measured improvement in communicating behavior (level 3).

Chapter V

Conclusions and Implications

THE HAMBLIN HIERARCHY IN PERSPECTIVE

The evaluation of a training course usually requires the examination of several criteria of success. The world of training, in addressing itself to the problems of the complexity of criteria applicable to evaluating training courses, in recent years has tended to accept a hierarchical model. The hierarchy is seen as being composed of several levels of criteria: favorable training effects at the lowest criterion level are seen to be necessary but not sufficient conditions for favorable training effects at the next highest criterion level, and so on to the highest criterion level.

Both Kirkpatrick (1967) and Hamblin (1974) have proposed hierarchical models of training evaluation; the latter's proposal is an extension of the former's. The usual statement of the Hamblin/Kirkpatrick hierarchy might lead one to conclude that to evaluate a training course would require the measurement of only one general training effect at each level of the hierarchy.

For example, one goal of the current study was to evaluate a training course in Michigan state government called Management II. The client organization wanted data collected relevant to each level of the Hamblin hierarchy. Therefore,

I was able to use the data collected for the client to determine to what extent the Hamblin hierarchy could be validated. This evaluation of Management II followed a commonly encountered statement of the Hamblin hierarchy:

--An improvement in reactions (level 1) will be a necessary but not sufficient condition for an improvement in learning (level 2) to occur.

--An improvement in learning (level 2) will be a necessary but not sufficient condition for improvements in job behavior (level 3) to occur.

--Improvements in job behavior (level 3) will be a necessary but not sufficient condition for improvements in organizational variables (level 4) to occur.

It will be observed that, in each of the above statements, nothing is said regarding the generality or specificity of the training effects to be measured at the levels of improvement in reactions (level 1), improvement in learning (level 2), improvements in job behavior (level 3), or improvements in organizational variables (level 4). Further, my review of the literature on training evaluation found no studies that had differentiated between general and specific training effects. I assumed that the above statements regarding the Hamblin hierarchy referred to general training effects because I was not sensitized to the importance of differentiating between general and specific training effects. Therefore, the current study measured general (rather than specific) training effects

at the reactions level (level 1) and general (rather than specific) training effects at the learning level (level 2).

However, it did not occur to me to measure general training effects at level 3 (improvements in job behaviors) and level 4 (improvements in organizational variables). The behavioral objectives of Management II focused on specific rather than general training effects at the level of improvements in job behavior (level 3). Further, general training effects at the level of improvements in organizational variables (level 4) could not be used because some simply were not relevant to a government agency (e.g., profits) and because the client organization wanted to focus on organizational variables (level 4) that would prove the effectiveness of the course in the short run; the latter condition eliminated certain general organizational variables (e.g., turnover) that tend to improve only in the long run.

Getting back to level 3 again, I relied upon the behavioral objectives of Management II, which identified three specific job behaviors (managing, communicating, and job training--all level 3) on which I could focus. Then on level 4 I also focused on two specific organizational variables (subordinate satisfaction with supervisory style and subordinate absenteeism--both level 4), the improvements in which might be expected to occur during the life of the study.

That is, the way it turned out, I could get general training effects at the two lowest levels of the Hamblin hierarchy (improvements in reactions and learning) and specific

training effects at the next two highest levels (improvements in each of the three job behaviors and subordinate satisfaction with supervisory style).

In retrospect, Hamblin probably intended a more pluralistic view of his hierarchy, with several operations required to measure the many training effects at each level of the hierarchy. Consider the improvements in reactions toward a course (level 1). These may be general, focusing on the course as a whole, or they may be specific, focusing on the different aspects of a course (e.g., each instructor, each topic presented). When a course is evaluated at the reactions level, either or both the general and the specific improvements in reactions toward the course could be measured.

Levels beyond reactions (level 1), that is, improvements in learning (level 2), improvements in job behavior (level 3), and improvements in organizational variables (level 4), also could include both general and specific training effects, all of which then could be measured independently. For example, in addition to general learning (level 2), a course on supervision might focus on improvements in such specific types of learning (level 2) as planning, organizing, motivating, disciplining, and so forth. Regarding improvements in job behaviors (level 3), the same course, in addition to general training effects, might focus on obtaining on-the-job improvements in such specific job behaviors as planning behavior, organizing behavior, motivating behavior, disciplining behavior, and so forth for each one of the other

topic areas covered in the course. It is logical to expect that, if the specific job behavior taught in a training course (level 2) is planning behavior, then on-the-job planning behavior would be the specific job behavior (level 3) in which an improvement would be expected. Finally, in addition to general training effects at the level of organizational variables (level 4), many specific training effects could be measured at this level, depending on the specific organizational goals of the course (e.g., improved employee morale, better work quality, reduced absenteeism). Obviously, an assessment of the training needs before the design of the course would show which specific job behavior improvements (level 3) would be necessary for which specific organizational variable improvements (level 4).

Also in retrospect, it is important to note that an unplanned partial multicontent-multimethod approach (Campbell and Fiske, 1959) was used to test the Hamblin hierarchy in the current study. Figure V-1 shows the aspects of the multicontent-multimethod approach involved in the current study.

The content dimension of the matrix of Figure V-1 shows that a training effect may fall anywhere along a continuum from very general content to very specific content. It could have been Hamblin's original assumption that to get specific training effects at a higher level of the hierarchy (e.g., improvement in planning behavior--level 3) it was necessary to get content-related specific training effects at the next

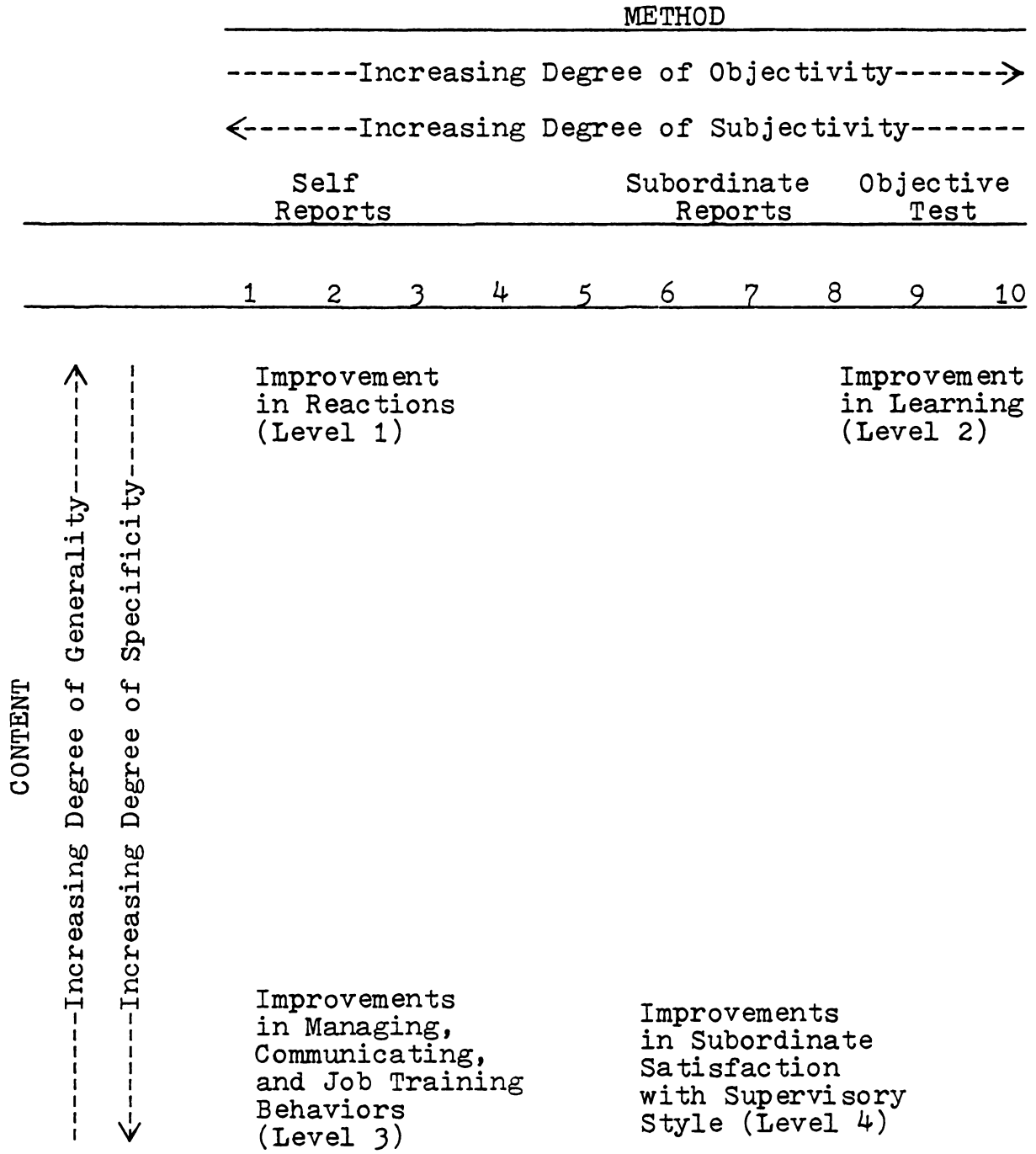


Figure V-1

The Multicontent-Multimethod Approach
Used to Test the Hamblin Hierarchy

lower level of the hierarchy (e.g., improvement in learning of planning behavior--level 2). However, as was stated previously, a review of the literature on training evaluation did not uncover a specific statement to this effect by Hamblin or any other researcher.

The method dimension of the matrix of Figure V-1 shows that the method used to measure the training effects at each level of the hierarchy may fall anywhere along a continuum from very subjective (a schematic value of 1 on the continuum) to very objective (a schematic value of 10 on the continuum). For example, the matrix of Figure V-1 shows that the improvement in reactions (level 1) was measured by means of self-reports, which tend to be quite subjective (perhaps achieving a value of 2 on the continuum). An objective test of knowledge (perhaps achieving a value of 9 on the continuum) was used to measure the improvement in learning (level 2). The improvements in each of the three specific job behaviors (managing, communicating, and job training--all level 3) were also measured by means of self-reports: the improvement in managing behavior (level 3) was measured by means of a questionnaire on which the experimental subject estimated the percentage of his job time spent managing instead of operating; the improvements in communicating behavior and job training behavior (both level 3) were measured by means of self-reports on Likert scales with seven items (communicating behavior) and six items (job training behavior). Finally, the improvement in subordinate

satisfaction with supervisory style (level 4) was measured by means of the supervisory scale of the Job Description Index, to which three of each of the experimental subject's subordinates responded (i.e., subordinate-reports were used). Because responses were obtained from three subordinates of each experimental subject, this method might tend to fall on the continuum at a value of about 7.

Also important regarding the multicontent-multimethod matrix, but not shown in Figure V-1, is the following continuum. One might expect that, when specific training effects at one level of the hierarchy (e.g., improvement in learning of specific planning behavior--level 2) are content-related to specific training effects at the next highest level (e.g., improvement in specific on-the-job planning behavior--level 3), the former training effects would be a necessary but not sufficient condition for the latter training effects to occur. However, general training effects at one level of the hierarchy are probably not content-related to either general training effects or specific training effects at another level of the hierarchy.

The multicontent-multimethod matrix of Figure V-1, including the above condition, has implications for the results of this study. How the results of the current study fit into the multicontent-multimethod matrix of Figure V-1 is the subject of the next section of this chapter.

HOW THE RESULTS FIT INTO THE MULTICONTENT-MULTIMETHOD MATRIX

This section of the chapter shows how the results of the current study fit into the multicontent-multimethod matrix of Figure V-1.

An improvement in general reactions (level 1)--measured by self-reports, a rather subjective method of measurement--was found to be a necessary but not sufficient condition for an improvement in general learning (level 2)--measured by a multiple-choice test of knowledge gained, a rather objective method of measurement. An improvement in general learning (level 2)--measured by the knowledge test, a rather objective method of measurement--was found to be a necessary but not sufficient condition for an improvement in one specific job behavior (communicating behavior--level 3)--measured by self-reports, a rather subjective method of measurement.

That is, an improvement in general learning (level 2) apparently could take place when preceded by an improvement in general reactions (level 1), and an improvement in specific communicating behavior (level 3) could take place when preceded by an improvement in general learning (level 2). Apparently, the improvement in general learning included enough communication learning content to produce an improvement in specific communicating behavior (level 3).

Contrary to the assumptions of the hierarchy, however, the improvement in managing behavior and job training behavior (both level 3)--two other specific job behaviors that were measured by self-reports--were not found to be

correlated to the improvement in general learning (level 2)--measured by an objective test. In other words, the improvement in general learning (level 2) apparently did not include enough specific managing and job training content to produce improvements in the specific job behaviors of managing behavior and job training behavior (both level 3).[4] Perhaps the knowledge test did not include certain specific content on managing and job training behavior that was in fact covered in the course.

The results of the current study also showed that the improvements in the three specific job behaviors (managing, communicating, and job training--all level 3, and all measured by self-reports) did not produce an improvement in a specific organizational variable (subordinate satisfaction with supervisory style--level 4--measured by subordinate reports, a relatively objective method of measurement. This result may not have been contrary to the assumptions of the hierarchy: the Hamblin model assumes that certain organizational events, external to the hierarchy, may influence the interfaces between the levels of the hierarchy. Such organizational events--several of which were uncovered by the empirical evidence reviewed in Chapter II--may have prevented the specific job behavior improvements (level 3) from producing the specific improvement in subordinate satisfaction with supervisory style (level 4). That is, the former

---

[4] A look at the test showed that the questions regarding communicating behavior were more specific to what was taught in the course than were the questions regarding managing and job training behaviors.

improvements (level 3) may have been necessary but not sufficient for the latter improvement (level 4).

When one looks at the results of the current study in light of the multicontent-multimethod matrix of Figure V-1, two problems are obvious. The first problem relates to the content dimension of the matrix, and the second problem relates to the method dimension.

1. The current study focused on general training effects at the two lowest levels of the hierarchy (improvements in reactions and learning) and specific training effects at the next two highest levels (improvements in each of the three job behaviors and subordinate satisfaction with supervisory style). Therefore, as shown in Figure V-1, one correlational test of the hierarchy focused on two general training effects (improvement in general reactions [level 1] and improvement in general learning [level 2]), three other correlational tests of the hierarchy focused on one general training effect (improvement in general learning [level 2]) and each of three specific training effects (improvements in the three job behaviors [level 3]), and the last three correlational tests of the hierarchy focused only on specific training effects (improvements in the three specific job behaviors [level 3] and improvement in subordinate satisfaction with supervisory style [level 4]). The problem here may have been that the assessment of training needs failed to uncover training effects at a higher level of

the hierarchy that would have elements identical to the training effects at the next lower level of the hierarchy. It is probably reasonable to assume that training effects at one level of the hierarchy are necessary but not sufficient for training effects at the next highest level only when the two sets of training effects have identical elements.

2. Regarding the second problem encountered in the current study, Figure V-1 shows that the training effects (general or specific) at the different levels of the hierarchy were measured by methods that differed greatly in their degree of objectivity. Therefore, the level of confidence in the results obtained at the different interfaces between the levels of the hierarchy could vary depending upon the degree of objectivity of the method used to obtain those results. A more complete multicontent-multimethod test of the hierarchy would apply several methods of measurement to each interface between the levels of the hierarchy of training effects.

Looking at the results of the current study in light of the multicontent-multimethod matrix of Figure V-1 uncovers implications for future use of the Hamblin hierarchy of training effects. These implications fall into two categories: the first category has to do with testing the hierarchy theory, and the second relates to training evaluators.

IMPLICATIONS FOR FUTURE TESTS OF THE HIERARCHY THEORY

The current study uncovered two major implications for future tests of the hierarchy of training effects. Each of these implications is related to the multicontent-multimethod matrix of Figure V-1.

The first major implication is that training effects at each level of the hierarchy tend to fall along a continuum ranging from very general to very specific training effects. To properly test the hierarchy, future researchers should attempt to include training effects at each level from the full range of points along this continuum. Also with regard to the content dimension, one might expect that, when specific training effects at one level of the hierarchy (e.g., improvement in learning of specific planning behavior--level 2) are content-related to specific training effects at the next highest level (e.g., improvement in specific on-the-job planning behavior--level 3), the former training effects would be a necessary but not sufficient condition for the latter training effects to occur. However, with the exception of general reactions (level 1), chances are not great that general training effects at one level of the hierarchy are content-related to either general or specific training effects at another level. Future testing is needed to determine the extent to which this condition is valid. It is important that the assessment of training needs attempt to identify the relatively specific training effects at one level of the hierarchy that might be expected to be

content-related to certain relatively specific training effects at the next highest level of the hierarchy.

The final implication for future tests of the hierarchy relates to the method dimension of the multicontent-multimethod matrix of Figure V-1. This dimension shows that the method used to measure the training effects at each level of the hierarchy may fall anywhere along a continuum from very subjective (a schematic value of 1 on the continuum) to very objective (a schematic value of 10 on the continuum). Future training researchers should use several measurement methods falling along the full range of this continuum in order to test adequately each interface between the levels of the hierarchy.

The above implications for future tests of the Hamblin hierarchy also hold meaning for training specialists who intend to evaluate training courses, but who do not specifically intend to test the hierarchy. These implications for training specialists are stated in the next section of this chapter.

IMPLICATIONS FOR TRAINING EVALUATORS

The implications of the current study for training evaluators reflect the two major implications for testing the hierarchical theory. That is, if training evaluators intend to use the hierarchy method of evaluation they should build their training evaluation designs with the following two implications in mind.

--Attempt to include training effects at each level of the

hierarchy from the full range of points along the content dimension of the multicontent-multimethod matrix of Figure V-1. Further, they might attempt to identify--through assessment of training needs--the relatively specific training effects at one level of the hierarchy that might be expected to be content-related to certain relatively specific training effects at the next highest level of the hierarchy.

--Attempt to use several measurement methods falling along the full range of the method continuum of the multicontent-multimethod matrix of Figure V-1.

Obviously, these implications were also relevant to the client organization. Consequently, these implications, as well as the specific results regarding the Management II training course, were reported to the client organization. This feedback to the client organization is the topic of the next section of this chapter.

FEEDBACK TO THE CLIENT ORGANIZATION

The client organization--the Personnel Development Division in Michigan state government--was particularly interested in the findings regarding the Management II course, their most frequently offered course.

The following results regarding the Management II course were reported to the client organization:

--Management II led to an improvement in general reactions (level 1) and an improvement in general learning (level 2); apparently, the improvement in general learning

(level 2) was preceded by the improvement in general reactions (level 1).

--Regarding the next highest interface in the hierarchy of training effects, Management II led to an improvement in one specific job behavior (communicating behavior--level 3) that was preceded by the improvement in general learning (level 2). Apparently, the improvement in general learning (level 2) included enough communication learning content to produce an improvement in specific communicating behavior (level 3). Management II also led to improvements in managing behavior and job training behavior (level 3)--two other specific job behaviors--but these improvements were not found to be preceded by the improvement in general learning (level 2). Apparently, the improvement in general learning (level 2) did not include enough specific managing and job training content to produce improvements in the specific job behaviors of managing behavior and job training behavior (both level 3).

--Management II did not lead to an improvement in subordinate satisfaction (level 4)--a specific organizational variable--even though improvements were found in each of the three specific job behaviors (managing, communicating, and job training--all level 3). Apparently, certain organizational events--perhaps similar to those uncovered by the empirical evidence reviewed in Chapter II--prevented the specific job behavior improvements

(level 3) from producing the specific improvement in subordinate satisfaction with supervisory style (level 4). That is, the former improvements (level 3) may have been necessary but not sufficient for the latter improvement (level 4).

Other feedback to the client organization included the implications specified in the prior section of this chapter.

APPENDIX

Exhibit I: Instruments Administered to Experimental and Control Subjects

PRE-TRAINING
QUESTIONS FOR THE TRAINEE

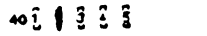
ATTITUDES TOWARD MANAGEMENT II

Here are some statements about the Management II training course. Indicate how much you agree or disagree with each:

	Strongly disagree	Disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Agree	Strongly agree
I would personally rate the course highly	1	2	3	4	5	6	7
Management II can give me little practical knowledge and skills	1	2	3	4	5	6	7
The knowledge gained in Management II should be very useful on my job	1	2	3	4	5	6	7
Management II can reduce the absenteeism of my subordinates by improving my supervisory style	1	2	3	4	5	6	7
I expect to learn a great deal from Management II	1	2	3	4	5	6	7
Management II will not influence the way I work	1	2	3	4	5	6	7
My department has a low opinion of Management II	1	2	3	4	5	6	7
Management II will have no noticeable effect upon the performance of my subordinates	1	2	3	4	5	6	7
I will be able to use the skills and knowledge learned in Management II back on the job	1	2	3	4	5	6	7
I do not consider Management II to be an important course for people in my job	1	2	3	4	5	6	7
Management II can result in improved employee morale because of better supervision	1	2	3	4	5	6	7
Most trainees do not seem to learn much in Management II	1	2	3	4	5	6	7
Management II is not pertinent to my job situation	1	2	3	4	5	6	7
Management II will have no effect upon the turnover in my department	1	2	3	4	5	6	7
Information learned in Management II will be forgotten quickly	1	2	3	4	5	6	7
Management II is not an effective training course	1	2	3	4	5	6	7

TEST OF SUPERVISORY KNOWLEDGE

The following questions measure your knowledge of the topics covered by the Management II training course. Please use a #2 pencil, and choose your answer by darkening in the desired space on the answer form provided. Darken the entire space, but try not to mark outside of the chosen space. If you decide to change an answer please erase the old answer completely. Assuming you chose answer (2) for question 40, your answer sheet should look as follows:



Please write your name in the space at the top of the answer form, and specify your last name in the spaces to the right of the form by darkening in the appropriate spaces. As you have probably guessed by now, this test is scored by computer; your cooperation in completing the answer form as neatly as possible would be greatly appreciated.

1. Management may best be defined as the activity of

   (1) keeping costs low and maximizing profits.
   (2) building and directing an organization.
   (3) maximizing the income of both employers and employees.
   (4) getting things done through people and directing the efforts of individuals toward a common objective.

2. On the first managerial level, it is necessary for the manager to possess managerial and technical skills. As the manager rises in the hierarchy,

   (1) it is likely that technical skills become more important.
   (2) he will practice his managerial skills more and use his technical skills less.
   (3) he should insist on exercising and practicing his technical skills so that he will always know as much or more than his subordinates.
   (4) he should never lose the outlook of the specialist.

3. Although there is much overlapping in the five managerial functions, it is logical that one should take place first. Which of the following is it?

   (1) organizing
   (2) motivating
   (3) planning
   (4) controlling
   (5) maintaining

4. Modern philosophies of management tend to say that

   (1) free enterprise and competition are destructive in terms of human values.
   (2) people possess capacity for exercising initiative, accepting responsibility, and making worthwhile contributions.
   (3) the average worker has an inherent dislike for work, avoids responsibility, and lacks ambition.
   (4) none of the above.

5. A need is stable (unchanging) over a long period of time.

   (1) true
   (2) false

6. Behavioral scientists have determined that people are remarkably alike in their needs and don't differ much from one need to another.

   (1) true
   (2) false

7. One who is attempting to become the best he can at his job is self-actualizing.

   (1) true
   (2) false

8. Which statement best describes the relationship between factors which merely maintain morale and factors which motivate?

   (1) theoretically, if all maintenance factors were met, workers would be highly motivated.
   (2) motivation factors involve the content of the job; maintenance factors involve the environmental context around the job.
   (3) absence of motivating factors leads to high dissatisfaction with the job.

9. The first point to be considered in planning is

   (1) determine the number of people assigned.
   (2) determine the method of instruction
   (3) review the facts
   (4) determine the location.

10. Communications put in writing will not be misunderstood.

    (1) true
    (2) false

11. The fundamental principle of effective communication is to make the other person accept your point of view.

    (1) true
    (2) false

12. There are many advantages of oral communication over written communications. Which of the following is <u>not</u> an advantage of oral communication?

    (1) a greater number of people can express themselves more easily and more completely by voice.
    (2) oral communication normally takes less time than written communication.
    (3) there is normally a degree of formality and permanency connected with oral communication.
    (4) the human voice can impart the message with meaning and enphasis.

13. Planning is not physical in nature; rather, it is mental preparation for the future.

    (1) true
    (2) false

14. The attitude of the supervisor toward training directly affects the climate of learning for his subordinates.

    (1) true
    (2) false

15. Plateaus of learning are common in learning a skilled task.

    (1) true
    (2) false

16. The boss should know all the answers to keep the respect of his group.

    (1) true
    (2) false

17. Which of the following statements regarding line and staff is correct?

    (1) the terminology "line and staff" is a way of distinguishing between the productive and non-productive efforts in an enterprise.
    (2) line personnel are those who are engaged in the primary activities, and staff are those who only indirectly contribute.
    (3) line and staff are characteristics of authority relationships.
    (4) staff is inferior to line authority.

18. Staff people should assume the responsibility for operating the innovations they originate.

    (1) true
    (2) false

19. Developing employees is primarily a staff function.

    (1) true
    (2) false

20. Which of the following is the "hierarchy of human needs"?

    (1) physiological, social, esteem, security, self-actualization.
    (2) security, physiological, social, esteem, self-actualization.
    (3) physiological, security, social, esteem, self-actualization.
    (4) none of the above.

21. Which of the following is not an activity in the function of controlling?

    (1) auditing
    (2) training
    (3) measuring
    (4) evaluating
    (5) correcting

22. A supervisor performs operating work at the expense of his or her primary job -- management work.

    (1) true
    (2) false

23. Which of the following are reasons why a supervisor may give operating work priority over managing work?

(1) operating work is often more familiar.
(2) ability to solve difficult operating or technical problems has traditionally been considered the trademark of a successfull supervisor in many organizations.
(3) operating work often provides more immediate personal satisfaction than does managing work.
(4) all of the above.

24. Behavior often happens without cause.

(1) true
(2) false

25. Changes in an individual's behavior from situation to situation can only be explained in terms of environmental conditions (that is, forces external to him).

(1) true
(2) false

26. Self-esteem is

(1) how a person views himself.
(2) how a person thinks others view him.
(3) how a person reaches his own potential.
(4) both (1) and (2) above.

27. Effective feedback has several characteristics. Which of the following is one of them?

(1) it is general rather than specific.
(2) it considers primarily the needs of the receiver of the feedback, not the needs of the giver.
(3) it is directed at behavior which the receiver is unable to control.
(4) it is descriptive rather than evaluative.

28. In general, the earlier the feedback, the better.

(1) true
(2) false

29. Having the receiver of feedback repeat or rephrase the feedback he has received,

(1) frustrates the receiver.
(2) helps to insure clear communication.
(3) confuses the receiver.
(4) insults the reciever.
(5) both (2) and (4) above.

30. Inaction can be a form of communication.

(1) true
(2) false

31. To be a more effective listener, one should

(1) ask questions.
(2) empathize (attempt to understand the other person's feelings).
(3) stop talking.
(4) all of the above.
(5) both (2) and (3) above.

32. The six steps of job training are listed below (they are also called the 6 D's):

a. Demonstration
b. Discussion
c. Determine the training need
d. Detailed explanation
e. Drill
f. Define the training purpose

Which of the following is the correct order in which they should be performed?

(1) b,a,d,c,f,e
(2) c,d,f,a,e,b
(3) c,f,d,a,e,b
(4) f,c,d,a,e,b

33. One step in job training is allowing for drill. Which of the following purposes does this serve?

(1) it helps the trainee to understand why the skill is needed.
(2) it allows the trainee to try the skill.
(3) it allows the trainer to demonstrate the skill to the trainee.
(4) it provides for an analysis of the skills required.

34. Supervisors can be concerned about people, production, both, or neither. A supervisor who is very concerned about people would

(1) allow employees much freedom in their work.
(2) encourage the use of uniform procedures.
(3) needle employees for greater effort.
(4) decide what should be done and how it should be done.

35. A supervisor who is very concerned about production would

(1) let employees do the work the way they think best.
(2) turn the employees loose on a job, and let them go to it.
(3) be willing to make changes.
(4) schedule the work to be done.

36. The more effective supervisors probably have

(1) a high concern for production and a low concern for people.
(2) a high concern for production and a high concern for people.
(3) a low concern for production and a high concern for people.
(4) a low concern for production and a low concern for people.

QUESTIONS FOR THE SUBJECT

State auditors and the Management Advisory Committee have asked that state training programs be evaluated to determine if they are improving behavior on the job. Consequently, the Personnel Development Division has asked a research team from Michigan State University to evaluate the Management II training course. Data on 75 supervisors will be gathered from the supervisors themselves, their bosses, and some of their subordinates. All data will be mailed in sealed envelopes directly to the MSU research team, and data on individuals will not be released to any component of state government. The Personnel Development Division is interested only in summaries of the data covering the entire subject group.

You have been chosen as one of the subjects for the study, and the following is the first of a series of questionnaires you will be asked to complete. After you have filled out this questionnaire, place it into the envelope provided to you, seal the envelope, and give it to your boss. He or she will place it into another sealed envelope with other data collected from your subordinates and mail the latter to the MSU researchers. Your boss will not see the data specified by your subordinates.

1. Please specify your name _____. This is needed only so we may combine all of your data for the analysis.

2. Indicate the degree to which you believe you have exhibited the following behaviors when you have communicated with your subordinates:

	Never		Sometimes		Often		Almost always
Showed a desire to listen	1	2	3	4	5	6	7
Put the subordinate at ease	1	2	3	4	5	6	7
Showed an understanding of the subordinate's feelings	1	2	3	4	5	6	7
Exhibited patience	1	2	3	4	5	6	7
Avoided arguing	1	2	3	4	5	6	7
Avoided criticizing	1	2	3	4	5	6	7
Eliminated distractions (closed the door, stopped receiving phone calls, etc.)	1	2	3	4	5	6	7

3. It is probable that your subordinates' jobs have changed, even if only slightly, over the last year or two. Possibly you have had to teach them new skills, how to do things differently, etc. This is called job training. Assuming you have used job training to teach current subordinates new skills, or even to break in a new employee, specify below the frequency with which you follow the steps listed.

	Never		Sometimes		Often		Almost always
Determined the need for training (analyzed skills required)	1	2	3	4	5	6	7
Defined the purpose of the skill to the subordinate	1	2	3	4	5	6	7
Explained the new skill in detail to the subordinate	1	2	3	4	5	6	7
Demonstrated the new skill to the subordinate	1	2	3	4	5	6	7
Allowed the subordinate to try the new skill himself, and pointed out any errors made	1	2	3	4	5	6	7
Discussed with the subordinate any questions or difficulties he may have had	1	2	3	4	5	6	7

4. A supervisor or manager concerns himself with two general types of work: managing work and operating work. Managing work can be defined as any work which must be performed by the supervisor himself, because it cannot be performed as well by subordinates or staff gruups. It includes planning, organizing, maintaining, motivating, and controlling the work of others; it means getting the work done through people. Operating work is all other work the supervisor performs, in effect accomplishing the result himself.

   Specify below the percentage of time you believe you spend on managing work and operating work during a typical work week. Try to avoid non-typical work weeks, such as when an unusual amount of overtime is required.

	Percentage of Time	
Managing work	_____ %	Please be sure that these add
Operating work	_____ %	to 100%

POST-TRAINING
QUESTIONS FOR THE TRAINEE

ATTITUDES TOWARD MANAGEMENT II

Here are some statements about the Management II training course. Indicate how much you agree or disagree with each:

	Strongly disagree	Disagree	Slightly disagree	Neither agree nor disagree	Slightly agree	Agree	Strongly agree
I would personally rate the course highly	1	2	3	4	5	6	7
Management II has given me little practical knowledge and skills	1	2	3	4	5	6	7
The knowledge gained in Management II will be very useful on my job	1	2	3	4	5	6	7
Management II will reduce the absenteeism of my subordinates by improving my supervisory style	1	2	3	4	5	6	7
I have learned a great deal from Management II	1	2	3	4	5	6	7
Management II will not influence the way I work	1	2	3	4	5	6	7
My department has a low opinion of Management II	1	2	3	4	5	6	7
Management II will have no noticeable effect upon the performance of my subordinates	1	2	3	4	5	6	7
I will be able to use the skills and knowledge learned in Management II back on the job	1	2	3	4	5	6	7
I do not consider Management II to be an important course for people in my job	1	2	3	4	5	6	7
Management II will result in improved employee morale because of better supervision	1	2	3	4	5	6	7
Most trainees do not seem to learn much in Management II	1	2	3	4	5	6	7
Management II is not pertinent to my job	1	2	3	4	5	6	7
Management II will have no effect upon the turnover in my department	1	2	3	4	5	6	7
Information learned in Management II will be forgotten quickly	1	2	3	4	5	6	7
Management II is not an effective training course	1	2	3	4	5	6	7

Please indicate your level of satisfaction with the following aspects of Management II:

	Very dissatisfied			Neither satisfied nor dissatisfied			Very satisfied
Instructor #1:							
His or her knowledge of the subject	1	2	3	4	5	6	7
His or her ability to impart the knowledge to you	1	2	3	4	5	6	7
His or her classroom style	1	2	3	4	5	6	7
Instructor #2:							
His or her knowledge of the subject	1	2	3	4	5	6	7
His or her ability to impart the knowledge to you	1	2	3	4	5	6	7
His or her classroom style	1	2	3	4	5	6	7
Instructor #3:							
His or her knowledge of the subject	1	2	3	4	5	6	7
His or her ability to impart the knowledge to you	1	2	3	4	5	6	7
His or her classroom style	1	2	3	4	5	6	7
The facilities -- consider seating arrangements, air conditioning, smoking arrangements, insulation from outside distractions, etc.	1	2	3	4	5	6	7

Please comment on any items above receiving a response of 3 or less:_____

_____

_____

QUESTIONS FOR THE SUBJECT

As you may remember, a team of researchers from Michigan State University is examining the effectiveness of the February 28th Management II course. The following is the last series of questionnaires you will be asked to fill out. Many of the questions are the same as those you answered earlier; we need your responses again so we can determine if your knowledge or attitudes have changed in the last two months. That is, do the effects of the course last very long?

While answering questions on this and the next page, please focus on your behavior since March 7th, if possible. Please remember that your responses will be seen only by the MSU researchers.

After you have filled out the questionnaires, place them into the envelope provided to you, seal the envelope, and give it to your boss. He or she will place it into another envelope with other questionnaires and mail it to the researchers. Your boss will not see your responses.

1. Please specify your name _____. This is needed only so we may combine all of your data for the analysis.

2. Indicate the degree to which you believe you have exhibited the following behaviors when you have communicated with your subordinates since March 7th, 1977:

	Never		Sometimes		Often		Almost always
Showed a desire to listen	1	2	3	4	5	6	7
Put the subordinate at ease	1	2	3	4	5	6	7
Showed an understanding of the subordinate's feelings	1	2	3	4	5	6	7
Exhibited patience	1	2	3	4	5	6	7
Avoided arguing	1	2	3	4	5	6	7
Avoided criticizing	1	2	3	4	5	6	7
Eliminated distractions (closed the door, stopped receiving phone calls, etc.)	1	2	3	4	5	6	7

3. It is probable that your subordinates' jobs have changed, even if only slightly, over the last year or two. Possibly you have had to teach them new skills, how to do things differently, etc. This is called job training. Assuming you have used job training to teach current subordinates new skills, or even to break in a new employee, since March 7th, 1977, please specify the degree to which you have performed the steps listed below:

	Never		Sometimes		Often		Almost always
Determined the need for training (analyzed skills required)	1	2	3	4	5	6	7
Defined the purpose of the skill to the subordinate	1	2	3	4	5	6	7
Explained the new skill in detail to the subordinate	1	2	3	4	5	6	7
Demonstrated the new skill to the subordinate	1	2	3	4	5	6	7
Allowed the subordinate to try the new skill himself, and pointed out any errors made	1	2	3	4	5	6	7
Discussed with the subordinate any questions or difficulties he may have had	1	2	3	4	5	6	7

4. A supervisor or manager concerns himself with two general types of work: managing work and operating work. Managing work can be defined as any work which must be performed by the supervisor himself, because it cannot be performed as well by subordinates or staff groups. It includes planning, organizing, maintaining, motivating, and controlling the work of others; it means getting the work done through people. Operating work is all other work the supervisor performs, in effect accomplishing the result himself.

Specify below the percentage of time you believe you spend on managing work and operating work during a typical work week. Try to avoid non-typical work weeks, such as when an unusual amount of overtime is required. Please remember to consider only the time since March 7th, 1977:

	Percentage of Time	
Managing work	_____ %	Please be sure that these add to 100%
Operating work	_____ %	

Questions on Factors Influencing Behavior Change

The information you provide on this page may be the most important information you will give us. Please be very frank in your answers. Remember that your responses will be seen only by the Michigan State University researchers.

The purpose of Management II is to improve your ability to supervise. However, you may have found it difficult to apply the new knowledge back on the job; you may have encountered any of several barriers to change. Examples of these barriers include pressure from the boss, policies or procedures which are incompatible with the new knowledge, resistance to change on the part of your subordinates, job pressures, etc.

Listed below are three changes in supervisory behavior which are expected to occur because of Management II. For each behavior change which you were unable to implement, please list the reasons in the spaces provided. Remember: the more you tell us, the better chance we have of redesigning the course to overcome these barriers.

More time spent on managing activities, such as planning, organizing, motivating, maintaining, controlling; less time on operating activities.

_____

_____

_____

_____

_____

More effective verbal communicating (better removal of barriers to communication, better listening, etc.)

_____

_____

_____

_____

_____

Increased use of the six D's of job training

_____

_____

_____

_____

_____

Exhibit II: Instruments Administered to Superiors

STATE OF MICHIGAN

WILLIAM G. MILLIKEN, Governor

DEPARTMENT OF CIVIL SERVICE

LEWIS CASS BUILDING, LANSING, MICHIGAN 48913

RICHARD A. ROSS, State Personnel Director

February 18, 1977

Dear _____:

In accordance with the recommendations of the Management Advisory Committee and the state auditors, we are evaluating the effectiveness of the Management II training course. The study is being conducted by a team of researchers from Michigan State University; it focuses on the February 28 course in Detroit.

Data will be collected on two types of subjects: trainees, who will attend the course; and "control subjects", who are supervisors like the trainees but who have not attended Management II in the past and will not attend the February 28 course. A comparison of data on trainees and controls allows us to determine if behavior changes during the time of the study are actually caused by the training course.

We need your assistance in this study, and we hope you will proceed as follows:

1. for each trainee from your division, choose one control subject. This can be any first level supervisor like the trainee. For example, if you are sending two trainees to the course, pick two control subjects on whom data can be gathered.

2. for each trainee and each control subject, fill out the questionnaire entitled "QUESTIONS FOR THE SUPERIOR".

3. give each control subject one small envelope and one questionnaire entitled "QUESTIONS FOR THE SUBJECT". He should fill this out, seal it in the envelope, and return it to you. You need not give this questionnaire to trainees because they will fill them out during the course.

4. for each trainee and each control subject, choose at random three of that person's subordinates. Give to each subordinate one small envelope and one questionnaire entitled "QUESTIONS FOR THE EMPLOYEE". Each should fill out the questionnaire, seal it in the envelope, and give it to you.

5. place all of the above envelopes and your questionnaires into the large pre-addressed envelope provided to you. Please mail this to the researchers no later than March 4, 1977; it is important that this data be collected before the trainee returns from the course (because it is to reflect his job behavior before the course). Do not allow individuals to see other peoples' responses to the questionnaires; these should not even be discussed because more data will be collected in about two months.

QUESTIONS FOR THE SUPERIOR

Your responses to the following questions will be seen only by the MSU research team. Only summary data for the entire training group and the entire control group will be provided to the Personnel Development Division and the auditors; data on individuals will be kept confidential.

1. What is the name of the trainee or control person for whom you are filling out this questionnaire?

_____

2. A supervisor or manager concerns himself with two general types of work: managing work and operating work. Managing work can be defined as any work which must be performed by the supervisor himself, because it cannot be performed as well by subordinates or staff groups. It includes planning, organizing, maintaining, motivating, and controlling the work of others; it means getting the work done through people. Operating work is all other work the manager performs, in effect accomplishing the result himself.

   For the trainee or control person in question, specify below the percentage of time you believe he or she spends on managing work and operating work during a typical work week. Try to avoid non-typical work weeks, such as when an unusual amount of overtime is required.

	Percentage of Time	
Managing work	_____ %	Please be sure that these add
Operating work	_____ %	to 100%

To: Date: April 27, 1977

Subject: Management II Evaluation Study

As you may remember, a team of researchers from Michigan State University is evaluating the effectiveness of the February 28th Management II course. Your assistance in this study is greatly appreciated.

Earlier, we received from you data regarding the job behavior regarding the following persons:

Trainees	Control Subjects

Remember that a comparison of data on trainees with data on control subjects allows us to determine the effectiveness of the course. We again need to collect data on the job behavior of the above persons, and we hope you will proceed as follows:

1. fill out the questionnaire entitled "QUESTIONS FOR THE SUPERIOR".

2. give each trainee and each control subject one small envelope and one questionnaire entitled "QUESTIONS FOR THE SUBJECT". He or she should fill this out, seal it in the envelope, and return it to you.

3. four to six weeks ago, some of the subordinates of these trainees and control subjects filled out a questionnaire entitled "QUESTIONS FOR THE EMPLOYEE". To each of these same subordinates please give one small envelope and one of the questionnaires, "QUESTIONS FOR THE EMPLOYEE". Each should fill out the questionnaire, seal it in the envelope, and return it to you.

4. place all of the above envelopes and your questionnaire into the large pre-addressed envelope and mail it to the MSU researchers by May 4, 1977. Please do not allow individuals to see other peoples' questionnaire responses.

All data will be treated confidentially; only the MSU research team will see the data on individuals. You are welcome to receive a copy of the summarized results, however. If you have any questions, please contact Ken Dawson at 373-3385 or Ron Clement at 353-5415.

Again, thank you for your help.

Best regards,

Kenneth Dawson
Training Coordinator
Personnel Development Division

QUESTIONS FOR THE SUPERIOR

Your responses to the following questions will be seen only by the MSU research team. Only summary data for the entire training group and the entire control group will be provided to the Personnel Development Division and the auditors; data on individuals will be kept confidential.

1. What is the name of the trainee or control person for whom you are filling out this questionnaire?

   _____

2. A supervisor or manager concerns himself with two general types of work: managing work and operating work. Managing work can be defined as any work which must be performed by the supervisor himself, because it cannot be performed as well by subordinates or staff groups. It includes planning, organizing, maintaining, motivating, and controlling the work of others; it means getting the work done through people. Operating work is all other work the manager performs, in effect accomplishing the result himself.

   For the trainee or control person in question, specify below the percentage of time you believe he or she spends on managing work and operating work during a typical work week. Try to avoid non-typical work weeks, such as when an unusual amount of overtime is required. Please consider only the time period since March 7th, 1977.

	Percentage of Time	
Managing work	_____ %	Please be sure that these add
Operating work	_____ %	to 100%

Exhibit III: Instruments Administered to Subordinates

QUESTIONS FOR THE EMPLOYEE

State auditors and the Management Advisory Committee have asked that state training programs be evaluated to determine if they are improving behavior on the job. Consequently, the Personnel Development Division has asked a research team from Michigan State University to evaluate the Management II training course. Data on 75 supervisors will be gathered from the supervisors themselves, their bosses, and some of their subordinates. The data will be mailed in sealed envelopes directly to the MSU research team, and data on individuals will not be released to any component of state government. The Personnel Development Division is interested only in summaries of the data covering the entire subject group.

Your supervisor has been chosen as one of the subjects on whom data will be gathered, and the following is one of two questionnaires you will be asked to complete. After you have filled out the questionnaire, place it into the envelope provided to you, seal the envelope, and give it to your supervisor's boss. He or she will place it into another sealed envelope containing other data and mail the envelope to the MSU researchers.

1. Please specify the name of your supervisor _____. This is needed only so we may combine all data for that person in the analysis.

2. Indicate the degree to which your supervisor exhibits the following behaviors when he or she verbally communicates with you:

	Never		Sometimes		Often		Almost always
Shows a desire to listen	1	2	3	4	5	6	7
Puts you at ease	1	2	3	4	5	6	7
Shows an understanding of your feelings	1	2	3	4	5	6	7
Exhibits patience	1	2	3	4	5	6	7
Avoids arguing	1	2	3	4	5	6	7
Avoids criticizing	1	2	3	4	5	6	7
Eliminates distractions (closes the door, stops receiving phone calls, etc.)	1	2	3	4	5	6	7

3. It is probable that your job has changed, even if only slightly, over the past year or two. Possibly you have had to learn new skills, how to do some things differently, etc. You may have learned a whole new job. Often the supervisor is responsible for teaching employees new skills or breaking them in on a new job. This is called job training. Assuming your supervisor has trained you in a new skill, specify below the degree to which he or she performed the steps listed:

	Never		Sometimes		Often		Almost always
Determined the need for training (analyzed skills required)	1	2	3	4	5	6	7
Defined the purpose of the new skill for you	1	2	3	4	5	6	7
Explained the new skill to you in detail	1	2	3	4	5	6	7
Demonstrated the new skill to you	1	2	3	4	5	6	7
Allowed you to try the new skill and pointed out any errors you made	1	2	3	4	5	6	7
Discussed with you any questions you may have had	1	2	3	4	5	6	7

4. The following scale allows you to describe the style of your supervisor. If you believe the item describes your supervisor, place a "Y" in the space provided; if you believe the item does not describe your supervisor, place an "N" in the space provided; if you cannot decide, place a "?" in the space provided:

__ Asks my advice

__ Hard to please

__ Impolite

__ Praises good work

__ Tactful

__ Influential

__ Up-to-date

__ Doesn't supervise enough

__ Quick-tempered

__ Tells me where I stand

__ Annoying

__ Stubborn

__ Knows job well

__ Bad

__ Intelligent

__ Leaves me on my own

__ Around when needed

__ Lazy

QUESTIONS FOR THE EMPLOYEE

As you may remember, a team of researchers from Michigan State University is examining the effectiveness of the February 28th Management II course. Your supervisor has been chosen as one of the subjects on whom data will be gathered, and the following is the last questionnaire you will be asked to complete. Your responses below will be seen only by the researchers.

While filling out this questionnaire, think of your supervisor's job behavior only since March 7th (which was the first work day after the course). After you have filled out the questionnaire, place it into the envelope provided to you, seal the envelope, and give it to your supervisor's boss. He or she will place it into another envelope with other data and mail it to the researchers.

1. Please specify the name of your supervisor _____. This is needed only so we may combine all data for that person in the analysis.

2. Indicate the degree to which your supervisor exhibits the following behaviors when he or she verbally communicates with you (remember -- focus on the period since March 7th):

	Never		Sometimes		Often		Almost always
Shows a desire to listen	1	2	3	4	5	6	7
Puts you at ease	1	2	3	4	5	6	7
Shows an understanding of your feelings	1	2	3	4	5	6	7
Exhibits patience	1	2	3	4	5	6	7
Avoids arguing	1	2	3	4	5	6	7
Avoids criticizing	1	2	3	4	5	6	7
Eliminates distractions (closes the door, stops receiving phone calls, etc.)	1	2	3	4	5	6	7

3. It is probable that your job has changed, even if only slightly, over the past year or two. Possibly you have had to learn new skills, how to do some things differently, etc. You may have learned a whole new job. Often the supervisor is responsible for teaching employees new skills or breaking them in on a new job. This is called job training. Assuming your supervisor has trained you in a new skill since March 7th, 1977, specify below the degree to which he or she performed the steps listed:

	Never		Sometimes		Often		Almost always
Determined the need for training (analyzed skills required)	1	2	3	4	5	6	7
Defined the purpose of the new skill for you	1	2	3	4	5	6	7
Explained the new skill to you in detail	1	2	3	4	5	6	7
Demonstrated the new skill to you	1	2	3	4	5	6	7
Allowed you to try the new skill and pointed out any errors you made	1	2	3	4	5	6	7
Discussed with you any questions you may have had	1	2	3	4	5	6	7

4. The following scale allows you to describe the style of your supervisor. If you believe the item describes your supervisor, place a "Y" in the space provided; if you believe the item does not describe your supervisor, place an "N" in the space provided; if you cannot decide, place a "?" in the space provided:

__ Asks my advice

__ Hard to please

__ Impolite

__ Praises good work

__ Tactful

__ Influential

__ Up-to-date

__ Doesn't supervise enough

__ Quick-tempered

__ Tells me where I stand

__ Annoying

__ Stubborn

__ Knows job well

__ Bad

__ Intelligent

__ Leaves me on my own

__ Around when needed

__ Lazy

REFERENCES

REFERENCES

Brayfield, A. H., and Crockett, W. H. Employee attitudes and employee performance. Psychological Bulletin, 1955, 52, 396-424.

Campbell, D. T., and Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Campbell, D. T., and Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand-McNally, 1966.

Campbell, J. P., Dunnette, M. D., Lawler, E. E., and Weick, K. E. Managerial behavior, performance, and effectiveness. New York: McGraw-Hill, 1970.

Catalanello, R. F., and Kirkpatrick, D. L. Evaluating training programs--the state of the art. Training and Development Journal, 1968, 22, 2-10.

Fleischmann, E. A. Leadership climate, human relations training, and supervisory behavior. Personnel Psychology, 1953, 6, 205-222.

Fromkin, H. L., Brandt, J., King, D. C., Sherwood, J. J., and Fisher, J. An evaluation of human relations training for police. Catalog of Selected Documents in Psychology, 1975, 5, 206-207.

Goldstein, A. P., and Sorcher, M. Changing supervisory behavior. New York: Pergamon Press, Inc., 1974.

Goldstein, I. L. Training: program development and evaluation. Monterey, Calif.: Brooks-Cole Publishing Co., 1974.

Goodacre, D. M., III. Experimental evaluation of training. Journal of Personnel Administration and Industrial Relations, 1955, 2, 143-149.

Hamblin, A. C. Evaluation and control of training. Maidenhead, Berkshire, England: McGraw-Hill Book Co., Ltd., 1974.

Hand, H. H., Richards, M. D., and Slocum, J. W., Jr. Organizational climate and the effectiveness of a human relations training program. Academy of Management Journal, 1973, 16, 185-195.

Handyside, J. D. An experiment with supervisory training. London: National Institute of Industrial Psychology, 1956.

Hariton, T. Conditions influencing the effects of training foremen in human relations principles. Unpublished doctoral dissertation, University of Michigan, 1951.

House, R. J. Leadership training: some dysfunctional consequences. Administrative Science Quarterly, 1968, 12, 556-571.

Kirkpatrick, D. L. Evaluation of training. In R. L. Craig and L. R. Bittel (Eds.), Training and Development Handbook, New York: McGraw-Hill, 1967.

Kerlinger, F. N., and Pedhazer, E. Multiple regression in behavioral research. New York: Holt, Rinehart, and Winston, 1973.

Latham, G. P., Wexley, K. N., and Purcell, E. D. Training managers to minimize rating errors in the observation of behavior. Journal of Applied Psychology, 1975, 60, 550-555.

Martin, H. O. The assessment of training. Personnel Management, 1957, 39, 88-93.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., and Best, D. H. Statistical package for the social sciences. New York: McGraw-Hill, 1975.

O'Reilly, A. P. Skill requirements: supervisor-subordinate conflict. Personnel Psychology, 1973, 26, 75-80.

Vroom, V. H. Work and motivation. New York: Wiley, 1964.

Wexley, K. N., and Yukl, G. A. Organizational behavior and personnel psychology. Homewood, Illinois: Richard D. Irwin, Inc., 1977.

REFERENCE NOTES

1. Hamblin, A. C. In a letter to the researcher, July 18, 1977.