

The Third in a Series of Four Articles . . .

Techniques For Evaluating Training Programs

Part 3 - Behavior

DR. DONALD L. KIRKPATRICK*

Advanced Management Development Administrator
International Minerals and Chemical Corporation
Skokie, Illinois

In the two previous articles in this series, we talked about techniques for evaluating training programs in terms of (1) REACTION and (2) LEARNING. It was emphasized that in our evaluations, we can borrow techniques but we cannot borrow results.

A personal experience may be the best way of starting this third article dealing with changes in behavior. When I joined The Management Institute of The University of Wisconsin in 1949, one of my first assignments was to sit through a one-week course on "Human Relations for Foremen and Supervisors." During the week I was particularly impressed by a foreman named Herman from a Milwaukee company. Whenever a conference leader asked a question requiring a good understanding of human relations principles and techniques, Herman was the first one who raised his hand. He had all the answers in terms of good human relations approaches. I was

very much impressed and I said to myself "if I were in industry, I would like to work for a man like Herman."

It so happened that I had a first cousin who was working for that company. And oddly enough, Herman was his boss. At my first opportunity, I talked with my cousin, Jim, and asked him about Herman. Jim told me that Herman may know all the principles and techniques of human relations, but he certainly does not practice them on the job. He performed as the typical "bull-of-the-woods" who had little consideration for the feelings and ideas of his subordinates.

At this time I began to realize there may be a big difference between knowing principles and techniques and using them on the job.

Robert Katz, Professor at Dartmouth wrote an article in the July-August 1956 issue of the *Harvard Business Review*. The article was called "Human Rela-

*Formerly Assistant Director, The Management Institute, University of Wisconsin, Madison, Wisconsin.

tions Skills Can Be Sharpened." And he said: "if a person is going to change his job behavior, five basic requirements must exist":

1. He must want to improve.
2. He must recognize his own weaknesses.
3. He must work in a permissive climate.
4. He must have some help from someone who is interested and skilled.
5. He must have an opportunity to try out the new ideas.

It seems that Katz has put his finger on the problems that exist in a transition between learning and changes in behavior on the job.

Evaluation of training programs in terms of on the job behavior is more difficult than the reaction and learning evaluations described in the two previous articles. A more scientific approach is needed and many factors must be considered. During the last few years a number of attempts have been made and more and more effort is being put in this direction.

Several guideposts are to be followed in evaluating training programs in terms of behavioral changes:

1. A *systematic* appraisal should be made of on-the-job performance on a *before-and-after* basis.
2. The appraisal of performance should be made by one or more of the following groups (The more the better):
 - A. The person receiving the training
 - B. His superior or superiors
 - C. His subordinates
 - D. His peers or other people thor-

oughly familiar with his performance.

3. A statistical analysis should be made to compare before and after performance and relate changes to the training program.
4. The post-training appraisal should be made three months or more after the training so that the trainees have an opportunity to put into practice what they have learned. Subsequent appraisals may add to the validity of the study.
5. A control group (not receiving the training) should be used.

Some of the best evaluation studies are briefly described below.

The Fleishman-Harris Studies¹

To evaluate a training program that had been conducted at the Central School of The International Harvester Company, Fleishman developed a study design and a battery of research instruments for measuring the effectiveness of the training. Seven paper-and-pencil questionnaires were used and the trainees, their superiors, and their subordinates were all surveyed.

To supplement the data that Fleishman had discovered, Harris conducted a follow-up study in the same organization. He used a before-and-after measure of job performance and worked with experimental and control groups. He obtained information from the trainees themselves as well as from their subordinates.

Survey Research Center Studies²

The Survey Research Center of the University of Michigan has contributed

much to evaluation of training programs in terms of on-the-job behavior. To measure the effectiveness of a human relations program conducted by Dr. Norman Maier at the Detroit Edison Co., and to measure the results of an experimental program called "feedback," a scientific approach to evaluation was used. A basic design was to use a before-and-after measure of on-the-job performance with experimental as well as control groups. The supervisors receiving the training as well as their subordinates were surveyed in order to compare the results of the research. The instrument used for measuring these changes was an attitude and opinion survey designed and developed by the Survey Research Center.

***The Lindholm Study*³**

This study was carried out in the home office of a small insurance company during the period of October, 1950 to May, 1951. A questionnaire developed as part of the research program of the Industrial Relations Center of the University of Minnesota was used. It was given on a before-and-after basis to the subordinates of those who took the training. No control group was used. A statistical analysis of the before-and-after results of the attitude survey determined the effectiveness of the program in terms of on-the-job behavior.

***The Blocker Study*⁴**

A different approach was used in the study conducted in an insurance company having approximately 600 employees. Fifteen supervisors who took a

course on "Democratic Leadership" were analyzed during the three-month period following the course. Eight of the supervisors were classified as democratic and seven were classified as authoritarian based on their behavior prior to the program.

During the three-month period immediately following the program, the changes in behavior of the supervisors were analyzed through a study of their interview records. They used standard printed forms which made provision for recording the reason for the interview, attitude of the employee, comments of the supervisor, and action taken, if any. Each supervisor was required to make a complete record of each interview. They did not know that these records were to be used for an evaluation study. There were a total of 376 interviews with 186 employees.

The interview records were classified as authoritarian or democratic. The changes in interview approach and techniques were studied during the three month period following the course to determine if on-the-job behavior of the supervisors changed.

***The Tarnopol Approach*⁵**

In his article called "Evaluate Your Training Program," Tarnopol suggests the approach to use as well as a specific example of an evaluation experiment. He believes in the employee attitude survey given on a before-and-after basis using control as well as experimental groups. He stresses that "in our experience, five employees is a good minimum for measuring the behavior of their supervisor." He also stresses that "although canned questionnaires are avail-

able, it is advisable to use measuring instruments that are specifically suited to the requirements of both your company and your training program."

In his employee attitude approach, Tarnopol has suggested inserting some neutral questions which do not relate to the training being given. This is an added factor in interpreting the results of the research.

***The Moon-Hariton Study*⁶**

Their study was made in an Engineering Section of a department of the General Electric Company in 1956. The staff of the General Electric Company was assisted by a representative of the Psychological Corporation.

In the spring of 1958, two years after the adoption of a new appraisal and training program, a decision was made to attempt to evaluate its effectiveness. It was felt that the opinion of the subordinates about changes in the managers' attitudes and behavior would provide a better measure than what the managers themselves thought about the benefits of the program. Thus a questionnaire was designed to obtain the subordinates' views about changes in their managers. Nevertheless, it was felt that the opinions of the manager would add to the picture. Accordingly, they were also surveyed.

The questionnaire asked the respondents to compare present conditions with what they were two years ago. In other words, instead of measuring the attitudes before and after the program, the subordinates and the managers were asked to indicate what changes had taken place during the last two years.

***The Buchanan-Brunstetter Study*⁷**

At the Republic Aviation Corporation, an attempt was made to measure the results of a training program. The questionnaire was used and an experimental and a control group were measured. The experimental group had received the training program during the past year while the control group was going to receive it during the following year. The subordinates of the supervisors in each one of these groups were asked to complete a questionnaire which related to the on-the-job behavior of their supervisor. After answering the questionnaire in which they described the job behavior of their supervisor, they were asked to go over the questionnaire again and to place a check opposite any items: "(1) which you think are *more* effectively done now than a year ago; (2) which you think are *less* effectively done now than a year ago."

In this experiment as well as in the Moon-Hariton approach, the subordinates were asked to indicate what changes in behavior had taken place during the last year. This was done because a before measure of their behavior had not been made.

***The Stroud Study*⁸**

A new training program called "Personal Factors in Management" was evaluated at the Bell Telephone Company of Pennsylvania by Peggy Stroud. Several different approaches were used to compare the results and obtain a more valid indication of on-the-job behavioral changes that resulted from the program. The first step was the formulation of a questionnaire to be filled out by four

separate groups: (1) conferees (2) controllees (supervisors not taking the course) (3) superiors of the conferees (4) superiors of the controllees.

The first part of the questionnaire was the "Consideration Scale" taken from the leader behavior description questionnaire originated in the Ohio State leadership studies. The second part of the questionnaire was called the Critical Incident section in which the conferee and control groups were asked to describe four types of incidents that had occurred on the job. The third and final section of the questionnaire applied to the conferees only. They were asked to rate the extent to which they felt the training course had helped them achieve each of its five stated objectives.

It was decided to conduct an extensive evaluation of the training program after the program had begun. Therefore it was not possible to make a before and after comparison. In this study, an attempt was made to get the questionnaire respondents to compare on-the-job behavior before the program with that following the program. According to Miss Stroud, it would have been better to measure behavior prior to the program and then compare it to behavior measured after the program.

This study called "Evaluating A Human Relations Training Program" is one of the best attempts this writer has discovered. The various evaluation results are compared and fairly concrete interpretations made.

The Sorensen Study⁹

The most comprehensive research that has been done to evaluate the effectiveness of a training program in terms of

on-the-job behavior was made at the Crotonville Advanced Management Course of the General Electric Company. It was called the "Observed Changes Enquiry."

The purpose of the "enquiry" was to answer these questions:

1. Have manager graduates of General Electric's Advanced Management Courses of 1956 been observed to have changed in their manner of managing?
2. What inferences may be made from similarities and differences of changes observed in graduates and non-graduates?

First of all, the managers (graduates and non-graduates alike) were asked to indicate changes they had observed in their own manner of managing during the previous twelve months. Secondly, subordinates were asked to describe changes they had observed in the managers during the past twelve months. Thirdly, their peers (looking sideways) were asked to describe changes in behavior. And finally, the superiors of the control and experimental groups were asked to describe the same changes in behavior. This gave Sorensen an excellent opportunity to compare the observed changes of all four groups.

In this extensive research, Sorensen used experimental as well as control groups. He also used four different approaches to measure observed changes. These include the man himself, his subordinates, his peers, and his superiors. In this research, he did not use a before-and-after measure but rather asked each of the participants to indicate what changes, if any, had taken place during the past year.

Summary and Conclusions

The purpose of this article has been to describe briefly some of the best experiments that have been used to measure effectiveness of training programs in terms of on-the-job behavior. Only the methods and instruments used in these studies have been mentioned. The results, although interesting, cannot be borrowed by other training directors but the techniques can.

For those interested in evaluating in terms of behavioral changes, it is strongly suggested that these studies be carefully analyzed. The references following this article indicate where the detailed articles can be found.

Once more I would like to emphasize that the future of training directors and their programs depends to a large extent on their effectiveness. To determine effectiveness, attempts must be made to measure in scientific and statistical terms. This article, dealing with changes in behavior resulting from training programs, indicates a very complicated procedure. But it is worthwhile and necessary if training programs are going to increase in effectiveness and their benefits made clear to top management.

It is obvious that very few training directors have the background, skill and time to engage in extensive evaluations. It is therefore frequently necessary to call on statisticians, research people, and consultants for advice and help.

REFERENCES

1. Fleishman, E. A., Harris, E. F., Buntt, H. E., "Leadership and Supervision in Industry." Personnel Research Board, Ohio State University, Columbus, Ohio, 1955, page 110.
2. Mann, Dr. Floyd, "Human Relations in the Industrial Setting." Survey Research Center, University of Michigan, Ann Arbor, Michigan.
3. Lindholm, T. R., "Supervisory Training and Employee Attitudes," *Journal of the ASTD*, Nov.-Dec., 1953.
4. Blocker, C. E., "Evaluation of a Human Relations Training Course," *Journal of the ASTD*, May-June, 1955.
5. Tarnopol, Lester, "Evaluate Your Training Program," *Journal of the ASTD*, Mar.-Apr., 1957.
6. Moon, C. G., Hariton, Theodore, "Evaluating an Appraisal and Feedback Training Program," *Personnel*, Nov.-Dec., 1958.
7. Buchanan, P. C., Brunstetter, P. H., "A Research Approach to Management Improvement," *Journal of the ASTD*, Jan. and Feb., 1959.
8. Stroud, P. V., "Evaluating a Human Relations Training Program," *Personnel*, Nov.-Dec., 1959.
9. Sorensen, Olav, "The Observed Changes Enquiry," Manager Development Consulting Service, General Electric Company, Crotonville, New York, May 15, 1958.