A COMPARATIVE ANALYSIS OF THREE FORMS OF
EVALUATING MANAGEMENT TRAINING PROGRAMS

THESIS

Presented to the Graduate Council of the
North Texas State University in Partial
Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

John P. Hale, B.B.A.
Denton, Texas
December, 1985

The practice of training evaluation has not kept pace with prescription, and evaluations being done are frequently negligent of appropriate controls needed to draw valid conclusions.

A comparison was made of training outcomes contrasting results obtained using carefully controlled scientific approaches with those from a more popular less scientific approach. The research design involved the collection and analysis of data from a single organization's managerial training program. Three different methods of training evaluation were studied: an "immediate reaction" rating sheet, a self-report participant survey, and a similar survey completed by the participants' subordinates.

Both survey results showed no significant changes in on-the-job behavior six weeks after training. In contrast the "immediate reaction" ratings were positive, implying the training program was a "success." Conclusions were drawn concerning the validity of methods compared.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>iv</td>
</tr>
<tr>
<td>A COMPARATIVE ANALYSIS OF THREE FORMS OF EVALUATING MANAGEMENT TRAINING PROGRAMS</td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Method</td>
<td>13</td>
</tr>
<tr>
<td>Subjects</td>
<td></td>
</tr>
<tr>
<td>Instruments</td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>19</td>
</tr>
<tr>
<td>Discussion</td>
<td>24</td>
</tr>
<tr>
<td>Limitations of the Present Research</td>
<td></td>
</tr>
<tr>
<td>Directions for Future Research</td>
<td></td>
</tr>
<tr>
<td>Appendix</td>
<td>34</td>
</tr>
<tr>
<td>References</td>
<td>44</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Descriptive Statistics for Fire Officer Tenure.</td>
<td>20</td>
</tr>
<tr>
<td>2. Descriptive Statistics for Subordinate Tenure.</td>
<td>21</td>
</tr>
<tr>
<td>3. Statistical Summary of Factor Scores for the Fire Officer Sample.</td>
<td>22</td>
</tr>
<tr>
<td>4. Statistical Summary of Factor Scores for the Subordinate Sample</td>
<td>23</td>
</tr>
</tbody>
</table>
A COMPARATIVE ANALYSIS OF THREE FORMS OF EVALUATING MANAGEMENT TRAINING PROGRAMS

The issue of training program evaluation has been a subject of considerable examination over the years. In general, its intent has been rarely criticized and usually encouraged. A review of training literature by the author however, revealed the practice of evaluation has not kept pace with the prescription. Evaluation research was scarce and when present was frequently negligent of appropriate controls needed to draw reasonably valid conclusions. There appeared to be a need to refocus attention on evaluation designs and the validity of information various techniques could provide. In their recent article, Swierczek and Carmichael (1985) emphasized that the key to evaluation training was developing a design that was methodologically sound as well as practical.

The rationale for not doing evaluations were abundant among professional trainers. There was a feeling of possible "self-exposure" in conducting a well-controlled evaluation. As Bunker and Cohen (1978) pointed out, the most concrete excuse was often unstated: If it doesn't work, I don't want to know about it.

No one pushes us to do it. Worse yet, we often have a gnawing anxiety that if we do research, the results
might be negative. Being in the dark is blissful when there is no apparent alternative to an ineffective program (Zenger & Hargis, 1982, p. 11).

However, as Brown (1980) emphasized, upper management was no longer writing off training as a cost, but they were seeing it as an investment. With that change in perspective, it was becoming increasingly important to be able to determine the benefits directly attributed to training programs. Sound evaluation techniques were becoming important to training directors who were being held accountable for the results their programs were designed to achieve.

Evaluations could serve as a source of information for improving future training. Trainers who could be "tough-minded" about the actual impact their programs were having could do a better job of setting objectives and designing programs; they would have something solid to work with when they stepped up to face their classes (Carnarius, 1975). Zenger and Hargis (1982) discussed benefits that could evolve out of evaluation research. They pointed out that within organizations, the training department was often seeking ways to increase their influence and visibility, and one way to avoid large budget cuts in "down-times" was to provide empirical support for the programs in which they were involved. More specifically, training specialists were seeking credibility in their organizations, and assessment data on their programs could help strengthen management's trust and confidence in them and their contributions.
In their 1978 article, Bunker and Cohen hypothesized that despite stated and unstated rationalizations to the contrary, competent evaluation was the cornerstone of meeting both organizational and individual educational needs, and of improving the cost effectiveness of any training program. Among industry experts there seemed to be a growing consensus on the importance of doing training evaluation, but what type of evaluation was not quite as clear cut.

Kirkpatrick (1979) identified four types of formal evaluation procedures: measures of reaction, learning, behavior, and results desired. Reaction measures were usually administered in the form of opinion surveys or comment sheets, and were gathered on a daily or program ending basis. Despite their popularity and ease of collection, this method was filled with unexamined assumptions. Not only did it assume that enjoyable courses were effective, but also that the trainee was competent to judge the quality of the training, which was questionable at best. Amusing and entertaining courses probably got higher ratings than those programs which were highly factual and presented in a dry-as-dust manner, even though the dull fact may have been the most important information a man had received in his career (Thorley, 1972). In addition, as Mezoff (1981) pointed out, collecting trainee self-reports was often affected by "response-shift bias." The training had
influenced the participants' frame of reference concerning the program's content and accurate self-appraisals were unlikely. They were more likely to report what they thought was expected of them now that they were "trained," rather than reporting their actual behaviors or opinions.

Assessment of learning behaviors were often standardized paper/pencil tests and were administered to determine what principles, facts, and techniques were understood by the trainees. The tests were scored to provide an index of the "effectiveness" of the training program. However, on-the-job use of this learned information was usually not addressed in this evaluation format (Kirkpatrick, 1979). Although knowledge was often a prerequisite for behavior change, there was little or no benefit realized from training that did not transfer to the work environment. Although there were several standardized tests in print such as How Supervise? and the Supervisory Inventory on Human Relations, this measurement technique appeared to be rather superficial when applied to training as complex as human relations.

Behavioral measures attempted to evaluate specific variable changes that could be attributed to the training, and were often designed as objective samples of on-the-job performance. A more scientific approach was often employed because this type of evaluation was more difficult than the reaction and learning evaluations previously discussed.
(Kirkpatrick, 1979). Kirkpatrick (1979) outlined several
guidelines for conducting a well-controlled, behaviorally
oriented evaluation. He recommended a before-after two-
group design in which the behavioral measures were
collected from at least one group other than the trainees
themselves. He stressed that the post-training measure
should be made after a time period that allowed sufficient
opportunity for the trainees to put into practice what they
had learned. This evaluation method was more time consuming,
but offered the advantages of objective behavioral appraisal
and controls for secondary variance.

In discussing the fourth type of evaluation procedure,
Kirkpatrick (1979) claimed that from an evaluation stand-
point, it would have been best to evaluate training programs
directly in terms of results desired. However, there were
so many complicating factors involved that it was almost
impossible to evaluate certain kinds of programs in terms
of organization-wide results. Measuring reduction of costs,
turnover, absenteeism, grievances and/or increases in pro-
duction quality or quantity that were direct results of a
management training program would have been a "herculean
task" due to the necessity of controlling the number of
potential confounding variables. In a similar discussion,
Thorley (1972) wrote,

One thought is to use gross measures such as the profit
and growth of the company, but it becomes nearly
impossible to separate out the influence of such large forces as the economy, market conditions, credit availability and the like from the effects of individual contribution. (p. 35)

In considering these four types of "formal" evaluation procedures as they were applied to supervisory training, there had been several research studies conducted to determine and analyze the techniques that were actually being used by business, industry, and government in the evaluation of programs. In their 1968 study, Catalanello and Kirkpatrick surveyed 154 firms who had previously been involved in human relations training (all had previously ordered the Supervisory Inventory on Human Relations). Of the responses they received, 78% reported attempts to measure trainee reaction following the program. However, approximately half of the respondents also indicated they attempted to measure learning, behavior, and/or results in some fashion. This latter group was re-surveyed in an attempt to receive more specific information concerning the quality of the evaluations. The respondents from the second questionnaire indicated the heavy use of learning measures (91 percent) which was possibly an artifact of the initial sample selection criteria for this study. Twenty-one organizations (44 percent) claimed to be attempting behavioral measures, but of these approximately half did not use "before" measures, only two used statistical analysis of the
data collected, and only one employed a control group for comparison. Sixteen respondents (34 percent) claimed to be measuring specific results desired, such as lower absenteeism, turnover, or costs, yet most were doing so in a subjective way (observation, interviews, etc.). The authors concluded that most organizations were measuring reaction to training programs; however, when considering the more important and difficult steps in the evaluation process, little was being done and many of these efforts were superficial and subjective.

A survey of 258 training officers in manufacturing, health, finance, retail, utility, insurance, and governmental organizations was conducted by Smeltzer (1979) in an effort to establish the type and quantity of evaluations that have been conducted. In the management/supervisory training category, he reported that 73 percent of the respondents were measuring trainee reaction to their programs, 82 percent were using learning measures of some type, 19 percent were measuring behavior samples, and 12 percent were attempting result specific evaluations.

Brandenburg (1982) found similar trends in his survey of training professionals. He reported the most-used evaluative techniques, regardless of criticisms heaped upon them, were the "smile" indices, such as reaction questionnaires and comments. Cognitive and performance-based outcome measures were used less often, while the least used
techniques were those that required longitudinal follow-up of participants.

These surveys served to emphasize the fact that, although training evaluation was obviously considered important by specialists in the field, over the last 15 years there have been little or no advances made in the quality of evaluations conducted. For whatever reason, the "quick-and-easy" type of evaluation was vastly predominant, although no one was providing research to support its validity as a measurement tool.

Alternative strategies have been proposed that were designed to circumvent some of the complexities that accompany the more strenuous management training evaluations. Bolar (1970) advocated a "company-based approach" which studied management training outcomes through the performance evaluation system that was employed in a particular organization. These performance rating forms could be compared on a before-and-after basis to ascertain what, if any, behaviors had changed since training was completed. He also pointed out that this approach could be helpful in highlighting weaknesses in the evaluation system currently being used by the organization.

Salinger and Deming (1982) discussed the use of the modified critical incidents approach. This method entailed talking with training participants, after they had been back on the job, to ask about behavioral changes and to
obtain specific examples of any change mentioned. These data could then be summarized across the group interviewed and analyzed to determine the impact that the training might have had, along with major obstacles the participants encountered in attempting to change their behavior.

Measuring shifts in attitudes or changes in values that had occurred as a result of management training was proposed by Mindak and Anderson (1971). They wrote that these attitudes might be reflected in how flexible an executive has become, or how sensitive he might be to other organizational roles, or possibly how close he perceives himself to be to other executives. This could be accomplished with the use of a questionnaire given at various points in time to the training participants.

The use of questionnaires to solicit both participants' and their superior(s)' perceptions of improved job performance after training was another proposed evaluation technique. This procedure attempted to achieve the kind of validity check that might have been available had a before-and-after scheme or a control group been used (Hodge, Anthony, & Swindle, 1976; Holder, 1972).

In reviewing the various evaluation techniques discussed previously, they all appeared to have at least a certain amount of face validity. Secondary variance, however, had been largely ignored and any conclusions drawn from data collected in such a manner would be
pseudo-scientific at best. As stated in the Handbook of Industrial and Organizational Psychology,

We take the strong position that attempts at training evaluation which are not experimental, do not have pre-post measures, and do not have adequate control groups, are not really evaluations (Hinrichs, 1976; p. 830).

McKinney (1957), in a discussion of various evaluation techniques, drew a similar conclusion regarding training program evaluation. In his opinion, the use of any evaluation technique other than a controlled experimental design would be a waste of the company's time and money. Using such a design as a basic criterion could quickly narrow down the number of truly valid evaluation methods; however, the issue of what to measure was still debatable.

Concerning management training evaluation, behavioral measures satisfied the conditions of a dependent variable in scientific research. Zenger and Hargis (1982) claimed that behavioral measures represented the best combination of rigor, relevance, and economy. Rigor was provided because behavior was a concrete event that could be measured fairly unambiguously. Relevance was assured since supervisory training had its usual target changes in actual on-the-job behavior. Also behavioral measures could be drawn using employee questionnaires which were relatively economical.
Collecting data on a training participant's behavior could be done by supplying the trainee's subordinates or superiors with questionnaires concerning specific behaviors stressed in the training programs. Zenger and Hargis (1982) pointed out that improving relationships with subordinates was the most reasonable objective of supervisory skill training, and they felt that surveying this population was the optimum means of gathering data.

Compounding the importance of collecting accurate, relevant evaluative data was the weak historical success of management training programs. Fielder (1967, 1973) concluded that empirical studies that had evaluated the effectiveness of various leadership, executive development, and supervisory training programs had not shown them to be very effective. The vast majority of research had failed to show that training was related to increased performance. As Fielder (1973) pointed out, leadership behavior was a combination of deeply ingrained personality factors and habits of interacting with others which could not be changed easily. It was highly unlikely that a few hours or days of training would have a significant impact over time.

These industry experts, and others, advocated training evaluations that combined features to form scientific, measurable, and feasible research designs for assessing management training results. A desirable research study in this area would be a comparison of reported training outcomes
contrasting results obtained using a carefully controlled scientific approach with those typically used as popular approaches. Would the figures differ, and if so, what would be the implications for the validity of various evaluation methods?

The present study proposed to compare several procedures currently being used for evaluation management training programs and to draw conclusions on the validity of results received. The design included the collection and analysis of data from a managerial training program on three different methods of training evaluation: a subjective, "immediate reaction" participant reaction form used in the training program, an objective assessment of leadership dimensions as self-rated by training participants, and a similar rating of the same dimensions by the participants' subordinates. In the latter two methodologies, pre and post-training samples were taken, and control groups comparable to the training participants (actually a "waiting list" of future trainees) were employed. It was hypothesized that, in measuring the effects of training on behavior as complex as supervision, the more strenuous the experimental design and the more objective the behavioral data collected, the less change would be attributable to the training program. Three specific hypotheses were proposed:

1. Course-ending participant rating sheets would appear the most favorable due to recency and leniency biases, and to professional format the program employed. This type of
evaluation could be as easily influenced by an "entertainment factor" as by course content (Thorely, 1972).

2. The self-report participant description would suggest a less favorable evaluation of the training program since it rated a behavior which, although the target of the training was more complexly determined. While this approach would be a much more strenuous evaluation of the course's effect, its reliance on a self-report methodology could easily be influenced by the trainee's expectancies and/or biases.

3. The subordinate questionnaire would generate the least favorable assessment of the same training. An experimental design combined with a fairly objective measure of behavior should provide a much more rigorous test for transfer of training. Since it was unlikely that a one-week program with no follow-up training could significantly change behaviors over an extended period of time, it was expected that this third methodology would present the clearest picture of the overall effectiveness of such a training program (i.e., few behaviorally-observable outcomes would be noticeable to outside parties).

Method

Subjects

The subjects were drawn from approximately 120 fire officers that were eligible to attend the Fire Officer Training Program conducted by the City of Fort Worth Fire
Department. These individuals represented the mid-management category within the fire department and were the target group for the training program. The initial sample consisted of 50 fire officers. However, due to nonparticipation, the final sample totaled 43, a response rate of 86 percent. All participants in the study were assigned to the operations division of the fire department. Half of the subject total was designated the control group and did not attend the training program until after the completion of this study. Experimental group subjects were selected at random, taking into consideration the fire department's available manpower and vacation schedule. Approximately equal numbers of supervisors were selected from each of three work shifts. The control group was selected in a similar random fashion and contained comparable percentages of captains and lieutenants as the experimental group. Subjects varied in age and experience, but the random selection helped to control group differences on these variables.

Survey participants were instructed that participation was to be confidential and was in no way going to be reported other than in general, summative results. A description of the research study's purpose and potential benefits was given to all participants in advance of data collection (these descriptions are in Appendices B and C).

**Instruments**

A 24-item course evaluation sheet, to which participants responded by means of a 4-point Likert scale, was developed
by the training program directors to measure participant reaction to the course at week's end. The items dealt with various presentations and topics discussed throughout the training program, but was not comprehensive in its coverage (see Appendix A).

Two 37-item questionnaires, to which participants responded by means of a 7-point Likert scale, were developed to measure participants' and their subordinates' perceptions and attitudes concerning supervision behavior as it related to the training program. A portion of the questionnaire was developed in a previous social research project (Organizational Assessments of the Effects of Civil Service, 1980) and dealt with general supervisory skills and behaviors expected to be affected by the training. The remainder of the items were constructed to reflect specific content stressed throughout the training program. It was felt that such a questionnaire would be comprehensive enough to detect behaviors that changed due to the transfer of training. The two questionnaires' item content was identical with the exception of wording to differentiate the subordinates' from the superiors' viewpoints (see Appendices B and C).

A sample of 190 of the subordinate questionnaires was collected during the training program's pilot study and factor analyzed for this population of subjects. Factors were identified using oblique factor analysis and their reliability was established using Cronbach's alpha (see Appendix B for
question/factor loadings). These three factors provided a means of measuring relevant behaviors before-and-after training: leadership skills, communication, and management style. The reliability coefficients of these three factors were $r = .69$, $r = .70$, and $r = .70$, respectively. These behaviors related both to general dimensions of supervisory behavior (Hodge, Anthony & Swindle, 1976) and to the focus of the training.

Additional data were collected from the participants including demographic fire department seniority and time served under the supervisor being rated. This information was used to "disqualify" survey respondents who had not had the opportunity to serve under an officer for a minimum (six months) of time.

**Procedure**

The study was conducted in the City of Fort Worth's Fire Department, which was conducting a fire officers' training program designed to provide its management personnel with systematic exposure to a variety of supervision and employee relation skills. This was a long-term project jointly administered by the Fire Department's Administration and the City of Fort Worth's Training Department. Classes were held in groups of approximately fifteen officers and at three week intervals.

The program format was perhaps a prototypical management development effort placing heavy emphasis on case studies,
role playing, video presentations and discussion. The program content dealt with a variety of topics which included: understanding individual differences, role responsibility, effective leadership styles, power of positive reinforcement, listening skills and discipline as a management tool. Overall, program emphasis was placed on developing leadership qualities, interaction management, and increased communication. The week-ending course evaluation sheet was virtually the sole means by which the training program was evaluated, a typical approach of most real-life training evaluation efforts.

The officer training program has been a rather innovative concept in a system that promotes strictly on the basis of civil service examinations. The vast majority of their management personnel had received no formal supervisory training, and those that had, did so independently of the department. The organization was willing to provide a large sample of officers who were receiving supervisory training because they were most interested in ascertaining the amount of training that was transferring to on-the-job behaviors.

Subjects comprising the experimental group were those attending the fire officer's training program. The training consisted of a five day, 40 hour program that was designed to provide systematic exposure to a variety of supervision and employee relation skills to the department's management personnel. Subjects comprising the control group were
approximately equal in number and rank to the experimental group, but were held out of training until the completion of the research project.

Data were collected using the three training evaluation techniques. A one-time course evaluation sheet was completed by the experimental group as the training was completed. This had been standard procedure throughout the previous programs and was administered by the training director.

Survey data were collected from the participants of two training program groups, and comparable size control groups, on two different occasions. The initial collection for both groups took place the week prior to the subjects' training and constituted the baseline information used for subsequent comparisons. The follow-up questionnaires were administered to both groups six weeks post-training. The survey design was as follows:

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Group</th>
<th>Before Observation</th>
<th>Treatment</th>
<th>After Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>R → X Y → 1</td>
<td>Experimental</td>
<td>W₁</td>
<td>X</td>
<td>Y₁</td>
</tr>
<tr>
<td>R → X Y → 2</td>
<td>Control</td>
<td>W₂</td>
<td>--</td>
<td>Y₂</td>
</tr>
</tbody>
</table>

R = Random Assignment  
X = Training Program  
W = Pre-Treatment Measure  
Y = Post-Training Measure

The third type of evaluation data collected was identical to the participant survey, but was administered to the subordinates of the members of both experimental and control groups. The time of the before-and-after measures was the same as the participant survey.
In all survey administrations, the fire department's in-house mail system was used to distribute and collect completed questionnaires. The questionnaires, with accompanying cover letter, were sent directly to the subordinates and subjects in both experimental and control groups. Envelopes were pre-addressed to the examiner and seals were provided to insure the protection of all relevant parties. The envelopes were held at the fire department's administrative offices where they were collected, along with the course rating sheets, by the examiner and later analyzed at North Texas State University.

Results

Information concerning demographic fire department seniority for the fire officers participating in the study is presented in Table 1. The experimental group is observed to be slightly more experienced than the control group on the three categories measured.

Table 2 contains fire department tenure, and time served under the supervisor being rated for the subordinate survey sample groups. Subordinates that had less than six months under the direct supervision of the officer to be rated were disqualified from the study. The experimental group is slightly more experienced, and has served slightly longer under their present supervisors, than has the control group.

Summaries of means and standard deviations of the three factor scores from the fire officer sample are presented in
Table 1
Descriptive Statistics for Fire Office Tenure

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fire Department Tenure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental Group</td>
<td>17.9</td>
<td>7.0</td>
<td>8-20</td>
</tr>
<tr>
<td>Control Group</td>
<td>16.3</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>17.1</td>
<td>6.7</td>
<td></td>
</tr>
<tr>
<td>Fire Officer Tenure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental Group</td>
<td>6.3</td>
<td>6.8</td>
<td>.5-20</td>
</tr>
<tr>
<td>Control Group</td>
<td>4.6</td>
<td>5.1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5.5</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td>Present Position Tenure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental Group</td>
<td>4.9</td>
<td>6.1</td>
<td>.5-20</td>
</tr>
<tr>
<td>Control Group</td>
<td>4.1</td>
<td>4.9</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4.5</td>
<td>5.5</td>
<td></td>
</tr>
</tbody>
</table>

Note. Experimental Group $n = 22$; Control Group $n = 21$; Total $n = 43$.

*Measured in years.*

Table 3. To determine the equivalence of the research groups before training and whether behaviors had been reported changed after training, analysis of variance (ANOVA) was performed on the participant survey data. A 2 x 2 repeated measures ANOVA design was employed on each of the three dependent measures from both sources of ratings.
Table 2
Descriptive Statistics for Subordinate Tenure

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fire Department Tenure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental Group</td>
<td>12.0</td>
<td>7.2</td>
<td></td>
</tr>
<tr>
<td>Control Group</td>
<td>10.7</td>
<td>7.4</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>11.4</td>
<td>7.3</td>
<td>1.0-29.0</td>
</tr>
<tr>
<td>Time Under Present Supervisor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental Group</td>
<td>2.6</td>
<td>2.9</td>
<td></td>
</tr>
<tr>
<td>Control Group</td>
<td>1.9</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2.3</td>
<td>2.4</td>
<td>.5-12.0</td>
</tr>
</tbody>
</table>

Note. Experimental Group _n_ = 39; Control Group _n_ = 34; Total _n_ = 73.

*aMeasured in years.

Neither the main effects nor the Group by Time interactions proved significant at the .05 level for any of the analyses. Both the experimental and control groups reported their behaviors to be at similar levels on the identified variables of leadership skills, communication, and management style before training program participation. Control group self-reports revealed stable performance on the variables across time. Training program participant self-reports on the variables disclosed negligible changes in one-the-job behavior over the six-week period.
Table 3
Statistical Summary of Factor Scores for the Fire Officer Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Leadership Skills</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretraining</td>
<td>98.7</td>
<td>6.9</td>
</tr>
<tr>
<td>Posttraining</td>
<td>98.1</td>
<td>9.7</td>
</tr>
<tr>
<td><strong>Communication</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretraining</td>
<td>42.4</td>
<td>6.1</td>
</tr>
<tr>
<td>Posttraining</td>
<td>42.5</td>
<td>5.7</td>
</tr>
<tr>
<td><strong>Management Style</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretraining</td>
<td>22.9</td>
<td>3.8</td>
</tr>
<tr>
<td>Posttraining</td>
<td>23.7</td>
<td>3.3</td>
</tr>
</tbody>
</table>

Note. Experimental Group n = 22; Control Group n = 21.

Statistical summaries of the three factors from the subordinate sample are presented in Table 4. An identical ANOVA was performed on the subordinates' survey data to test for pretraining equivalence and to identify any statistically significant changes by their supervisor on the three behavior variables. Neither the main effects nor the Group by Time interaction proved significant at the .05 level for any of the analyses. All supervisors' behaviors, before training program participation, were judged to be at approximately equivalent levels by the subordinate sample groups.
Control group ratings remained consistent across time. Experimental group ratings also showed no statistically significant changes on any behavior variable when reevaluated six weeks after their respective supervisors had completed the training program.

Table 4
Statistical Summary of Factor Scores for the Subordinate Sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Leadership Skills</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretraining</td>
<td>88.3</td>
<td>22.9</td>
</tr>
<tr>
<td>Posttraining</td>
<td>88.9</td>
<td>18.6</td>
</tr>
<tr>
<td>Communication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretraining</td>
<td>37.5</td>
<td>11.4</td>
</tr>
<tr>
<td>Posttraining</td>
<td>37.7</td>
<td>10.7</td>
</tr>
<tr>
<td>Management Style</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretraining</td>
<td>20.1</td>
<td>6.8</td>
</tr>
<tr>
<td>Posttraining</td>
<td>20.0</td>
<td>5.6</td>
</tr>
</tbody>
</table>

Note. Experimental Group n = 26; Control Group n = 35.

As part of standard procedure throughout all of the officer training programs, a week-ending participant rating sheet was administered by the training directors. Fire Department administrators use these evaluation forms to determine the "success" or "failure" of each training class.
At the close of the training week, the evaluation forms were completed by the participating officers, and the experimental groups were separated and analyzed for content. Numerical values were assigned to the Likert scale (4 = Strongly Agree, 3 = Agree, 2 = Disagree, 1 = Strongly Disagree) tabulated, and averaged to obtain an overall course rating. The course ratings for the 25 participants ranged from 2.88 to 3.5, and the group mean was 3.1 with a standard deviation of .19. As a group, the participating officers "more than" agreed that the training program was helpful, well presented, and transferable back to their jobs (see Appendix A).

Discussion

The results of the present study supported the general hypothesis that when measuring the effects of training on behavior as complex as supervision, the more rigorous the experimental design and the more objective the behavioral data collected, the less favorable the training program results will appear. Three separate procedures were employed to evaluate the same management training program, yet the assessments collected differed in their implications. A more detailed examination of the results is necessary in order to determine the conclusions to be drawn and the limitations of the present study.

The subjective, "immediate reaction" participant rating form has been the sole means of evaluating the fire officer training program up to the present time. It was hypothesized
that this evaluation method would appear the most favorable due to the recency and leniency biases, and to the professional format that the program employed. Analysis of the experimental group's ratings revealed an overall score of 3.14 on a possible scale of 4.00. The training program format with its use of case studies, role playing situations, and video presentations appeared to leave a very favorable impression on the participant. Whether or not their evaluations were influenced by the presentation style more than by its applicability is difficult to determine. The "on-the-spot" method of evaluation is by far the most popular format used by training professionals (Brandenburg, 1982), yet is negligent of attempting to control for biases or other forms of error; its scientific validity is virtually non-existent. Nevertheless, it is extensively used as "support" for continuing training programs, and in the present research casts an initial positive light on the program's effectiveness. Closer, more tightly controlled evaluations apparently contradict this course-ending rating score collected from the participants.

It was hypothesized that the self-report participant survey would suggest a less favorable evaluation of the training program since it rates complex behaviors, and this appears to have been an accurate prediction. The survey revealed that targeted behaviors were not improved posttraining regardless of initial training reactions. These results seem to imply that the training had no effect on the officer's on-the-job behavior
even after they had sufficient time (six weeks) to put their learning into practice. The behavior variables (leadership skills, communication, and management style) were virtually unchanged within both the experimental and control groups.

It must be pointed out that on one of the variables, leadership skills, the preratings were at such an elevated level that the range of possible improvement was restricted. The experimental group rated themselves as performing at an average of 5.8 efficiency (7.0 possible) on the 17 questions pertaining to leadership skills. The control group was only slightly less optimistic, rating themselves at a 5.7 level. However, a "ceiling effect" on the pre-test was not observed on the other two behavior variables. The two groups combined rated their communication efficiency to be at a 4.6 level and their management style to be at 4.7 before training. Even though these lower starting levels provided more room for improvement, there were still no reported changes in behaviors. It would appear that regardless of the proficiency before training, the participants were unaware of any "improvements" in their supervisory behaviors and reported it as such.

The anticipated influence that the trainees' expectancies and/or biases may have had on their self-reports did not materialize. As Mezoff (1981) indicates, many training program participants are inclined to report what they think is expected of them rather than reporting their actual behaviors. With no reported changes occurring on any of the variables, the
participants appeared to have been more "objective" in their appraisals than might have been expected considering their recent experiences.

A final hypothesis was that the subordinate questionnaire would generate the least favorable assessment of training program effectiveness. However, these results were identical to the participant's self-report survey in revealing no significant changes on any of the behavior variables measured. These combined results imply that there was virtually no transfer of training from the week long program to on-the-job performance. As pointed out in the initial research hypothesis, it is unlikely that a one-week training program with no follow-up could significantly change behaviors over an extended period of time. The present study's results support these earlier assumptions.

The question of which survey method may provide the most accurate portrayal of reality must be addressed. Although neither survey sample (supervisors or subordinates) reported any significant changes on the measured behaviors, there were differences observed on the stated levels of proficiency reported by the two groups. Supervisors as a group rated themselves higher on all three behavior variables in the pre-test than did their subordinates on those same variables. Management style was the area of strongest disagreement with the supervisors rating themselves at 4.70 average and their subordinates placing them at the 4.15 level. Leadership skill
ratings were the next most discrepant with an overall 5.75 self-rating from the supervisors and a 5.35 average from their subordinates. The behavior variable that the two groups came closest to agreeing on was communication; supervisors rating themselves at 4.6 and their subordinates ratings being at a 4.45 level.

Because most of the subordinate sample data contained multiple ratings of the appropriate supervisor, the reliability should be higher in this group's ratings. As a group, the supervisors saw themselves performing at a higher level than did their subordinates (implying leniency), which coupled with the greater reliability in the subordinate ratings, would seem to suggest using nonparticipant rating as the more "objective" measure of behavior. This would be consistent with the research hypotheses and is also compatible with Zenger and Hargis' (1982) recommendations for using nonparticipant raters. As they point out, improving subordinate relations is the objective of most supervisory skill training in the first place.

The three evaluation method results closely followed the hypothesized outcomes as they related to the training program's effectiveness. The course-ending rating sheet appeared most favorable, while the survey design results were less favorable. The subordinate survey seemed to be the most "objective," and therefore should present the clearest picture of the training program results as measured by the officer's on-the-job behavior.
The present study is unique in that it has allowed the results obtained using carefully controlled evaluation designs to be compared against results from a casual but a more popular approach. The findings appear discrepant. The training program participants gave the program and its content a positive rating upon completion of the week-long course. They were implying to the training directors that they felt the program was a "success" (or at least entertaining). In most management training programs this would be all the evaluative data collected (Brandenburg, 1982). Considering the overall response from the participants, the program would appear to have accomplished its purpose. But when a scientific evaluation approach was employed (using either participant self-reports or subordinate ratings) and secondary variance was controlled for (before-after two-group design), the same training program apparently had no observable effects on targeted on-the-job behaviors. It is safe to say that most training directors would not consider these results as positive. The present research would seem to provide disqualifying data on the "quick-and-easy" type of evaluation as a valid training measurement tool.

The findings that the management training program in question had little effect are consistent with those cited by Fiedler (1973). He discussed a research study conducted by Newport who, after surveying 121 large companies concluded that not one of the companies had obtained scientifically acceptable evidence that the leadership training for
their middle management had actually improved performance. The present study’s results are supportive of this disturbing trend and suggest that the efficacy of such training programs is questionable at best.

Controlled experimental designs coupled with behavioral measures provide a rigorous test for transfer of training. They are "evaluations" in the strictest sense of the word (Hinrichs, 1976; McKinney, 1957), and provide the participating organization with invaluable information from which they can redesign future programs or simply reevaluate training outcome expectancies. Fear of negative results can no longer be accepted as rationale for not doing well-controlled evaluations. Many programs will discover that expected results are not being realized (as in the present study), but until this knowledge is discovered unequivocally, there will be no advancements towards possible alternatives in the field. Research should begin on new training techniques and learning programs, and training directors must be held accountable if present methodologies are ineffective. There is far too much at stake, enormous amounts of time and money invested, for business and industry to continue to avoid conducting proper training evaluations. The knowledge and expertise is available to generate competent, accurate training evaluations and to then use this information to build powerful, effective programs. Management training evaluations, when performed properly, can become the cornerstone of successful, profitable organizations.
Limitations of the Present Research

One possible shortcoming of the survey employed in this study is that it was not totally specific to the training program under evaluation. A portion of the questionnaire was developed for a previous social research project. It was felt, however, that the questionnaire was comprehensive enough to detect behaviors that change due to the transfer of training since it dealt with both general supervisory skills and behaviors, as well as specific content stressed throughout the training program. Ideally, a survey used for a project of this type would be developed entirely from the training program's materials and content to insure complete relevance to all targeted behaviors. The cost of developing a "unique" survey for every situation would likely be prohibitive for most organizations, but nevertheless would be the theoretically proper procedure.

Another shortcoming of the present study was not being able to statistically compare the course-ending evaluation form to the two surveys. Different formats, scales, and question content precluded more precise analyses. A similar Likert scale would at least provide an opportunity for rating "average" comparisons to be made. In the present study, the course-ending form was already in use before the research data were collected and no alterations were possible.

A final possible limitation was the assumption that six weeks was sufficient time for the newly trained supervisors
to "put into action" what they had been taught during the training program. Behaviors as complex as supervision and leadership are slow to develop (Fiedler, 1973). Longitudinal follow-ups are extremely important in tracking the evolution of such behaviors, and should be conducted at various points in time as experience and opportunities expand.

**Directions for Future Research**

A review of the research literature in the training field reveals very little "scientific" research being done on management training program long-term effectiveness. An increase in longitudinal type studies conducted after management training could concentrate on the costs of ineffective evaluations. Perhaps if long range information was published on the amount of money and time lost on ineffective training programs that had used a "quick-and-easy" (but inaccurate) evaluation format, as opposed to a similar sized organization employing a well-controlled, scientific format, there would be an impact felt in the field.

Unfortunately, well-controlled scientifically sound evaluations are not feasible in many applied settings. Many smaller organizations either don't have the sample size necessary or the adequate time frame that it takes to conduct a thorough study. A format or sampling system that could improve the "lag time" in evaluation studies, while providing the same confidence in the results, would be extremely beneficial to the field.
Specific questionnaire items that were identified in this sample to measure leadership skills, communication, and management style could possibly be used in future research. These identified items could be given emphasis both in training programs and in subsequent evaluation efforts. The items (and factors) appear to measure different aspects of general supervisory behavior, and could provide a nucleus for future training programs in similar para-military environments.

Future researchers will also want to address more closely the feasibility and "success rate" of teaching behavior as complexly determined as that of leadership/managerial skills. Possibly this avenue is nothing but a "dead end" and emphasis should be placed instead on developing assessment techniques for identifying individuals who possess the personality traits and abilities that combine to make an inherently good manager.

In summary, the present study has shown that the type of training evaluation format employed can greatly influence evaluation outcomes, especially when training is aimed at complex behaviors such as leadership. Research in this area has implied that long-term behavior change may not be a feasible expectation of training. Accurate evaluations of training programs are needed to draw conclusions on the cost effectiveness of this approach to changing behavior, and to generate additional research on possible alternatives.
Appendix A

FIRE OFFICER TRAINING PROGRAM

COURSE EVALUATION

Please respond to each of the statements below about course content and activities using the following scale:

SA = Strongly Agree  D = Disagree
A = Agree            SD = Strongly Disagree

1. I clearly understood the theory about different "types" of Personality Styles (i.e., sensing-intuitive, thinking-feeling, etc.) and related examples discussed. SA A D SD

2. The individual profile I received using the Myers-Briggs Type Indicator was an accurate indicator of my own decision-making style. SA A D SD

3. The discussion of the different types of Personality Styles helped me to understand more about how I am different from others. SA A D SD

4. Understanding of these differences will help me in dealing with my subordinates, peers, and superiors. SA A D SD

5. The concept of "type A" and "type B" behaviors and their implications on health and stress are applicable to my occupation. SA A D SD

6. The classification of myself as Type A or Type B was consistent with how I view myself. SA A D SD

7. Understanding my values and how they differ from other generations' values is important to me personally. SA A S SD

8. Understanding my values and how they differ from other generations' values is relevant to my development as a Fire Officer. SA A D SD

9. The Morris Massey presentation and related exercises have helped me to become more understanding of others. SA A D SD

10. I would recommend the Massey presentation as a standard part of future development programs. SA A D SD
11. I found the case studies related to the LEAD Instrument helpful in understanding diverse styles of leadership. SA A D SD

12. The information I received from the LEAD Instrument has helped me to compare my own leadership style to alternative, possibly more effective styles. SA A D SD

13. This program dealt with skills which are relevant to my current management situations and problems. SA A D SD

14. I will find it too difficult to actually apply the skills and techniques which we practiced. SA A D SD

15. The handouts and exercises were clearly written and illustrated points or enhanced learning for me. SA A D SD

16. The trainers were highly effective in their roles. SA A D SD

17. My communication techniques will change very little as a result of this program. SA A D SD

18. The skill practices helped my understanding of the key principles and critical steps. SA A D SD

19. I feel better equipped to gain employee commitment as a result of this program. SA A D SD

20. The critical steps will help me plan more productive performance discussions. SA A D SD

21. The program was well worth the time I invested in it. SA A D SD

22. The trainers did not really use the skills themselves during the workshop. SA A D SD

23. The skill practices are where the real learning took place. SA A D SD

24. The action plan will help me actually transfer and use the skills on the job. SA A D SD
Appendix B

Dear Fire Officer:

As you know, the Fort Worth's Fire Department Administration has initiated a management training program with the intent of teaching fire officer personnel better leadership and employee relations skills. The administration has asked me to conduct a research study to evaluate the effectiveness of this program. As part of this evaluation, you are asked to complete the attached survey. It is constructed to find out how you feel about the type of supervision you are now providing. Your answers, along with those of other officers, will help us understand the current approaches used in the Fire Department, and will provide us a basis for comparison at a later date.

If this questionnaire is to be useful, we need frank and honest answers to the questions. We are interested in the way you think and feel about your work situation and your supervisory skills.

Your answers to these questions are to be completely confidential. All surveys will be taken to North Texas State University for analysis and safekeeping. Summaries of the results will combine the responses of many different officers and no one other than myself will have access to your individual responses, so candor is encouraged.

To determine whether change has occurred, we will be asking you to fill out another questionnaire in approximately one month. This comparison will provide the department with valuable information for making any needed changes in the training program. Improved management techniques and improved service for the City of Fort Worth are the goals of this training and evaluation.

After you complete the survey, please place it in the self-addressed envelope, attach seal to the closure of envelope, and return it through the department's in-house mail service as soon as possible.

Thank you in advance for your cooperation and assistance. If you have any questions concerning this project that have not been answered, please feel free to contact me at the number provided below.

John Hale
Graduate Student,
Department of Psychology
North Texas State University
Appendix B—Continued

Name ____________________________
Length of time with Fort Worth Fire Department ________________
Length of time as a fire officer ________________
Length of time in present position ________________

Use the following scale to make your responses. Circle the appropriate number to the right of the questions:

1 = Strongly Disagree
2 = Disagree
3 = Slightly Disagree
4 = Neither Agree nor Disagree
5 = Slightly Agree
6 = Agree
7 = Strongly Agree

As a Supervisor:

1. (S) . . . I reward subordinates for good performance. 1 2 3 4 5 6 7

2. (S) . . . I am someone they can trust. 1 2 3 4 5 6 7

3. . . . I insist that they work hard. 1 2 3 4 5 6 7

4. (S) . . . I am concerned about each subordinate as a person. 1 2 3 4 5 6 7

5. (S) . . . I encourage them to participate in making important decisions. 1 2 3 4 5 6 7

6. (S) . . . I admit when I make a mistake. 1 2 3 4 5 6 7

7. . . . I demand high quality work of subordinate. 1 2 3 4 5 6 7

8. (S) . . . I handle the administrative parts of my job extremely well. 1 2 3 4 5 6 7

9. (S) . . . I make sure that each subordinate knows what has to be done. 1 2 3 4 5 6 7

10. . . . I am simply following orders most of the time. 1 2 3 4 5 6 7

11. (S) . . . I help subordinates discover problems before they get too bad. 1 2 3 4 5 6 7
Appendix B--Continued

12. . . .I criticize those who perform poorly.

13. (S). . .I encourage them to speak up when they disagree with a decision.

14. (S). . .I help poor performers learn to do their job better.

15. (S). . .I blame subordinates when things go wrong.

16. (S). . .I give subordinates opportunities for training and development.

17. (M). . .Top level managers here encourage subordinates to speak up when they disagree with a decision.

18. (C). . .Top level managers make most important decisions without adequate input from lower level employees.

19. (M). . .People in other parts of this organization try to understand the problems my subordinates face on the job.

20. . . .This would be a better place to work if people helped each other in solving problems.

21. (C). . .There is a lot of conflict among people here.

22. (C). . .People here will do things behind you back.

23. (C). . .Around here it's important to protect yourself or you'll be blamed for problems.

24. (C). . .People are afraid to express their real views to top management.

25. (C). . .I feel I give too much supervision on the job.

26. (C). . .I feel the supervision I do give is poor.
27. (M). . .When problems come up on the job, 1 2 3 4 5 6 7
give a great deal of help in
handling those problems.

28. (M). . .When changes are being discussed 1 2 3 4 5 6 7
which will affect my subordinates' jobs, I usually consult them before
the changes are made.

29. (C). . .The planning, organizing, and 1 2 3 4 5 6 7
scheduling of the work in my unit
is handled poorly.

30. (C). . .Both I and those above me under-
stand the technical problems my
subordinates face in doing their
work.

31. (S). . .My subordinates feel free to talk 1 2 3 4 5 6 7
over job problems with me.

32. (M). . .My subordinates feel free to talk 1 2 3 4 5 6 7
over job problems with those above
me.

33. (S). . .I make an effort to listen and to 1 2 3 4 5 6 7
understand my subordinates' points
of view.

34. (S). . .I give the impression of being 1 2 3 4 5 6 7
more concerned with what my subor-
dinates do wrong rather than what
they do right.

35. (S). . .I treat those of my own generation 1 2 3 4 5 6 7
more fairly than I treat others.

36. (S). . .My subordinates have no say in 1 2 3 4 5 6 7
developing new work rules and
procedures.

37. . . .My subordinates have a good deal 1 2 3 4 5 6 7
to say in rating my performance
as a supervisor

Factor Loading Identification

(S) = Leadership Skills; (C) = Communication; (M) = Management Style
Dear Fire Department Employee:

As you may know, the Fort Worth's Fire Department Administration has initiated a management training program with the intent of teaching fire officer personnel better leadership and employee relations skills. The administration has asked me to conduct a research study to evaluate the effectiveness of this program. As part of this evaluation, you are asked to complete the attached survey. It is constructed to find out how you and others feel about the type of supervision you now receive. Your answers, along with those of other employees, will help us understand the current approaches to supervision used in the Fire Department, and will provide us a basis for comparison at a later date.

If this questionnaire is to be useful, we need frank and honest answers to the questions. We are interested in the way you think and feel about your work situation and the supervision you are receiving.

Your answers to these questions are to be completely confidential. All surveys will be taken to North Texas State University for analysis and safekeeping. Summaries of the results will combine the responses of many different employees and no one other than myself will have access to your individual responses. The supervisor you are evaluating will never see your answers, so candor is encouraged.

To determine whether change has occurred, we will be asking you to fill out another questionnaire in approximately one month. This comparison will provide the department with valuable information for making any needed changes in the training program. Improved management for your units and improved service for the City of Fort Worth are the goals of this training and evaluation.

After you complete the survey, please place it in the self-addressed envelope, attach the seal to closure of envelope, and return it through the department's in-house mail service as soon as possible.

Thank you in advance for your cooperation and assistance. If you have any questions concerning this project that have not been answered, please feel free to contact me at the number provided below.

John Hale
Graduate Student, Department of Psychology
North Texas State University (214) 484-1068
Appendix C--Continued

Name ____________________________

Supervisor I am rating ____________________________

Length of time with Fort Worth Fire Department ____________________________

Length of time under supervisor being rated ____________________________

Use the following scale to mark your responses. Circle the appropriate number to the right of the questions:

1 = Strongly Disagree
2 = Disagree
3 = Slightly Disagree
4 = Neither Agree nor Disagree
5 = Slightly Agree
6 = Agree
7 = Strongly Agree

My Supervisor:

1. . . .rewards me for good performance. 1 2 3 4 5 6 7
2. . . .is someone I can trust. 1 2 3 4 5 6 7
3. . . .insists that I work hard. 1 2 3 4 5 6 7
4. . . .is concerned about me as a person. 1 2 3 4 5 6 7
5. . . .encourages me to participate in making important decisions. 1 2 3 4 5 6 7
6. . . .admits when he/she makes a mistake. 1 2 3 4 5 6 7
7. . . .demands that I do high quality work. 1 2 3 4 5 6 7
8. . . .handles the administrative parts of his/her job extremely well. 1 2 3 4 5 6 7
9. . . .makes sure I know what has to be done. 1 2 3 4 5 6 7
10. . . .is simply following orders most of the day. 1 2 3 4 5 6 7
11. . . .helps me discover problems before they get too bad. 1 2 3 4 5 6 7
12. . . .criticizes people who perform poorly. 1 2 3 4 5 6 7
13. . . .encourages me to speak up when I disagree with a decision.
14. . . .helps poor performers learn to do their job better.
15. . . .blames subordinates when things go wrong.
16. . . .gives me opportunities for training and development.
17. . . .Top level managers make most important decisions without adequate input from lower level employees.
18. . . .Top level managers here encourage subordinates to speak up when they disagree with a decision.
19. . . .People in other parts of this organization try to understand the problems I face on my job.
20. . . .This would be a better place to work if people helped each other in solving problems.
21. . . .There is a lot of conflict among people here.
22. . . .People here will do things behind your back.
23. . . .Around here it's important to protect yourself or you'll be blamed for problems.
24. . . .People are afraid to express their real views to top management.
25. . . .I receive too much supervision on my job.
26. . . .The supervision I do receive is poor.
27. . . .When problems come up in doing my job, I receive a great deal of help in handling those problems.
28. When changes are being discussed which will affect my job, I am usually consulted before the changes are made.

29. The planning, organizing, and scheduling of the work on my shift is handled poorly.

30. My supervisor and those above him understand the technical problems I face in doing my work.

31. I feel free to talk over job problems with my supervisor.

32. I feel free to talk over job problems with the person above my supervisor.

33. My supervisor makes an effort to listen and to understand my point of view.

34. My supervisor seems more concerned with what I do wrong rather than what I do right.

35. My supervisor treats those of his own generation more fairly than he treats others.

36. I have no say in developing new rules and procedures.

37. I have a good deal of say in rating the performance of my supervisor.
References


Smeltzer, L. R. (1979). Do you really evaluate, or just talk about it? *Training, 16* (8), 6-8.
