

Summer 1987

## Factors Affecting the Efficacy of Rater Training: Information Type and Mode of Presentation

Todd A. Silverhart  
*Old Dominion University*

Follow this and additional works at: [https://digitalcommons.odu.edu/psychology\\_etds](https://digitalcommons.odu.edu/psychology_etds)



Part of the [Industrial and Organizational Psychology Commons](#)

---

### Recommended Citation

Silverhart, Todd A.. "Factors Affecting the Efficacy of Rater Training: Information Type and Mode of Presentation" (1987). Doctor of Philosophy (PhD), Dissertation, Psychology, Old Dominion University, DOI: 10.25777/xggd-d792  
[https://digitalcommons.odu.edu/psychology\\_etds/312](https://digitalcommons.odu.edu/psychology_etds/312)

This Dissertation is brought to you for free and open access by the Psychology at ODU Digital Commons. It has been accepted for inclusion in Psychology Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact [digitalcommons@odu.edu](mailto:digitalcommons@odu.edu).

FACTORS AFFECTING THE EFFICACY OF RATER TRAINING:  
INFORMATION TYPE AND MODE OF PRESENTATION

By

TODD A. SILVERHART

B.A. MAY 1981, MIDDLEBURY COLLEGE

M.S. AUGUST 1984, OLD DOMINION UNIVERSITY

A Dissertation Submitted to the Faculty of  
Old Dominion University in Partial Fulfillment of the  
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

PSYCHOLOGY

Old Dominion University

August, 1987

Approved by:

Terry L. Dickinson (Chair)

Glyn D. Coates

Robert M. McIntyre

William H. Silverman

## DEDICATION

This work is dedicated  
to the memory of my brother,  
Peter Andrew Silverhart

In many ways Peter is responsible for both the beginning and the end of this achievement. The beginning, since his interests in psychological research first attracted me to this discipline. The end, since he was so influential in my developing the characteristics of diligence and perseverance that were critical in bringing this project to completion. I hope I would have made him proud.

## ACKNOWLEDGEMENTS

There are several people whose roles have been instrumental in the development that I have experienced over the last five years, and whose support has been indescribably important to me. To all of these people, I offer my sincere appreciation.

There are certain individuals that require special acknowledgement. The first of these is Dr. Terry Dickinson. I have had the distinct privilege of working with Terry since my first year of graduate school. During that time, I have been fortunate enough to experience a special mentoring relationship. The word mentor is used with the utmost respect and truly describes the role that Terry has played in my development: wise adviser and loyal friend. I will never be able to thank him enough for his guidance, his support, and all that I have learned from him.

Special acknowledgement is also in order for the other members of my dissertation committee: Drs. Coates, McIntyre, and Silverman. All of them have contributed to the quality of this project and have provided support from the early stages. I thank them for that, and for accommodating last minute drafts so that I could achieve my goal of August graduation. In addition, all of them have contributed towards my development as an industrial/organizational psychologist. I would like to express special appreciation to Dr. Coates for his patience and assistance with numerous statistical problems over the years; to Bob McIntyre for his enthusiasm, expertise and eye for detail, and to Bill Silverman for my experiences at Giant Food and an opportunity to gain an insight on life in the business world.

Drs. Dalessio, Davis, Glickman, Kirby, and Morgan also all

deserve mention. Each of them have contributed to my development in their own special way, and as a result of them I am a far better psychologist.

There is not enough that can be said to my colleague, Dolph Johnson. Dolph's encouragement, companionship, and good humor have made the difficult times much easier to deal with. The experiences that we have shared over the course of the last year have bonded a friendship that is certain to endure.

Finally, I would like to acknowledge my parents. It is as a result of their never ending love, support, understanding, and concern that I am the person that I am today. It was from them that I learned the value of being the best that I can be. Now its time to move on, and apply that value to the challenges of a professional career.

## TABLE OF CONTENTS

	<u>Page</u>
LIST OF TABLES.....	viii
LIST OF FIGURES.....	ix
 Chapter	
I. INTRODUCTION.....	1
Rater Training.....	3
Review of training methods.....	4
Review of practice and feedback studies.....	5
Multiple Cue Probability Learning.....	10
Implications of MCPL for Rater Training.....	13
Research Hypotheses.....	17
Hypothesis 1.....	17
Hypothesis 2.....	18
II. METHOD.....	19
Participants.....	19
Stimulus Materials.....	19
Rating Forms.....	20
Design.....	20
Information type.....	21
Mode of information presentation.....	21
Control conditions.....	23
Target Score Development.....	23
Procedures.....	26
Evaluation of Dimension Training.....	28
Accuracy Measurement.....	29

Measures of Correlational Accuracy.....	29
Correlational accuracy per ratee.....	29
Correlational accuracy per dimension.....	29
Cronbach's Accuracy Statistics.....	30
Elevation.....	30
Differential elevation.....	30
Stereotype accuracy.....	30
Differential accuracy.....	30
Extended Accuracy Design.....	31
III. RESULTS.....	40
Evaluation of Dimension Training.....	40
Evaluation of Information Type and Mode.....	41
Correlational Accuracy.....	42
Correlational accuracy per ratee.....	42
Correlational accuracy per dimension.....	47
Cronbach's Accuracy Measures.....	50
Extended Accuracy Design.....	57
Evaluation of the Post-Experimental Questionnaire.....	62
IV. DISCUSSION.....	64
Information Type.....	64
Mode of Information Presentation.....	67
Factors Contributing to the Results.....	69
Analytical Approaches to Measuring Accuracy.....	71
Practical Implications.....	73
Summary and Conclusions.....	73
REFERENCES.....	75
APPENDIX A: Script for Interview Simulations.....	84

APPENDIX B: Definitions of the Performance Dimensions.....	123
APPENDIX C: The BARS Rating Forms.....	125
APPENDIX D: Checklist of Behaviors used by Expert Raters.....	129
APPENDIX E: Script for Behavioral Rationale Lecture.....	135
APPENDIX F: Target Scores for the Training and Experimental Interview Simulations.....	143
APPENDIX G: Informed Consent Form.....	145
APPENDIX H: Script of Videotaped Presentation.....	147
APPENDIX I: Background Information for the Interview Simulation Exercise.....	157
APPENDIX J: Script for Training Role Play.....	159
APPENDIX K: Post-Experimental Questionnaire.....	166
APPENDIX L: Tests to Evaluate General Dimension Training.....	169
APPENDIX M: Computational Formulas for Cronbach's (1955) Accuracy Measures.....	179



## LIST OF TABLES

<u>Table</u>	<u>Page</u>
1. Summary of the Analysis of Variance for the Psychometric Properties of the Target Scores .....	25
2. Psychometric Interpretations of the Basic Accuracy design.....	32
3. Psychometric Interpretations of the Extended Accuracy Design...	34
4. Error Terms for the Extended Accuracy Design.....	37
5. Means and Standard Deviations for the Correlational Accuracy per Ratee Measure.....	44
6. Summary of the Analysis of Variance for the Correlational Accuracy per Ratee Measure.....	45
7. Means and Standard Deviations for the Seven Ratees.....	46
8. Means and Standard Deviations for the Correlational Accuracy per Dimension Measure.....	48
9. Summary of the Analysis of Variance for the Correlational Accuracy per Dimension Measure.....	49
10. Means and Standard Deviations for Cronbach's Accuracy Statistics.....	51
11. Summary of the Analysis of Variance for Cronbach's Elevation Measure.....	53
12. Summary of the Analysis of Variance for Cronbach's Differential Elevation Measure.....	54
13. Summary of the Analysis of Variance for Cronbach's Stereotype Accuracy Measure.....	55
14. Summary of the Analysis of Variance for Cronbach's Differential Accuracy Measure.....	56
15. Summary of the Analysis of Variance for the Extended Accuracy Design.....	58

## LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1. Brunswick's (1943) Lens Model.....	11

## ABSTRACT

### FACTORS AFFECTING THE EFFICACY OF RATER TRAINING: INFORMATION TYPE AND MODE OF PRESENTATION

Todd A. Silverhart  
Old Dominion University  
Director: Dr. Terry L. Dickinson

In order to enhance the quality of performance ratings, researchers have directed their efforts towards training raters to evaluate performance more accurately. The purpose of the present study was to examine two factors that may affect the efficacy of rater training for improving the accuracy of performance ratings. One factor was the type of information that was presented during training (target score information, behavioral rationale for target scores, or a combination of target score and behavioral rationale). The second factor was the mode in which information was presented during training (feedback or feedforward). In addition to assessing the unique contribution that various types of information contribute to the success of rater training programs, the present study tested two hypotheses based on generalizing the multiple cue probability learning (MCPL) literature to the task of rating performance. The first hypothesis was that rater training that incorporates target score information, or combines target score information with a behavioral rationale for the expert ratings will result in less accurate performance ratings than rater training incorporating only the behavioral rationale. The second hypothesis was that performance ratings will be more accurate when raters receive training information by means of feedforward than when control training is provided in

which the training information is not presented.

A hundred and one undergraduate and graduate students served as participants in the study. These participants were randomly assigned either to one of six experimental conditions formed by crossing three levels of information type with two levels of the mode in which training information was presented, or to one of two training control conditions. Ratings were made of the videotaped performance of seven individuals conducting simulated performance evaluation interviews. The performance ratings were analyzed with correlational measures of accuracy, Cronbach's (1955) accuracy statistics, and Dickinson's (1987) extended accuracy design. The results of these analyses generally did not find the training to be effective. Support was not found for either hypothesis, although some findings did indicate that feedback was more effective than feedforward. The results are discussed in terms of differences between the MCPL paradigm and the task of performance rating. In addition, a number of possible explanations for the findings from the study are presented.

-

FACTORS AFFECTING THE EFFICACY OF RATER TRAINING:  
INFORMATION TYPE AND MODE OF PRESENTATION

CHAPTER ONE

Introduction

The measurement of performance has long been an area of interest to industrial/organizational psychologists (Landy & Farr, 1980). This interest can be attributed to the importance of the role that performance assessment plays in the functioning of a work organization, as well as the difficulties inherent in measuring performance accurately.

The importance of performance measurement stems from the reliance of personnel decisions on the accurate description of employee performance. These decisions determine administrative, developmental, and research programs for an organization (Landy & Farr, 1983).

A variety of possible sources of performance information exist in the workplace. These sources of information consist of objective measures (e.g., number of units produced or sales volume), personnel measures (e.g., absence record or rate of advancement), and judgmental measures (e.g., performance ratings). Although the ideal measurement method would combine objective, personnel, and judgmental sources of information (Landy & Trumbo, 1980), judgmental data are typically relied upon because of problems with the availability, reliability, and confounding of the other sources. Despite a reliance by decision makers in work organizations on judgmental measures, difficulties in

accurately describing performance result from intentional and inadvertent rating biases (Landy & Farr, 1980).

In order to enhance the quality of performance ratings, researchers have directed their efforts towards developing improved rating formats, and training raters to evaluate more accurately (Pulakos, 1986; Smith, 1986). Regardless of years of research, concentration on measurement instruments has not resulted in improvements to performance evaluations (Landy & Farr, 1980). Rater training, however, has had notable success for improving the quality of performance ratings (Smith, 1986).

Rater training is the focus of the present study. More specifically, the purpose is to examine two factors that may affect the efficacy of rater training for improving the accuracy of performance ratings. One factor that was examined was the type of information that is presented in rater training programs. Current rater training programs have combined target score information with the presentation of information about behaviors that are relevant to the behavioral dimensions being evaluated (i.e., a behavioral rationale for the ratings). The presentation of both types of information within the same training programs has not allowed for determination of the unique contributions that the different types of information make towards improving the accuracy of ratings. In order to have an understanding of why a training strategy is or is not successful, it is important to examine systematically the components that are involved in the strategy (Spool, 1978). In addition, findings from the multiple cue probability learning (MCPL) literature suggest that outcome information (i.e., target scores) may have an

adverse effect on the learning of complex judgment tasks. By manipulating the type of information presented during training, the generalizability of these MCPL findings to the task of rating performance was examined.

The second factor that was investigated in this study was also based on results from the MCPL literature. Although current training programs typically present information about the rating task in the form of feedback, studies using the MCPL paradigm suggest that when performing complex judgment tasks, the use of feedforward (i.e., presenting the information before the task has been performed) may be an effective training strategy. The present study investigated the effectiveness of providing training information in the form of feedforward.

#### Rater Training

Historically, rater training studies have been concerned with eliminating psychometric biases from ratings. For example, raters have been trained to change the shape of their rating distributions to avoid biases such as central tendency, leniency/severity, or halo (Latham, Wexley, & Pursell, 1975). This training was based on the assumption that reducing rating biases would result in more accurate ratings. More recently, however, it has been argued that the appropriate focus of rater training should be on the process of making accurate ratings (e.g., Borman, 1979; Hedge, 1982; McIntyre, Smith, & Hassett, 1984; Pulakos, 1984). Since the composition of the rating biases, in terms of true, systematic and error variance, is unknown (Bingham, 1939; Wherry & Bartlett, 1982), it seems clear that reducing rating biases does not directly result in more accurate ratings.

Furthermore, there is empirical evidence that training to reduce rating biases may lead raters to adopt response sets which decrease the rating biases but decrease accuracy as well (Bernardin & Pence, 1980).

Review of training methods. In a recent review, Smith (1986) described training methods utilized in rater training studies. In the review, the training methods fall into three categories: lecture, discussion, and practice and feedback.

The lecture method has generally proven to be ineffective at improving the accuracy of performance ratings. In studies by Borman (1975), McIntyre and Athey (1985), McIntyre et al. (1984), Smith, Hassett, and McIntyre (1982), and Vance, Kuhnert, and Farr (1978), the lecture method failed to improve the accuracy of ratings. In three studies where lectures did improve the accuracy of ratings, the lecture method was combined with either discussions (Hedge, 1982), practice and feedback (Smith, 1984), or both discussion and practice and feedback (Pulakos, 1984).

According to Smith (1986), only one study concerned with the accuracy of ratings has been reported that used discussion alone as the training method. In this study, discussion of stereotypes of effective and ineffective performance did not improve the accuracy of ratings (Bernardin & Pence, 1980). Other studies (McIntyre & Athey, 1985; Pulakos, 1984) have included the discussion method in some way and found improved accuracy. However, in these studies, the discussion method was supplemented by practice and feedback.

Finally, five studies reported that the practice and feedback method served to increase the accuracy of ratings (McIntyre & Athey,



1985; McIntyre et al., 1984; Pulakos, 1984; Silverhart & Dickinson, 1985a; Smith, 1984). It should be noted that a number of studies using practice and feedback have also been reported that were not successful at improving rating accuracy (Borman, 1979; Dickinson & Silverhart, 1986; Hassett, 1987; Silverhart & Dickinson, 1985a, 1985b).

Review of practice and feedback studies. Due to the promising results of practice and feedback in improving rating accuracy, closer review of the studies utilizing this method is warranted. The most common training that was administered in the studies incorporating practice and feedback is referred to as frame-of-reference training (Bernardin & Buckley, 1981). This is the training strategy that was utilized in the studies by Dickinson and Silverhart (1986), Hassett (1987), McIntyre and Athey (1985), McIntyre et al. (1984), Silverhart and Dickinson (1985a, 1985b), and Smith (1984).

Frame-of-reference training was developed by Bernardin and Buckley (1981) based on the rationale that raters will be able to make more accurate ratings of performance if they share a "common nomenclature" (Borman, 1979) with those individuals who define the importance of various behaviors to the organization. This can be accomplished by providing raters with a frame-of-reference that is the same as those whose ratings are used as the standard for comparison. The procedures typically used in the frame-of-reference training programs familiarize participants with the performance dimensions, and with examples of highly effective, average, and ineffective performance on each dimension. Participants then view rating stimuli and rate the performance of the target person in each

stimulus on each dimension. After participants rate each stimulus, target scores are presented so that participants can compare their ratings with experts' ratings. In addition, a behavioral rationale explaining the target scores is provided.

Pulakos (1984) has recommended a method similar to frame-of-reference training. The difference between Pulakos' (1984) rater accuracy training and frame-of-reference training is that rather than providing specific behavioral rationale for target scores, she discussed general behaviors that were indicative of various levels of effectiveness for each of the dimensions utilized.

Three sets of rating stimuli involving videotaped scenerios of performance have been utilized in these training studies. In one set, the target person is an instructor presenting a lecture (see Murphy, Garcia, Kerkar, Martin, & Balzer, 1982). This served as the stimuli for the studies by Hassett (1987), McIntyre and Athey (1985), McIntyre et al. (1984), and Smith (1984). In a second set of stimuli, the target person is a recruiter conducting an interview with a perspective job applicant (see Borman, 1977). These stimuli were used in the studies by Dickinson and Silverhart (1986), and Silverhart and Dickinson (1985a, 1985b). A third set of rating stimuli were used by Pulakos (1984). These stimuli, also developed by Borman (1977), depict discussions between a manager and a problem subordinate.

A review of the results of the studies in which practice and feedback has been used to improve rating accuracy reveals that three different measures of accuracy have served as the dependent variables. The measure most often used is distance accuracy ( Hassett, 1987; McIntyre & Athey, 1985; McIntyre et al., 1984; Silverhart & Dickinson,

1985a, 1985b; Smith, 1984). Distance accuracy is a measure of the deviation between ratings and target scores. All of the studies which focused on distance accuracy, with the exception of the studies by Hassett (1987) and Silverhart and Dickinson (1985b), found that rater training led to some improvements on this measure. McIntyre and Athey (1985) reported that frame-of-reference training resulted in distance accuracy scores that were significantly better than a placebo training group but not better than a no-training control. The placebo training group was a form of control in which non-specific information was presented to subjects to equate the amount of time and individual attention that was required in the frame-of-reference training.

McIntyre et al. (1984) found that frame-of-reference training led to greater distance accuracy than that obtained with rater error training and a no-training control, and that a combination of frame-of-reference and rater error training were more accurate than the rater error training by itself and the no-training control. Silverhart and Dickinson (1985a) reported that participants receiving frame-of-reference training showed significantly more distance accuracy than a no-training control group. Also, in the study by Smith (1984), it was reported that three types of training produced greater distance accuracy than did a control group, while there were no differences among the training groups. Although Smith (1984) reported that the measure of accuracy that he utilized was a measure of Cronbach's (1955) differential accuracy, it was computed as distance accuracy (i.e., the average sum of the absolute values of the deviations between the subject's ratings and the target scores). The three types of training that were used in Smith's (1984) study were observation

training, observation plus dimension training, and observation plus dimension plus performance standards training (i.e., frame-of-reference training).

Two studies did not find differences on the distance accuracy measure. Silverhart and Dickinson (1985b) compared frame-of-reference training with a no-training control, while Hassett (1987) compared two separate forms of frame-of-reference training (i.e., lecture presentation and group discussion presentation) and a no-training control. These two studies will be discussed further below.

It should be noted that while distance accuracy was the most commonly used measure of accuracy, the meaning of the measure has recently been called into question. Dickinson (1986) has demonstrated that distance accuracy is not a unidimensional measure of accuracy but rather is a composite made up of components of halo, leniency/severity and correlation accuracy. Dickinson suggests that interpretations of rater accuracy made on the basis of distance accuracy are ambiguous due to the multiple, underlying sources of inaccuracy.

A second measure of accuracy that was included in the studies being examined is referred to as correlation accuracy or correlational accuracy per ratee ( Dickinson & Silverhart, 1986; Hassett, 1987; McIntyre & Athey, 1985; McIntyre et al., 1984; Silverhart & Dickinson, 1985a, 1985b). This measure provides an index of how accurately a ratee's performance can be described with behavioral dimensions. Correlation accuracy in the McIntyre et al. (1984) study was found to be significantly higher in the frame-of-reference, and the combination of frame-of-reference and rater-error training groups than the rater error or the no-training control groups. In the studies by Dickinson

and Silverhart (1986), Hassett (1987), McIntyre and Athey (1985), and Silverhart and Dickinson (1985a, 1985b), no differences were found on the correlation accuracy measure between any of the training conditions that were investigated.

The third accuracy measure that was utilized is referred to as differential accuracy or correlational accuracy per dimension (Dickinson & Silverhart, 1986; Pulakos, 1984; Silverhart & Dickinson, 1985a, 1985b). This measure provides an index of how accurately a dimension can be used to rate a group of ratees. Using a measure of correlational accuracy per dimension, Pulakos (1984) reported that (a) the rater accuracy training groups were significantly more accurate than all other groups, (b) there were no differences between rater error training and a combination of rater accuracy/rater error training, and (c) the rater accuracy and combination rater accuracy/rater error training were both significantly more accurate than a no-training control. In the three studies that the present author was involved in (Dickinson & Silverhart, 1986; Silverhart & Dickinson, 1985a, 1985b), no differences were found between any of the training conditions on the correlational accuracy per dimension measure.

While previous research has shown some positive results for developing rater training programs that yield improvements in rating accuracy, the presence of the negative findings suggests that the training procedures that have been utilized may still not be sufficient for a complex rating task. It has been suggested that the difficulty in overcoming the biases to which performance ratings are vulnerable is a result of the complex cognitive processing required by

the rating task (Denisi, Cafferty, & Meglino, 1984; Feldman, 1981). In an attempt to understand better and to improve performance judgments, current researchers have focused on the cognitive processes that underlie performance measurement (Banks & Murphy, 1985). Generally, these cognitively oriented researchers have addressed the manner in which raters observe, categorize, recall, and integrate information concerning employees to be evaluated (Cooper, 1981; DeNisi, Cafferty, & Meglino, 1984; Feldman, 1981). Incorporating knowledge about these processes into rater training programs may be necessary for training to be effective given the complex nature of the task. In the decision making literature, there is an approach that focuses on how people learn to integrate information and may have generalizability to the rating of performance. This approach is referred to as the multiple cue probability learning (MCPL) paradigm.

#### Multiple Cue Probability Learning

The MCPL paradigm has its roots in the psychology of Egon Brunswick (1943, 1952). Brunswick considered the behavior of a decision maker to be a function of information within the environment in which the behavior occurred. The properties of the environment were considered to be probabilistically related. It was believed that in order to perform effectively in a given environment, a decision maker must learn the probability of different events occurring, as well as the interrelationships between different events. Brunswick developed a "Lens Model" to describe the relationship between an individual and his or her environment. An illustration of this model can be seen in Figure 1.

There are three essential components in the lens model. At the

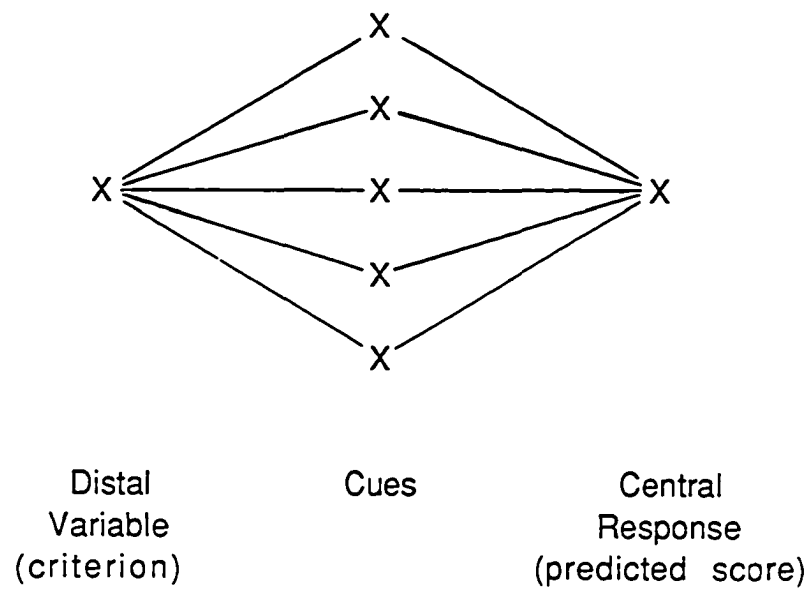


Figure 1. Brunswick's (1943) lens model.

far left is the distal variable or criterion. This represents the physical properties of an object in the environment. In the center of the model are a number of cues. These are events which are probabilistically related to the criterion. The weights of the cues and the form of the relationship between the cues and the criterion can vary across cues. The lines between the cues and the criterion represent the true relationship between these variables. The component on the far right of the model is referred to as the central response. This represents an individual's predicted score or judgmental decision about the criterion based on the properties of the object in the environment. The lines between the cues and the central response indicate the relationships between the cues and the decision maker's judgment.

Mathematical equations have been derived to describe the interrelationships of the various components of the lens model (Hursch, Hammond, & Hursch, 1964; Tucker, 1964). Furthermore, the lens model has been used to investigate the variables that affect how individuals learn to utilize information about the cues in making judgments. The variables studied have included the number of cues, the form of the cues, cue intercorrelations, the distribution of cue validities, the form of the relationship between the cues and the criterion, the predictability of the task, the effects of group discussion on the utilization of cues, and the effects of different forms of feedback (see Zellinger, 1981 for a review).

In essence, the MCPL paradigm is a technique for training individuals to make a judgment based on a set of cues that are probabilistically related to that criterion. This paradigm has



implications for rater training in a performance appraisal context, since the task required of raters in training is similar to that addressed in the MCPL paradigm. In rater training programs, the goal is to teach raters the relationship between the relevant behavioral cues of ratees and the ratings of dimensions made by experts. In terms of the MCPL paradigm, the relevant behaviors are probabilistic cues, and the experts' ratings are the distal variables (i.e., the criteria).

#### Implications of MCPL for Rater Training

Some of the findings from studies conducted within the MCPL paradigm have relevance for training on complex judgment tasks like performance rating. One area in particular that may have relevance to rater training is research on feedback in MCPL situations.

Feedback has typically been defined as information received by an individual about past performance that, in part, provides an indication of the accuracy or correctness of a response (Annett, 1969). As such, feedback, or knowledge of results has long been recognized as being critical for learning and motivation (Ammons, 1956; Annett, 1961; Wexley & Latham, 1981). More recently, however, it has been acknowledged that different types of feedback exist (Ilgen, Fisher, & Taylor, 1979). Rather than merely providing information about the correctness, accuracy, or appropriateness of a response (outcome feedback), it is possible for a feedback message to convey information that represents the "how and why" that underlies accuracy (referred to as information value feedback).

Within the MCPL paradigm, the relative efficacy of outcome and information value feedback has been studied extensively. Consistently

it has been reported that information-value feedback (called cognitive-oriented or lens-model feedback in the MCPL paradigm) is more effective for teaching complex judgment tasks than is outcome feedback (Adelman, 1981; Hammond & Summers, 1965; Lindell, 1976; Newton, 1965; Nystedt & Magnusson, 1973; Summers & Hammond, 1966; Todd & Hammond, 1965). Actually, outcome feedback has been shown to have a debilitating effect on the learning of these judgment tasks (Hammond, Summers, & Deane, 1973).

For example, in the study by Hammond, Summers, and Deane (1973), participants predicted a criterion score on the basis of three cues. Each cue was a value on one of three vertical, uncalibrated scales arranged on a 5 x 7 inch card. The task was developed to contain certain characteristics. The task was constructed so that a) there was a specific functional relationship between the cues and the criterion, b) the three cues had different validities, c) the inter-correlations between the cues were near zero, and d) only 88 percent of the variance in the criterion was accounted for by the cues. Subjects were presented with either outcome-feedback (i.e., the correct responses), information-value feedback (i.e., the functional relationship between the cues and the criterion, and the cue validities), or both outcome and information-value feedback. Results of the study indicated that outcome feedback was insufficient to learn the task. Furthermore, when outcome feedback was provided in addition to information-value feedback, the task was not learned as well as when only information-value feedback was provided.

As early as 1965, it was argued that outcome feedback may not be appropriate for learning multiple cue probability tasks. The reason

for this is that outcome feedback only allows for a comparison of end results when what is really important is the learning of relationships between cues and the criteria (Todd & Hammond, 1965). Hammond (1971) stated that the learning that occurs as a result of outcome feedback is slow and "stupid." Hammond meant that with outcome feedback, even when participants learn, they cannot explain what it is that they have learned. Furthermore, it has been suggested that when learning a task which involves probabilistic relationships between the cues and the criterion, outcome feedback messages may contain erroneous nonrepresentative information. In turn, the erroneous information can cause response inconsistencies that will have an adverse effect on performance (Hammond & Summers, 1972; Hendrix & Dudycha, 1981).

The implications of these findings for the improvement of rater training programs in a performance evaluation context is apparent. In studies that utilized practice and feedback as the training method, the information that was provided as feedback consisted of target scores and either a behavioral rationale that the target scores were based on (i.e., frame of reference training), or a discussion of behaviors that were indicative of different levels of effectiveness (i.e., Pulakos's rater accuracy training). The presentation of target scores is outcome feedback. The rater is provided with the opportunity to compare the ratings that he or she gave to what the experts have defined as the correct responses. Providing a behavioral rationale or general discussion of relevant behaviors serves to provide the rater with information about the characteristics of the task. That is, this information identifies relevant cues that raters should take into account when making a rating.

When two types of information are presented as feedback, as in the frame-of-reference and rater accuracy training programs, it is not possible to tell which of the forms of feedback are effective or whether the combination of the two is necessary. In addition, the results from the MCPL studies suggest that providing target scores as feedback may have an adverse effect on the effectiveness of training, even in the presence of information about characteristics of the task (i.e., information-value feedback).

Results from MCPL studies also have implications for when information about the rating task should be provided. It has been suggested that providing information about the task before the activity takes place may be an effective strategy for learning complex judgment tasks (Bogart, 1980; Hendrix & Dudycha, 1981). This technique is referred to as feedforward (Bjorkman, 1972).

Unfortunately, studies of the effectiveness of feedforward have not been conclusive. While there has been some indication that feedforward information improves performance on cognitive judgment tasks (Dudycha, Dudycha, & Schmitt, 1973; Newton, 1965; Magnusson & Nystedt, 1971; Nystedt & Magnusson, 1973), there are also some results that suggest feedforward is not effective (Hendrix & Dudycha, 1981; Holt, 1958). Furthermore, while Hendrix and Dudycha (1981) did include feedback and feedforward in the same design, no studies have compared the relative effectiveness of these two strategies.

Typically, in the MCPL paradigm, feedforward has consisted of information about the weights assigned to the cues, and the form of the relationship between the cues and the criteria. In the context of the rater training studies, providing target score information serves

to identify the criteria (i.e., the distal variables). Providing the experts' behavioral rationale serves to identify the target behaviors (or cues). Since previous research has indicated that feedforward information may improve performance on cognitive judgment tasks (Dudycha, Dudycha, & Schmitt, 1973; Newton, 1965; Magnusson & Nystedt, 1971; Nystedt & Magnusson, 1973), providing target score or behavioral rationale information as feedforward may improve the effectiveness of rater training.

### Research Hypotheses

The purpose of the present study is to examine two factors that may affect the efficacy of rater training for improving the accuracy of performance ratings. One factor that will be examined is the type of training information that is presented in rater training programs. Based on the findings in the MCPL literature that outcome information is less effective than information value feedback for learning complex cognitive judgment tasks, the following hypothesis was made:

Hypothesis 1: Rater training that incorporates target score information, or combines target score information with a behavioral rationale for the expert ratings will result in less accurate performance ratings than rater training incorporating only the behavioral rationale.

The second factor that will be investigated is whether training information in rater training programs is presented as feedback or feedforward. While current training programs typically present training information in the form of feedback, studies within the MCPL paradigm suggest that when performing complex judgment tasks, the use of feedforward may be an effective strategy (Dudycha, Dudycha, &

Schmitt, 1973; Newton, 1965; Magnusson & Nystedt, 1971; Nystedt & Magnusson, 1973). Based on this MCPL the following hypothesis was made:

Hypothesis 2: Performance ratings will be more accurate when raters receive training information by means of feedforward than when control training is provided in which the training information is not presented.

In addition, the present study explored the relative effectiveness of providing training information through feedback and feedforward. Since these two strategies for presenting information in training had not previously been compared, directional hypotheses were not made.

## CHAPTER TWO

### Method

#### Participants

The participants in this study were 101 undergraduate and graduate students from Old Dominion University (51 male, 50 female). All participants were compensated 40 dollars for their involvement in the study. The data from five participants were not included in the analyses. Three of these did not return for the second session, and two had previously participated in research which utilized the same rating stimuli. The mean age of the final sample was 21.69. Experimental sessions were assigned randomly to groups of participants.

#### Stimulus Materials

Ten videotaped sessions of a simulated performance evaluation interview were used as rating stimuli in this study. This interview simulation was one of five exercises administered prior to the present study in a management assessment center conducted for the purpose of developing performance measurement stimuli. Performance of 45 assessment center participants in the interview simulation was videotaped from which ten (five male, five female) were selected in such a way as to represent a range of effectiveness on a variety of performance dimensions. Subsequently, these videotaped role-play performances were transcribed and reenacted on film by drama students with previous acting experience. The reason for the reenactment was

to improve the technical quality of the final stimuli and to protect the anonymity of the original assesseees. To ensure that the final stimuli accurately represented the levels of performance exhibited in the original videotapes, the actors familiarized themselves thoroughly with the scripts, reviewed the original videotapes, and then received videotaped feedback of their performance before the final stimuli were filmed. The length of the videotaped simulations ranged from six to 15 minutes. Scripts of the ten interview simulations are included in Appendix A.

### Rating Forms

Participants rated the performance exhibited on the videotaped interview simulations on five-point behaviorally anchored rating scales (BARS). These scales had previously been developed in an effort to investigate the construct validity of assessment center simulation exercises and ratings of the behavioral dimensions (Campbell, 1986). The procedures used in the scale development are described in detail by Campbell (1986). Performance ratings were obtained on three behavioral dimensions: problem analysis, problem solution, and sensitivity. Definitions of the three behavioral dimensions, and the BARS rating forms are included in Appendices B and C, respectively.

### Design

This study utilized a factorial design that crossed three levels of "information type" (target score, behavioral rationale, combination target score/behavioral rationale) with two levels of "mode of information presentation" (feedback or feedforward). In addition, dimension-training control and a no-training control conditions were



included in the design.

Information type. Participants in all conditions received three practice trials. In each of these practice trials, participants observed a videotaped interview simulation, and took notes on the behaviors believed to be relevant to the problem analysis, problem solution, and sensitivity dimensions. Participants in the target score information conditions were presented with the mean expert ratings on each of the three dimensions for each of the practice trials. These target scores were presented orally as well as displayed with an overhead projector. Participants in the behavioral rationale information conditions were provided with a lecture describing the behaviors that experts considered in determining their ratings. This lecture focused on a checklist of the actual behaviors that the expert raters considered in making their ratings. In addition, participants viewed segments from the videotaped interview simulations that illustrated the behaviors that the expert raters focused on. The combination target score and behavioral rationale conditions consisted of presenting all of the information from the target score and the behavioral rationale conditions described above. The checklist of the behaviors that were considered by the expert raters and the script for the behavioral rationale lecture are included in Appendices D and E, respectively.

Mode of information presentation. In the feedback conditions, participants viewed and rated each of the three practice interview simulations. After making these ratings, they received feedback information. In the feedforward conditions, participants received all information before viewing each of the three practice interview

simulations. Because participants in the target score and combination of target score and behavioral rationale feedforward conditions were provided with the target scores before viewing each of the practice interview simulations, ratings of the practice interview simulations were not made in these conditions.

To ensure that the subjects attended to the information provided by the feedback or feedforward, certain condition-specific procedures were employed. In the target score feedback condition, participants plotted their ratings and the target scores on a graph. This provided the participants with a visual comparison of their ratings and the target scores. Participants in the target score feedforward condition simply plotted the target scores on a graph before viewing each of the practice simulations.

In the behavioral rationale feedback condition, after the behavioral rationale for the experts' ratings was provided, the participants completed a checklist which contained the behaviors that the expert raters focused on in determining the target ratings. The participants indicated which behaviors on the checklist they believed they considered in determining their ratings. This allowed the participants the opportunity to compare the behaviors that the expert raters focused on with the behaviors that they considered in determining their ratings. Participants in the feedforward behavioral rationale condition checked behaviors on the checklist as they saw them occur while viewing each of the practice simulations. The conditions that combined target score and behavioral rationale information each experienced both of the procedures that the appropriate (either feedback or feedforward) target score and

behavioral rationale groups received.

Control conditions. Two conditions were included in the design of the study to control for the information presented in training. One control condition was a dimension-training control group. This control condition was intended to generalize to "basic" training given in most organizational contexts. This group received the same training regarding the dimension definitions, proper use of the rating forms, and familiarization with the behavioral anchors on the rating forms as the experimental conditions. No information was provided, however, concerning the performance viewed in the three practice trials.

The second control condition was a no-training control group. This group did not receive the training on the dimension definitions or the behavioral anchors on the rating forms that all of the other conditions received. After receiving instruction for the proper use of the rating forms, this group was given time to read and become familiar with the definitions and anchors. In addition, no information was provided regarding the performance viewed in the three practice trials.

#### Target Score Development

Target scores for the ten interview simulations were generated by five advanced graduate students in an industrial/organizational psychology doctoral program. These "expert" raters had all received considerable training in performance rating processes, the performance evaluation interview, and assessment center simulation exercises. In developing the target scores, experts viewed each interview simulation at least two times and familiarized themselves with scripts of the

simulations. For each videotaped ratee, after all experts expressed confidence that they recognized the relevant behaviors in a simulation, they completed a behavioral checklist and rated each behavioral dimension. The behavioral checklist contained 15 behaviors for each of the three dimensions which were observed in the original assessment center performances (Campbell, 1986). Experts used scripts and notes they took during the viewing of the simulations to complete the checklist and make their ratings. Counter-balancing techniques were used to control for the order in which the checklist and rating forms were completed. Expert raters were assigned to two groups and for each assessee the groups alternated the order in which the checklist and rating forms were completed.

After each expert completed the checklist and rating form, they discussed all discrepancies on the checklist. The goal of this discussion was to establish a group consensus for behaviors to be cited as target behaviors. After the discussion on the target behaviors, the expert raters discussed differences in BARS ratings and then re-rated each simulation. The target scores were defined as the means of the expert raters' final ratings. These scores are presented in Appendix F.

The psychometric quality of these final ratings was analyzed by means of a three-way analysis of variance. The analysis included factors of raters, assessees and dimensions. Results of the analysis indicated that there was a significant effect for ratees, a significant effect for dimensions, and a significant ratees x dimensions interaction ( $p < .01$ ). A summary of this analysis can be seen in Table 1.

Table 1

Summary of the Analysis of Variance for the Psychometric Properties of the Target Scores.

Source	DF	MS	F-Ratio
Raters (Rr)	4	0.0567	0.47
Ratees (R)	9	12.2696	102.56*
Dimensions (D)	2	20.6467	6.99* <sup>a</sup>
Rr x R	36	0.1196	No Test
Rr x D	8	0.0467	0.47
R x D	18	2.9207	29.65*
Rr x R x D	72	0.0985	

<sup>a</sup>  
Quasi-F-Ratio.

\* $\underline{p} < .01$ .

The significant main effect for ratees indicated that the ratings contain the desirable property of convergent validity (Cambell & Fiske, 1975; Kavanagh, MacKinney, & Wolins, 1970). Furthermore, the significant ratees x dimensions interaction indicates the presence of discriminant validity. Importantly, there were no rater effects, indicating agreement among the raters on the target scores.

### Procedures

At the onset of the first experimental session, informed consent was received from all participants. The form utilized for obtaining informed consent is included in Appendix G. A videotaped presentation, in which the experimenter served as narrator, was then shown to the participants. The script for this presentation is included in Appendix H. In the presentation, the experimenter explained that the purpose of the research was to examine how people observe and evaluate performance. It was noted that of particular concern was the manner in which the performance of managers or potential managers was observed and evaluated in management assessment centers. Since the rating stimuli that were used in the study were from an actual assessment center exercise and based on actual assessment center performances, it was believed that describing the participants' rating task in terms of an assessment center would have the greatest face validity.

Following a brief description of management assessment centers, the scenario of the assessment center exercise that was to be evaluated by the participants (i.e., the interview simulation) was explained. To enhance understanding of the simulation exercise, participants were provided with the description of the scenario that

was provided to the original assessment center participants before participating in the interview simulation. This description is included in Appendix I. In addition, a videotaped example of an interview simulation, in which the experimenter played the role of the assessee, was shown to the participants. The script for this interview simulation can be seen in Appendix J.

The behavioral dimensions that were used to evaluate the effectiveness of the assessee's performance in the interview simulations were then presented. Definitions of each dimension were provided to the participants. For all conditions except the no-training control, the experimenter read aloud and discussed each of these definitions.

Participants were then instructed in proper use of the BARS rating format. It was explained that the evaluation forms contained five behavioral statements meant to represent five different levels of performance and that the participants should circle the statement which best reflected the level of performance on each dimension that was viewed in an interview. Furthermore, instructions were provided to clarify what was meant by the "could be expected" format of the BARS forms. It was emphasized that based on the behaviors demonstrated in the interview, the participants should select the statement on the rating form that described the level of performance that they could consistently expect of the assessee. It was also pointed out that some of the behaviors that the participants would observe in the interviews were actually listed as statements on the rating forms. It was emphasized that simply because those behaviors may have occurred in an interview it did not mean that the statement

describing that behavior should be selected as a rating. Rather, it was explained that before making a rating, all behaviors relevant to a dimension that were exhibited in an interview should be taken into consideration. The actual statement chosen for a rating should reflect the level of performance that could consistently be expected of the assessee.

Following the instructions for proper use of the rating scales, the three practice interviews were viewed. At this point the experimental manipulations of the information type and mode occurred. After completion of the practice trials, the first session was terminated.

Participants returned the following day for a second session. At this time, brief refresher training was provided of the information presented on the dimension definitions and use of the rating scales. Random identification numbers were assigned to all participants to ensure anonymity of their ratings. Participants viewed and rated seven experimental interview simulations after receiving instructions to take notes as they watched the interviews to use when making their ratings. Following the rating of the seven interviews, a post-experimental questionnaire was administered and the participants were debriefed about the nature of the research project. The post-experimental questionnaire is included in Appendix K.

#### Evaluation of Dimension Training

Three tests were administered for the purpose of evaluating whether training regarding the dimension definitions was effective. These tests were given at the beginning of the training session (i.e., pre-training), at the end of the training session (i.e., post-



training), and at the beginning of the rating session the following day (i.e., pre-rating). The tests consisted of matching exercises in which behavioral statements had to be matched with appropriate dimension titles (i.e., problem analysis, problem solution, sensitivity). The statements were selected from a checklist describing possible behaviors that could be exhibited in the interview simulations (Campbell, 1986). The pre-training and pre-rating tests contained all of the same statements. Copies of the three tests and the instructions given to participants are included in Appendix L.

### Accuracy Measurement

Three approaches to accuracy measurement were utilized in this study. These included measures of correlational accuracy, Cronbach's (1955) accuracy statistics and an extension of Dickinson's (1987) analysis of variance accuracy design.

### Measures of Correlational Accuracy

Correlational accuracy per ratee. This is a measure of how accurately a ratee's performance can be described using the behavioral dimensions (e.g., Hoffman & Dossett, 1984; McIntyre et al., 1984). Correlational accuracy per ratee was calculated for each ratee by computing the r-to-z transformed correlations of the ratings on the three dimensions with the target scores.

Correlational accuracy per dimension. This is a measure of how accurately a dimension can be used to rate a group of ratees (e.g., Borman, 1979; Cardy & Kehoe, 1984; Pulakos, 1984). Correlational accuracy per dimension was calculated for each dimension by computing the r-to-z transformed correlations of the ratings for the seven ratees on that dimension with the target scores.

### Cronbach's Accuracy Statistics

In situations where ratings are made of more than one ratee, on more than one dimension of performance, each raters' overall accuracy can be expressed as the sum of four components: elevation, differential elevation, stereotype accuracy, and differential accuracy (Borman, 1977; Cronbach, 1955; Murphy, Garcia, Kerkar, Martin, & Balzar, 1982). Computational formulas for each of these components are included in Appendix M.

Elevation. Elevation represents a rater's overall average rating, taken across all ratees and all dimensions. Accuracy on this component reflects consistency between the rater's overall average rating and the overall average rating for the target scores.

Differential elevation. Differential elevation represents the average rating that a rater gives a ratee, taken across all of the dimensions. Accuracy in terms of differential elevation suggests that a rater has rank ordered the ratees in the same order as the target scores.

Stereotype accuracy. Stereotype accuracy represents the average rating, taken across all ratees, for each of the performance dimensions. This component of accuracy reflects the ranking of dimensions or the extent to which the ratings reflect the same relative strengths of the group of ratees as the target scores.

Differential accuracy. Differential accuracy represents the extent to which a rater accurately rates each individual ratee. For each ratee, ratings on each of the dimensions are compared with the target scores. This measure has been suggested to be the most important of the four accuracy components (Cronbach, 1955).

### Extended Accuracy Design

Dickinson (1987) described an analysis of variance approach to evaluate the variance in performance ratings. This approach utilizes a basic accuracy design which includes the factors of rating sources, ratees, and dimensions. In the context of this design, the sources for ratings are the participant raters, and the target scores derived from experts. The sums of squares obtained from the analysis of these factors can be interpreted in terms of Cronbach's accuracy statistics (see Dickinson, 1986). The psychometric interpretations of the factors in the basic accuracy design can be found in Table 2.

Elevation accuracy is reflected in the variance due to the rating sources. Variation due to the rating sources indicates differences between the raters overall mean rating and the mean of the target scores. Thus, the more variance accounted for by this factor, the less accurate is the rater.

Differential elevation accuracy is reflected in the interaction between rating sources and ratees. This source of variation suggests that the rater orders the ratees differently than the order reflected in the target scores. Since an accurate rater would order the ratees the same as the expert raters, variance due to this interaction is undesirable.

Stereotype accuracy is indicated in the interaction between rating sources and dimensions. Variation due to this interaction represents differences between the raters and target scores in terms of the relative amount that the dimensions are exhibited by the group of ratees. The larger these differences, and thus the larger the magnitude of the interaction, the more inaccuracy there is reflected

Table 2

Summary Table for the Psychometric Interpretations of the Basic Accuracy Design.

Source	Psychometric interpretation
Rating sources (S)	Elevation accuracy
Ratees (R)	Convergent validity
Dimensions (D)	Trait bias
S x R	Differential elevation accuracy
S x D	Stereotype accuracy
R x D	Discriminant validity
S x R x D	Differential accuracy

in the ratings.

Differential accuracy is reflected in the three-way interaction between rating sources, ratees, and dimensions. This accuracy component measures how a rater describes the individual differences among ratees relative to the target scores. Since agreement between the rater and target scores indicates accuracy, large sources of variance due to this interaction are undesirable.

In the present study, the basic accuracy design was extended to include the between-subjects experimental variables of information type, information mode, and contrasts for both the dimension-training control and no-training control with the mean of the experimental conditions. The interactions of these additional variables with the factors in the basic accuracy design allow for assessment of the differential effects of the experimental treatments on the accuracy and validity of the ratings. A summary of the extended design, along with the psychometric interpretations of all of the effects is included in Table 3. The appropriate error terms for the design can be found in Table 4.

-

Table 3

Summary Table for the Psychometric Interpretations of the Extended Accuracy Design.

Source	Psychometric interpretation
Information type (IT)	Research condition
Information mode (M)	Research condition
IT x M	Research condition
Dimension-training control (DTC)	Research condition
No-training control (NTC)	Research condition
Rating sources (S)	Elevation accuracy
S x IT	Elevation accuracy differing by information type
S x M	Elevation accuracy differing by information mode
S x IT x M	Elevation accuracy differing by information type and mode
S x DTC	Elevation accuracy differing by dimension-training control
S x NTC	Elevation accuracy differing by no-training control
Ratees (R)	Convergent validity
R x IT	Convergent validity differing by information type
R x M	Convergent validity differing by information mode
R x IT x M	Convergent validity differing by information type and mode
R x DTC	Convergent validity differing by dimension-training control

Table 3 (continued)

Source	Psychometric interpretation
R x NTC	Convergent validity differing by no-training control
Dimensions (D)	Trait bias
D x IT	Trait bias differing by information type
D x M	Trait bias differing by information mode
D x IT x M	Trait bias differing by information type and mode
D x DTC	Trait bias differing by dimension-training control
D x NTC	Trait bias differing by no-training control
S x D	Stereotype accuracy
S x D x IT	Stereotype accuracy differing by information type
S x D x M	Stereotype accuracy differing by information mode
S x D x IT x M	Stereotype accuracy differing by information type and mode
S x D x DTC	Stereotype accuracy differing by dimension-training control
S x D x NTC	Stereotype accuracy differing by no-training control
S x R	Differential elevation accuracy
S x R x IT	Differential elevation accuracy differing by information type
S x R x M	Differential elevation accuracy differing by information mode

Table 3 (concluded)

Source	Psychometric interpretation
S x R x IT x M	Differential elevation accuracy differing by information type and mode
S x R x DTC	Differential elevation accuracy differing by dimension-training control
S x R x NTC	Differential elevation accuracy differing by no-training control
R x D	Discriminant validity
R x D x IT	Discriminant validity differing by information type
R x D x M	Discriminant validity differing by information mode
R x D x IT x M	Discriminant validity differing by information type and mode
R x D x DTC	Discriminant validity differing by dimension-training control
R x D x NTC	Discriminant validity differing by no-training control
S x R x D	Differential accuracy
S x R x D x IT	Differential accuracy differing by information type
S x R x D x M	Differential accuracy differing by information mode
S x R x D x IT x M	Differential accuracy differing by information type and mode
S x R x D x DTC	Differential accuracy differing by dimension-training control
S x R x D x NTC	Differential accuracy differing by no-training control



Table 4

Summary Table for the Error Terms for the Extended Accuracy Design.

Source	Error Term
Information type (IT)	$(Rr/)+(RxIT)-(RxRr/)$
Information mode (M)	$(Rr/)+(RxM)-(RxRr/)$
IT x M	$(Rr/)+(RxIT \times FT)-(SxRxRr)$
Dimension-training control (DTC)	$(Rr/)+(RxDT C)-(SxRxRr)$
No-training control (NTC)	$(Rr/)+(RxNT C)-(SxRxRr)$
Rating sources (S)	$(SxRr/)+(SxA)-(SxRxRr/)$
S x IT	$(SxRr/)+(SxRxIT)-(SxRxRr/)$
S x M	$(SxRr/)+(SxRxM)-(SxRxRr/)$
S x IT x M	$(SxRr/)+(SxRxIT \times M)-(SxRxRr/)$
S x DTC	$(SxRr/)+(SxRxDT C)-(SxRxRr/)$
S x NTC	$(SxRr/)+(SxRxNT C)-(SxRxRr/)$
Rates (R)	$(RxRr/)$
R x IT	$(RxRr/)$
R x M	$(RxRr/)$
R x IT x M	$(RxRr/)$
R x DTC	$(RxRr/)$
R x NTC	$(RxRr/)$
Dimensions (D)	$(DxRr/)+(DxA)-(DxRxRr/)$
D x IT	$(DxRr/)+(DxRxIT)-(DxRxRr/)$
D x M	$(DxRr/)+(DxRxM)-(DxRxRr/)$
D x IT x M	$(DxRr/)+(DxRxIT \times M)-(DxRxRr/)$
D x DTC	$(DxRr/)+(DxRxDT C)-(DxRxRr/)$

Table 4 (continued)

Source	Error Term
D x NTC	$(DxRr) + (DxRxNTC) - (DxRxRr/)$
S x D	$(SxDxRr/) + (SxDxA) - (SxDxRxRr/)$
S x D x IT	$(SxDxRr/) + (SxDxRxIT) - (SxDxRxRr/)$
S x D x M	$(SxDxRr/) + (SxDxRxM) - (SxDxRxRr/)$
S x D x IT x M	$(SxDxRr/) + (SxDxRxIT x M) - (SxDxRxRr/)$
S x D x DTC	$(SxDxRr/) + (SxDxRxDT C) - (SxDxRxRr/)$
S x D x NTC	$(SxDxRr/) + (SxDxRxNTC) - (SxDxRxRr/)$
S x R	$(SxRxRr/)$
S x R x IT	$(SxRxRr/)$
S x R x M	$(SxRxRr/)$
S x R x IT x M	$(SxRxRr/)$
S x R x DTC	$(SxRxRr/)$
S x R x NTC	$(SxRxRr/)$
R x D	$(DxRxRr/)$
R x D x IT	$(DxRxRr/)$
R x D x M	$(DxRxRr/)$
R x D x IT x M	$(DxRxRr/)$
R x D x DTC	$(DxRxRr/)$
R x D x NTC	$(DxRxRr/)$
S x R x D	$(SxDxRxRr/)$
S x R x D x IT	$(SxDxRxRr/)$
S x R x D x M	$(SxDxRxRr/)$
S x R x D x IT x M	$(SxDxRxRr/)$

Table 4 (concluded)

Source	Error Term
S x R x D x DTC	(SxDxRxRr/)
S x R x D x NTC	(SxDxRxRr/)

Note. Rr/ = Raters nested within IT, M, DTC, and NTC.

## CHAPTER THREE

### Results

#### Evaluation of Dimension Training

Planned comparisons were conducted to determine the difference between the mean of the experimental training conditions and the two control conditions on the pre-training, post-training, and pre-rating tests. No significant differences were detected between the mean of the experimental groups and either the dimension-training control ( $F(1,83) = 0.37, p > .05$ ) or the no-training control ( $F(1,83) = 0.98, p > .05$ ) for the number of correct responses on the pre-training test. For the post-training test, the planned comparisons indicated that the experimental groups scored significantly higher than did the no-training control ( $F(1, 83) = 7.61, p < .01$ ). No difference was found for the number of correct responses between the experimental groups and the dimension-training control ( $F(1,83) = 1.15, p > .05$ ).

For the pre-rating test, no significant differences were indicated for the number of correct responses between the experimental conditions and either the dimension-training control ( $F(1,83) = .44, p > .05$ ) or the no-training control ( $F(1,83) = 1.00, p > .05$ ).

In sum, these results suggest that before training there was no difference between the experimental and control conditions in terms of understanding the dimensions. After training, the experimental groups were able to match specific behaviors to the appropriate dimensions more correctly than the no-training control. This

difference, however, was not detected the next day when the participants returned to complete the rating task.

### Evaluation of Information Type and Mode

Three approaches were utilized to analyze the effects of information type and mode on the quality of the performance ratings. These were: 1) measures of correlational accuracy, 2) Cronbach's (1955) accuracy measures and, 3) the extended accuracy design (Dickinson, 1987).

Hypothesis 1 stated that rater training that incorporates target score information, or combines target score information with a behavioral rationale for the expert ratings will result in less accurate performance ratings than rater training incorporating only the behavioral rationale. With the first two approaches, Hypothesis 1 was assessed with tests of the main effect for the type of information presented during training (i.e., information type).

Hypothesis 2 stated that performance ratings will be more accurate when raters receive training information by means of feedforward than when control training is provided in which the training information is not presented. With the first two approaches, Hypothesis 2 was tested with planned comparisons between the mean of the conditions receiving feedforward and each of the two control training conditions. Furthermore, exploratory comparisons of the effectiveness of feedforward and feedback were conducted with tests of the main effect for the mode of information presented during training (i.e., information mode).

With the extended accuracy design (Dickinson, 1987), Hypothesis 1 was assessed with tests of the interactions that involved both rating

sources and information type (i.e., Rating Sources x Information Type, Rating Sources x Dimensions x Information Type, Rating Sources x Ratees x Information Type, Rating Sources x Ratees x Dimensions x Information Type). Hypothesis 2 was assessed with tests of the interactions that involved rating sources and contrasts between the control conditions and the feedforward groups (i.e., Rating Sources x Dimension-Training Control versus Feedforward, Rating Sources x No-Training Control versus Feedforward, Rating Sources x Dimensions x Dimension-Training Control versus Feedforward, Rating Sources x Dimensions x No-Training Control versus Feedforward, Rating Sources x Ratees x Dimension-Training Control versus Feedforward, Rating Sources x Ratees x No-Training Control versus Feedforward, Rating Sources x Ratees x Dimensions x Dimension-Training Control versus Feedforward, Rating Sources x Ratees x Dimensions x No-Training Control versus Feedforward). Finally, comparisons of the effectiveness of feedforward and feedback were conducted with tests of the interactions that involved rating sources and information mode (i.e., Rating Sources x Mode, Rating Sources x Dimensions x Mode, Rating Sources x Ratees x Mode, Rating Sources x Ratees x Dimensions x Mode).

#### Correlational Accuracy

Correlational accuracy per ratee. A repeated measures analysis of variance was conducted on the correlational accuracy per ratee measure. This analysis included three levels of information type (target scores, behavioral rationale, combination target scores and behavioral rationale), two levels of the mode in which information was presented during training (feedback, feedforward), and seven repeated measures on the ratees variable. Contrasts between each of the

control groups and the mean of the six experimental conditions were also included in the analysis. Group means and standard deviations for the r-to-z transformed correlations are included in Table 5.

The analysis of the r-to-z transformed correlations (Jones, 1968) indicated that there was a significant effect for ratees ( $p < .01$ ); however, no other main effects or interactions were found to be significant ( $p > .05$ ). Furthermore, neither of the two control groups was found to be significantly different from the mean of the experimental groups ( $p > .05$ ). A summary of this analysis is included in Table 6.

A Newman-Keuls post hoc analysis of the significant effect for ratees indicated that the first ratee was rated significantly more accurately than each of the other six ratees. Means and standard deviations on the correlational accuracy per ratee measure for each of the seven ratees are included in Table 7.

Planned comparisons of the mean of the feedforward conditions with each of the two control groups on the correlational accuracy per ratee measure were conducted. These analyses indicated that there was not a significant difference between raters receiving feedforward and either of the two controls ( $p > .05$ ).

The results of the analyses on the correlational accuracy per ratee measure do not provide support for either of the hypotheses. Specifically, Hypothesis 1 was not supported since a significant effect for information type was not indicated. Hypothesis 2 was not supported since the groups receiving feedforward were not found to be more accurate than either of the control conditions. Furthermore, the finding that there was not a significant effect for the mode of

Table 5

Mean r-to-z Transformed Scores and Standard Deviations for the  
Correlational Accuracy per Ratee Measure.<sup>a</sup>

Information Mode	<u>Information Type</u>		
	TS	BR	TS/BR
Feedback	0.6672 (0.7597)	0.6799 (1.0477)	0.8003 (0.9857)
Feedforward	1.0393 (1.0446)	0.1009 (1.1842)	0.1555 (0.7477)
<u>Control Groups</u>			
	DTC	NTC	
	0.8346 (0.5684)	0.3346 (0.9997)	

Note. TS = Target score; BR = Behavioral rationale; TS/BR =  
 Combination target score and behavioral rationale; DTC = Dimension  
 training control; NTC = No-Training control.

<sup>a</sup>  
 Standard deviations are included in parentheses.



Table 6

Summary of Analysis of Variance for Correlational Accuracy per Ratee.

Source	DF	MS	F-Ratio	<sup>b</sup> VC	<sup>b</sup> ICC
IT	2	10.1570	<sup>a</sup> 1.08	.00072	.0002
M	1	10.1542	<sup>a</sup> 3.19	.00346	.0008
IT x M	2	4.1202	<sup>a</sup> 0.56	-.00320	.0000
DTC	1	4.8950	<sup>a</sup> 0.78	-.00039	.0000
NTC	1	0.6696	<sup>a</sup> 0.18	-.00086	.0000
Rr/IT x M x DTC x NTC	88	6.0063		.01046	
R	6	26.2892	6.19*	.07653	.0175
R x IT	12	7.6778	1.81	.03571	.0082
R x M	6	1.4313	0.34	-.01957	.0000
R x IT x M	12	5.5943	1.32	.02801	.0064
R x DTC	6	4.5019	1.06	.00102	.0002
R x NTC	6	1.9127	0.45	-.00947	.0000
R x Rr/IT x M x DTC x NTC	528	4.2500		4.25000	

Note. If a source's variance component was negative, that value was used in the denominator to compute intraclass correlation coefficients, but the source's coefficient was set to zero. VC = variance component; ICC = intraclass correlation coefficient; IT = information type; M = mode; DTC = dimension-training control; NTC = no-training control, R = ratees; Rr = raters.

<sup>a</sup> Quasi-F-Ratio.

<sup>b</sup> VC's and ICC's computed according to Vaughan and Corballis (1969).

- \*p < .01.

Table 7

Mean r-to-z Transformed Scores and Standard Deviations on the  
Correlational Accuracy per Ratee Measure for each of the Seven  
<sup>a</sup>  
Ratees.

Ratee One	Ratee Two	Ratee Three	Ratee Four	Ratee Five	Ratee Six	Ratee Seven
1.5532 (2.965)	0.6227 (1.329)	0.3161 (1.083)	0.7008 (2.329)	0.8471 (2.527)	0.0687 (2.328)	0.0522 (1.684)

<sup>a</sup>

Standard deviations are included in parentheses.

information presentation suggests there is no difference in effectiveness between feedback and feedforward.

Correlational accuracy per dimension. A repeated measures analysis of variance was also conducted on the correlational accuracy per dimension measure. The design for this analysis was the same as the design described in the section above except that repeated measures were included on the dimensions variable instead of rates. Group means and standard deviations of the r-to-z transformed correlations for correlational accuracy per dimension are included in Table 8.

The analysis yielded significant main effects for information mode ( $p < .05$ ), and dimensions ( $p < .01$ ). No other main effects or interactions were found to be significant, and neither of the control groups were found to be significantly different than the mean of the experimental groups ( $p > .05$ ). A summary of this analysis can be found in Table 9.

Examination of the group means for information mode indicated that participants who received feedback ( $M = 0.7656$ ) rated significantly more accurately than participants whose training incorporated feedforward ( $M = 0.6138$ ). In addition, a Newman-Keuls post hoc analysis of the means for the dimensions indicated that ratings on the sensitivity dimension ( $M = 0.95$ ) were significantly more accurate than both the problem analysis ( $M = 0.49$ ) and the problem solution dimensions ( $M = 0.58$ ).

Planned comparisons of the mean of the feedforward conditions with each of the two control groups on the correlational accuracy per dimension measure were conducted. These analyses indicated that there

Table 8

Mean r-to-z Transformed Scores and Standard Deviations for the  
Correlational Accuracy per Dimension Measure.<sup>a</sup>

Information Mode	<u>Information Type</u>		
	TS	BR	TS/BR
Feedback	0.7768 (0.3827)	0.8300 (0.2996)	0.6903 (0.3661)
Feedforward	0.6959 (0.2620)	0.5301 (0.2323)	0.6154 (0.2247)
<u>Control Groups</u>			
	DTC	NTC	
	0.6657 (0.4192)	0.5845 (0.3111)	

Note. TS = Target score; BR = Behavioral rationale; TS/BR = Combination target score and behavioral rationale; DTC = Dimension training control; NTC = No-Training control.

<sup>a</sup> Standard deviations are included in parentheses.

Table 9

Summary of Analysis of Variance for Correlational Accuracy per Dimensions.

Source	DF	MS	F-Ratio	VC	ICC
IT	2	0.1306	0.43	-.00041	.0000
M	1	1.2456	4.07*	.00109	.0060
IT x M	2	0.2958	0.97	-.00002	.0000
DTC	1	0.0178	0.06	-.00019	.0000
NTC	1	0.3417	1.12	.00002	.0001
Rr/IT x M x DTC x NTC	88	0.3059		.00192	
D	2	5.5174	32.95*	.01238	.0679
D x IT	4	0.2428	1.45	.00035	.0019
D x M	2	0.0527	0.31	-.00027	.0000
D x IT x M	4	0.0320	0.19	-.00009	.0000
D x DTC	2	0.0932	0.56	-.00010	.0000
D x NTC	2	0.4072	2.43	.00032	.0018
D x Rr/IT x M x DTC x NTC	176	0.1674		.16743	

Note. If a source's variance component was negative, that value was used in the denominator to compute intraclass correlation coefficients, but the source's coefficient was set to zero. VC = variance component; ICC = intraclass correlation coefficient; IT = information type; M = mode; DTC = dimension-training control; NTC = no-training control, D = dimensions; Rr = raters.

\*p < .01.

was not a significant difference between raters receiving feedforward and either of the two controls ( $p > .05$ ).

Since a significant effect for information type was not indicated, and no differences were detected between the groups receiving feedforward and the experimental groups, the analyses on the correlational accuracy per dimension measure do not provide support for either of the two hypotheses. However, the significant effect for information mode does suggest that in terms of correlational accuracy per dimension, feedback is more effective than feedforward.

#### Cronbach's Accuracy Measures

An analysis of variance was conducted on each of Cronbach's (1955) accuracy components (elevation, differential elevation, stereotype accuracy, differential accuracy). These analyses each included three levels of information type (i.e., target score, behavioral rationale, combination target score and behavioral rationale), two levels of the mode by which information was presented during training (i.e., feedback, feedforward), and contrasts between each of the two control groups and the mean of the experimental groups. Group means and standard deviations for each of the four accuracy components are included in Table 10.

The results of the analyses indicated that for all four of the components of accuracy there were no significant effects due to information type, information mode, or the information type by information mode interaction ( $p > .05$ ). It should be noted that for differential accuracy, the  $F$  value for information mode did approach significance ( $p < .10$ ). Examination of the means suggested that participants who received feedforward ( $M = 0.7488$ ) were higher on the

Table 10

Group Means and Standard Deviations for Cronbach's Accuracy Measures.<sup>a</sup>

Cronbach's Measures				
	Elevation	Differential Elevation	Stereotype Accuracy	Differential Accuracy
FB-T S	0.3333 (0.2279)	0.6275 (0.2732)	0.5666 (0.2762)	0.6440 (0.1511)
FB-BR	0.4325 (0.4460)	0.6498 (0.1498)	0.4566 (0.1537)	0.6591 (0.1116)
FB-T S/BR	0.4286 (0.2956)	0.7597 (0.2915)	0.5940 (0.2214)	0.6982 (0.2369)
FF-T S	0.3690 (0.2648)	0.7156 (0.1912)	0.5802 (0.2010)	0.6965 (0.1835)
FF-BR	0.5833 (0.2880)	0.8206 (0.1824)	0.5567 (0.2023)	0.7614 (0.1892)
FF-T S/BR	0.4405 (0.2992)	0.6893 (0.1430)	0.6049 (0.2328)	0.7885 (0.2079)
DTC	0.4841 (0.3267)	0.8558 (0.2838)	0.3304 (0.1737)	0.7449 (0.1860)
NTC	0.5595 (0.3100)	0.7841 (0.2629)	0.6255 (0.2004)	0.7980 (0.1748)

Note. For all four measures, smaller scores represent more accurate ratings. FB = Feedback; FF = Feedforward; TS = Target score; BR = Behavioral rationale; TS/BR = Combination target score and behavioral rationale; DTC = Dimension training control; NTC = No-Training control.

<sup>a</sup>

Standard deviations are included in parentheses.

differential accuracy measure, and therefore less accurate, than participants who received feedback ( $\underline{M} = 0.6671$ ).

The analyses also indicated that for all four accuracy components, the no-training control group did not significantly differ from the experimental groups ( $\underline{p} > .05$ ). The dimension-training control group, however, was more accurate than the experimental groups on the stereotype accuracy measure ( $\underline{p} < .01$ ), and less accurate on the differential elevation measure ( $\underline{p} < .05$ ). These analyses are summarized in Tables 11, 12, 13, and 14.

For each of the four components of accuracy, planned comparisons were made between the mean of the feedforward conditions and both of the control groups. For elevation, differential elevation, and differential accuracy, no differences were detected between the feedforward conditions and either of the control conditions ( $\underline{p} > .05$ ). However, a significant difference was found on the stereotype accuracy measure for the comparison of the feedforward groups and the dimension-training control ( $F(1,47) = 12.70$ ,  $\underline{p} < .01$ ). Examination of the means indicated that the feedforward conditions ( $\underline{M} = .5806$ ) had significantly larger stereotype accuracy scores than did the dimension-training control ( $\underline{M} = .3304$ ). Since smaller stereotype accuracy scores represent more accurate ratings, these findings suggest that the feedforward conditions were less accurate than the dimension-training control condition.

In summary, the results of the analyses on Cronbach's (1955) components of accuracy do not lend support for either of the hypotheses being evaluated in this study. Since no significant effects were indicated for information type, there is no support for



Table 11

Summary of the Analysis of Variance on Cronbach's Elevation Measure.

Source	DF	MS	F-Ratio	VC	ICC
IT	2	0.1476	1.51	.00034	.0035
M	1	0.0787	0.80	-.00007	.0000
IT x M	2	0.0331	0.34	-.00023	.0000
DTC	1	0.0288	0.29	-.00014	.0000
NTC	1	0.1693	1.73	.00014	.0015
Error	88	0.0980		.09800	

Note. If a source's variance component was negative, that value was used in the denominator to compute intraclass correlation coefficients, but the source's coefficient was set to zero. VC = variance component; ICC = intraclass correlation coefficient; IT = information type; M = mode; DTC = dimension-training control; NTC = no-training control.

Table 12

Summary of the Analysis of Variance on Cronbach's Differential Elevation Measure.

Source	DF	MS	F-Ratio	VC	ICC
IT	2	0.0279	0.53	-.00017	.0000
M	1	0.0710	1.35	.00006	.0012
IT x M	2	0.0902	1.71	.00026	.0461
DTC	1	0.2174	4.12*	.00033	.0063
NTC	1	0.0559	1.06	.00001	.0001
Error	88	0.0527		.05270	

Note. If a source's variance component was negative, that value was used in the denominator to compute intraclass correlation coefficients, but the source's coefficient was set to zero. VC = variance component; ICC = intraclass correlation coefficient; IT = information type; M = mode; DTC = dimension-training control; NTC = no-training control.

\*p < .05.

Table 13

Summary of the Analysis of Variance on Cronbach's Stereotype Accuracy Measure.

Source	DF	MS	F-Ratio	VC	ICC
IT	2	0.0550	1.24	.00007	.0016
M	1	0.0310	0.70	-.00005	.0000
IT x M	2	0.0155	0.35	-.00020	.0000
DTC	1	0.5413	12.21*	.00101	.0223
NTC	1	0.0444	1.00	.00000	.0000
Error	88	0.0443		.04430	

Note. If a source's variance component was negative, that value was used in the denominator to compute intraclass correlation coefficients, but the source's coefficient was set to zero. VC = variance component; ICC = intraclass correlation coefficient; IT = information type; M = mode; DTC = dimension-training control; NTC = no-training control.

\* $p < .01$ .

Table 14

Summary of the Analysis of Variance on Cronbach's Differential Accuracy Measure.

Source	DF	MS	F-Ratio	VC	ICC
IT	2	0.0322	0.96	-.00001	.0000
M	1	0.1202	3.57*	.00030	.0089
IT x M	2	0.0041	0.12	-.00021	.0000
DTC	1	0.0141	0.42	-.00004	.0000
NTC	1	0.0833	2.48	.00010	.0297
Error	88	0.0337		.03370	

Note. If a source's variance component was negative, that value was used in the denominator to compute intraclass correlation coefficients, but the source's coefficient was set to zero. VC = variance component; ICC = intraclass correlation coefficient; IT = information type; M = mode; DTC = dimension-training control; NTC = no-training control.

\* $p < .10$ .

the hypothesis that rater training involving target score information would be less accurate than training providing only a behavioral rationale for ratings (i.e., Hypothesis 1). The hypothesis that training providing feedforward information would be more accurate than control training (i.e., Hypothesis 2) also was not supported. In fact, the feedforward groups were found to be significantly less accurate than the dimension-training control group. Finally, the finding of no significant effects for information mode on the accuracy components suggests no difference in the effectiveness of providing information in the form of feedforward or feedback.

#### Extended Accuracy Design

Analysis of variance procedures were used to evaluate ratings with the extended accuracy design (Dickinson, 1987). This design included the factors from the basic accuracy design (i.e., rating sources, ratees, and dimensions), as well as the interaction of these factors with the between-subjects experimental variables of information type, information mode, and contrasts for both the dimension-training control and no-training control with the mean of the experimental conditions. The interactions of these additional variables with the factors in the basic accuracy design allow for assessment of the differential effects of the experimental treatments on the accuracy of the ratings. A summary of the results from this analysis is included in Table 15.

The results from this analysis suggested that there were significant effects for all of the accuracy components in the basic accuracy design ( $p < .05$ ). These significant effects were for rating sources (elevation), rating sources x ratees (differential elevation),

Table 15

Summary of the Analysis of Variance on the Extended Accuracy Design.

Source	DF	MS	F-Ratio	VC	ICC
IT	2	1.2821	0.88 <sup>a</sup>	-.00003	.0000
M	1	1.1124	0.77 <sup>a</sup>	-.00003	.0000
IT x M	2	1.2345	0.56 <sup>a</sup>	.00005	.0000
DT C	1	0.0370	0.02 <sup>a</sup>	-.00011	.0000
NT C	1	2.2451	1.06 <sup>a</sup>	.00001	.0000
Rr/IT x M x DT C x NT C	88	2.6289		.00193	.0011
S (rating sources)	1	180.4586	5.40* <sup>a</sup>	.01215	.0069
S x IT	2	1.2821	0.89 <sup>a</sup>	-.00002	.0000
S x M	1	1.1124	0.77 <sup>a</sup>	-.00003	.0000
S x DT C	1	0.0370	0.02 <sup>a</sup>	-.00021	.0000
S x NT C	1	2.2451	1.06 <sup>a</sup>	.00001	.0000
S x IT x M	2	1.2345	0.56 <sup>a</sup>	-.00016	.0000
S x Rr/IT x M x DT C x NT C	88	2.6289		.00010	
D (dimensions)	2	178.1646	3.24 <sup>a</sup>	.02037	.0115
D x IT	4	0.2781	0.30 <sup>a</sup>	-.00022	.0000
D x M	2	0.3079	0.46 <sup>a</sup>	-.00006	.0000
D x DT C	2	2.8090	4.26* <sup>a</sup>	.00021	.0001
D x NT C	2	0.0171	0.03 <sup>a</sup>	-.00005	.0000
D x IT x M	4	0.1055	0.14 <sup>a</sup>	-.00021	.0000
D x Rr/IT x M x DT C x NT C	176	1.0754		.00116	

Table 15 (continued)

Source	DF	MS	F-Ratio	VC	ICC
R (ratees)	6	360.1201	249.38**	.20757	.1175
R x IT	12	0.2597	0.18	-.00206	.0000
R x M	6	0.2513	0.17	-.00138	.0000
R x DTC	6	1.0355	0.72	-.00028	.0000
R x NTC	6	0.9420	0.65	-.00034	.0000
R x IT x M	12	1.0076	0.70	-.00152	.0000
R x Rr/IT x MxDTCxNTC	528	1.4441		.00526	
S x D	2	120.0884	8.28* <sup>a</sup>	.01746	.0000
S x D x IT	4	0.2781	0.30 <sup>a</sup>	-.00022	.0000
S x D x M	2	0.3079	0.46 <sup>a</sup>	-.00006	.0000
S x D x DTC	2	2.8090	4.26* <sup>a</sup>	.00021	.0001
S x D x NTC	2	0.0171	0.03 <sup>a</sup>	-.00005	.0000
S x D x IT x M	4	0.1055	0.14 <sup>a</sup>	-.00021	.0000
S x D x Rr/IT x MxDTCxNTC	176	1.0753		.00232	.0013
S x R	6	32.2535	22.34**	.03566	.0202
S x R x IT	12	0.2597	0.18	-.00411	.0000
S x R x M	6	0.2513	0.17	-.00276	.0000
S x R x DTC	6	1.0355	0.72	-.00055	.0000
S x R x NTC	6	0.9420	0.65	-.00068	.0000
S x R x IT x M	12	1.0076	0.70	-.00303	.0000
S x R x Rr/ IT x MxDTCxNTC	528	1.4441		.01053	.0060

Table 15 (concluded)

Source	DF	MS	F-Ratio	VC	ICC
D x R	12	54.5617	79.53**	.09353	.0529
D x R x IT	24	0.5532	0.81	-.00069	.0000
D x R x M	12	0.2859	0.42	-.00139	.0000
D x R x DTC	12	0.2706	0.39	-.00866	.0000
D x R x NTC	12	0.1751	0.26	-.01065	.0000
D x R x IT x M	24	0.3532	0.51	-.00347	.0000
D x R x Rr/ IT x M x DTC x NTC	1056	0.6861		.68610	.3884
S x D x R	12	14.1188	20.58**	.04664	.0264
S x D x R x IT	24	0.5532	0.81	-.00138	.0000
S x D x R x M	12	0.2859	0.42	-.00278	.0000
S x D x R x DTC	12	0.2706	0.39	-.00289	.0000
S x D x R x NTC	12	0.1751	0.26	-.00355	.0000
S x D x R x IT x M	24	0.3532	0.51	-.00694	.0000
S x D x R x Rr/ IT x M x DTC x NTC	1056	0.6861		.68610	

Note. If a source's variance component was negative, that value was used in the denominator to compute intraclass correlation coefficients, but the source's coefficient was set to zero. VC = variance component; ICC = intraclass correlation coefficient; IT = information type; M = mode; DTC = dimension-training control; NTC = no-training control, Rr = raters.

<sup>a</sup> Quasi-F-Ratio.

\* $p < .05$ . \*\* $p < .01$ .



rating sources x dimensions (stereotype accuracy), and rating sources x ratees x dimensions (differential accuracy). It should be noted that significant effects for each of these four accuracy components suggest inaccuracies in the ratings.

Of particular concern in this analysis were the interactions of the accuracy effects (i.e., elevation, differential elevation, stereotype accuracy, differential accuracy) with the between-subjects experimental variables (i.e., information type, information mode, and the two control conditions). One significant interaction between the accuracy effects and experimental variables was obtained ( $p < .05$ ). This interaction represented the effect for stereotype accuracy differing by dimension-training control (Rating Sources x Dimensions x Dimension-Training Control). Post hoc analysis of this three-way interaction with the Newman-Keuls procedure suggested that the dimension-training control group exhibited greater stereotype accuracy than did the experimental groups. For the mean ratings from the experimental groups, there were significant differences ( $p < .05$ ) between the participants' ratings and the target scores on two of the three dimensions (i.e., problem analysis and problem solution). However, for the dimension-training control group, no significant differences were found between the participants' ratings and the target scores on any of the dimensions ( $p > .05$ ).

Planned comparisons were also conducted to examine differences on each of the accuracy components between the feedforward groups and each of the two control conditions. These comparisons were made by analyzing the interactions of the sources of variance representing the accuracy components and contrasts of the feedforward and control

conditions. The interaction of rating sources and dimensions with the contrast of the feedforward and dimension-training control groups was the only significant effect ( $F(2,52) = 4.68, p < .05$ ). This interaction suggested different amounts of stereotype accuracy for the feedforward and dimension-training control conditions.

Newman-Keuls post hoc analysis of the interaction indicated that for the feedforward groups there were significant differences between the ratings and target scores on the problem analysis and problem solution dimensions ( $p < .05$ ). For the dimension-training control, significant differences between the ratings and target scores were not detected for any of the three dimensions ( $p > .05$ ). These findings suggest that the dimension-training control group was more accurate on the stereotype accuracy measure than the feedforward groups.

In summary, the results from the extended accuracy design do not provide support for either of the two hypotheses. No support was provided for Hypothesis 1 since none of the interactions with information type were found to be significant. Furthermore, Hypothesis 2 was not supported since the only difference that was detected with the comparisons of feedforward and the control groups indicated that the feedforward groups were less accurate (i.e., stereotype accuracy). With regards to the comparison of the effectiveness of feedforward and feedback, no differences were indicated with the analysis of the extended accuracy design.

#### Evaluation of the Post-Experimental Questionnaire

Responses to each of the eleven items on the post-experimental questionnaire measuring the participants' reactions to the study were analyzed with a 3 X 2 analysis of variance (Information Type x

Information Mode). In addition, for each of the items, contrasts were examined between each of the two control groups and the mean of the experimental groups. No significant effects were found for information type, mode, or the type by mode interaction on any of the eleven items ( $p > .05$ ).

Analysis of the contrasts between each of the control groups and the mean of the experimental groups indicated that there were significant differences between the no-training control and the experimental groups on two of the items (i.e., items 7 and 12). Participants in the no-training control group indicated that they felt the information presented in session one was less understandable ( $F(1,82) = 5.31, p < .05$ ), and more confusing ( $F(1,82) = 4.69, p < .05$ ) than did the experimental groups.

## CHAPTER FOUR

### Discussion

#### Information Type

The first objective of this study was to examine systematically the contributions that different types of information presented in rater training programs make towards improving the accuracy of performance ratings. Of particular interest were rater training programs based on providing raters with a frame-of-reference in common with the experts that define the standards for performance (Bernardin & Buckley, 1981). Training programs based on the frame-of-reference notion have incorporated information concerning expert raters' target scores and the behavioral rationale that the target scores are based on. Since previous research has provided mixed results, this study was an attempt to understand the effectiveness of the frame-of-reference training approach. It was hoped that examining the different types of information that have been used with this training approach would allow for determination of the individual contributions that the different components make towards improving the accuracy of ratings.

It was hypothesized that the use of target score information (i.e., outcome information) within a rater training program would have a debilitating effect on the effectiveness of the training. This hypothesis was based on findings within the MCPL paradigm which have suggested that outcome information impedes the learning of complex

cognitive judgment tasks.

The results from the study suggest that the different types of information presented in the rater training programs did not have differential effects on the accuracy of the performance ratings. No significant effects for information type were indicated for correlational accuracy per ratee, correlation accuracy per dimension, or any of Cronbach's (1955) accuracy measures. Furthermore, in the extended analysis of variance accuracy design, no significant interactions were found which would indicate different levels of accuracy for the different types of training information.

These results are inconsistent with the results reported in previous research. Several researchers have reported that information-value feedback is more effective for teaching complex judgment tasks than is outcome feedback (Adelman, 1981; Hammond & Summers, 1965; Lindell, 1976; Newton, 1965; Nystedt & Magnusson, 1973; Summers & Hammond, 1966; Todd & Hammond, 1965). All of this previous research has been conducted within the MCPL paradigm.

Perhaps the inconsistency between the results of the present study and previous research is a question of generalizability of the MCPL paradigm task to the performance rating task. While the performance rating task is certainly analogous to the MCPL task, clear differences between the two can be identified. For instance, in the MCPL studies mentioned above, the cues that were presented to subjects were numerical in nature. In the present study, the cues were actual behaviors. Anderson (1977) compared numerical cues with verbal cues and found that the verbal cues were rated significantly less consistently than were the numerical ones. It is possible that

behavioral cues also might cause inconsistency.

Another difference between the MCPL and rating task is that with the MCPL task the cues are readily identifiable. In the rating task, the behaviors (or cues) relevant to the dimensions of interest, were presented in the midst of a stream of other irrelevant behaviors. The requirement to distinguish between relevant and irrelevant information adds considerably more complexity to the rating task.

In addition, the cognitively-oriented information about the task that is presented in typical MCPL research is different than the information presented in the present study. Typically, MCPL research includes such information as the statistical relationships between the cues and criteria, or the relative weighting of the cues. In the present study, the information about the task that was presented was less quantitative, and merely identified which behaviors were relevant for the different dimensions. More detailed information concerning how relevant behaviors were weighted in determining target ratings for each dimension was unavailable.

One further issue concerning the generalizability of the MCPL paradigm to performance rating tasks needs to be addressed. The MCPL paradigm focuses on the process of evaluation. That is, the concern is on how cues are evaluated to arrive at some criterion response. Evaluation, however, is only one of the processes which is involved in performance rating. In its simplest case, performance rating involves both observation and evaluation (Thornton & Zorich, 1980). In more complicated cases, additional cognitive processes such as encoding or recall also are involved (Cooper, 1981; Denisi, et al., 1984; Feldman, 1981). To focus on the process of evaluation without addressing the

influence of these other variables may serve to create artificial problems.

It should be noted that the present study is one of several studies that have not found frame-of-reference training to improve the accuracy of performance ratings (Dickinson & Silverhart, 1986; Hassett, 1987; McIntyre & Athey, 1983; Silverhart & Dickinson, 1986a, 1986b). Important questions remain to be answered concerning the circumstances under which this training approach will be effective.

#### Mode of Information Presentation

A second objective of the study was to investigate different modes of presenting information during rater training. While successful rater training programs have typically presented information in the form of feedback, studies in the MCPL paradigm suggest that presenting information in the form of feedforward may be an effective strategy for learning complex cognitive tasks. It was hypothesized that the use of feedforward would be an effective training strategy and would result in improved rater accuracy when compared with control conditions. In addition, this study allowed for exploratory comparisons between the use of feedforward and feedback.

The results from the study generally do not support the hypothesis that the use of feedforward would be an effective training strategy for improving the accuracy of performance ratings. When the mean of the feedforward conditions was compared with the control groups, no differences were found on either of the correlational accuracy measures. With both Cronbach's accuracy statistics and the extended accuracy design, significant effects were found for stereotype accuracy. The difference for stereotype accuracy with both

of these approaches, however, indicated that the feedforward conditions were significantly less accurate than the dimension-training control.

Examination of the means for the groups receiving feedforward on the correlational accuracy per ratee measure reveals an interesting finding relevant to the effectiveness of feedforward. The group that received target score feedforward information was very accurate on the correlational accuracy per ratee measure. The mean correlational accuracy per ratee score for the target score/ feedforward group ( $\bar{Z} = 1.0393$  or  $\bar{r} = .77$ ) is notably larger than the values reported in previous studies (Hoffman & Dossett, 1984,  $\bar{r} = .41$ ; McIntyre et al., 1984,  $\bar{r} = .55$ ; Dickinson & Silverhart, 1986,  $\bar{r} = .38$ ). This result suggests that rater training programs which present target score information in the form of feedforward may be effective at improving performance ratings. In light of this finding, additional research investigating feedforward as a training strategy is warranted.

In terms of the comparison between the effectiveness of feedforward and feedback, the results of the study do not justify strong conclusions. The groups which received feedback were found to be significantly more accurate on the correlational accuracy per dimension measure than those receiving feedforward. The corresponding effect, however, was not found with either Cronbach's stereotype accuracy, or the rating sources by dimensions by mode interaction in the extended accuracy design (i.e., stereotype accuracy). Consistency with regards to the results on these three measures would be expected since they all measure the use of dimensions across ratees.



### Factors Contributing to the Results

The results from this study suggest that the type and mode of information presented in rater training do not differentially effect the accuracy of performance ratings. One contributing factor for this finding is suggested by the evaluations of dimension training which was given to all experimental groups and the dimension-training control group. The purpose of the dimension training was to provide the participants with an understanding of how the behavioral dimensions were being defined and how the rating forms should be used to make ratings on each of the dimensions. The difference between the experimental groups and the no-training control group that was detected at the end of training (i.e., the post-training test) was not detected the following day before the rating task was performed (i.e., the pre-task test). While this finding may be attributable to insensitivity of the measures, it is possible that the effect of the dimension training diminished during the time between when the training was administered and the rating task performed. Although diminishing effects for training have previously been reported (Goldstein, 1980), the conclusion that information loss was responsible for the present results is alarming since the time in question is so short.

In another study which used similar dimension training (Johnson, 1987), differences between experimental and control groups were found on both post-training and pre-task evaluations. Johnson's (1987) study differed from the present study in that a behavioral checklist was used as a heuristic in the dimension training, and the dimension training was combined with rater training based on a cognitive

modeling strategy. Perhaps the checklist is a necessary heuristic. In addition, perhaps the cognitive modeling approach to rater training is more effective than the frame-of-reference approach in the present study. Future research should investigate these differences.

A second factor which may have contributed to the results of this study is the rating forms that were utilized. Comments on the post-experimental questionnaire indicated that participants had difficulty differentiating between the effectiveness of various anchors on the BARS forms and therefore had difficulty in selecting appropriate ratings. Rigorous scaling procedures, based on Taylor's (1968) ranking technique, were utilized in the development of these rating forms (see Campbell, 1986). While this ranking method required agreement among raters as to the relative ordering of various behaviors that were selected as anchors, it did not necessitate the selection of anchors representing a full range of effective and ineffective behaviors. Perhaps the use of a rating procedure in the development of the rating forms would produce anchors which are more easily differentiated.

Finally, a third factor that may have contributed to the present results involves a combination of the participants' motivation and the complexity of the rating task. It was hoped that compensating the participants for their time would provide sufficient motivation to ensure attentiveness to the training materials and the rating stimuli. Remuneration, however, was contingent only on being physically present during the training and the rating sessions. If the perception could be developed in participants that the compensation was at least in part contingent on their performance, then one could be more confident

that sufficient motivation did exist to ensure attentiveness. Motivation of the participants is extremely important in performance rating research because of the complexity of the task that is involved. Perhaps, even with motivated participants, the nature of the task is too complicated to train naive individuals to perform in single sessions that last two to three hours.

#### Analytical Approaches to Measuring Accuracy

An issue raised by this research concerns what is the most appropriate approach to analyze the accuracy of performance ratings. This issue is quite complicated, since in addition to the three approaches that were used in this study, other approaches (e.g., signal detection theory; Lord, 1985) and statistics (e.g., distance accuracy) are endorsed in the literature. With the existence of several approaches for measuring the accuracy of ratings, there is a strong need for description of the relative merits and shortcomings of the various approaches.

Of the approaches used in the present study, a question exists as to the rigor of the correlational accuracy measures. One area of concern relates to the stability of these measures, especially correlational accuracy per ratee. Correlations based on a small number of scores are likely to be unreliable. In the present study correlational accuracy per ratees was computed based on only three sets of ratings and correlational accuracy per dimensions was based on seven.

In addition, since these measures are based on correlations, variance is assumed both in participants' ratings and the target scores. In cases where no variance is present in a set of ratings,

the accuracy score becomes zero. In the present study, there were a substantial number of cases (16 percent) in which participants provided the same numerical rating on all three dimensions for a particular ratee and thus received a correlational accuracy per ratee score of zero.

Both Cronbach's accuracy statistics and the extended accuracy designs allow researchers to evaluate the manner in which experimental manipulations affect Cronbach's (1955) accuracy components. In the present study, the two approaches resulted in relatively consistent findings. Both approaches detected that with stereotype accuracy the dimension-training control group was more accurate than the experimental groups and the feedforward groups. A significant advantage with the extended accuracy design is that when differences are detected, the nature of the differences can be investigated with post hoc analyses. In the present study, post hoc analyses indicated that for stereotype accuracy there were not significant differences between the ratings and target scores on any of the dimensions for the dimension-training control group. For the mean of the experimental groups and the feedforward groups, however, there were significant differences between the ratings and target scores on the problem analysis and problem solution dimensions. The ability to trace inaccuracies in the ratings to such specific levels makes a significant contribution to an understanding of the ratings.

It should be noted that the analyses with Cronbach's accuracy statistics detected one significant finding that was not detected with the extended accuracy design. This finding was that the experimental groups had significantly greater differential elevation than the

dimension-training control group. Differences in the statistical power that each of the approaches to measuring the accuracy of ratings possess is an area that future researchers will have to explore.

#### Practical Implications

Implications of the results of this study are limited by the paucity of differences that were detected. Two findings, however, do have implications that are worthy of being mentioned. These findings were the significant effect for ratees on the correlational accuracy per ratee measure and the significant effect for dimensions on the correlational accuracy per dimension measure. These significant effects suggest that some ratees and some dimensions may be more difficult to make accurate ratings on than others. When developing a rater training program, information concerning particular behavioral dimensions, or the characteristics of ratees that are difficult to rate could be useful in designing training content to deal with these difficulties.

#### Summary and Conclusions

In summary, the results from this study indicate that the rater training was not effective at improving the accuracy of performance ratings. Support was not found for the hypothesis that training with target score information, or target score information combined with a behavioral rationale for the expert ratings would result in less accurate ratings than training only incorporating the behavioral rationale. Since no differences were found between any of the types of training information, conclusions about their unique contribution to the effectiveness of training cannot be made. Support was also not found for the hypothesis that ratings would be more accurate when

ratets receive training information by means of feedforward than when control training was provided in which the training information was not presented.

### References

- Adelman, L. (1981). The influence of formal, substantive, and contextual task properties on the relative effectiveness of different forms of feedback in multiple-cue probability learning tasks. Organizational Behavior and Human Performance, 27, 423-442.
- Anderson, B.L. (1977). Differences in teachers' judgment policies for varying numbers of verbal and numerical cues. Organizational Behavior and Human Performance, 19, 68-88.
- Annett, J. (1961). The role of knowledge of results in learning: A survey. U.S. NAVTRACDEVCEEN, Technical Document Report 342-3.
- Annett, J. (1969). Feedback and human behavior. Baltimore, MD: Penguin Books.
- Ammons, R.B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formulation. The Journal of General Psychology, 54, 279-299.
- Banks, C.G., & Murphy, K.R. (1985). Toward narrowing the research-practice gap in performance appraisal. Personnel Psychology, 38, 335-345.
- Bernardin, H.J., & Buckley, M.R. (1981). Strategies in rater training. Academy of Management Review, 6, 205-212.
- Bernardin, H.J., & Pence, E.C. (1980). Effects of rater training: Creating new response sets and decreasing accuracy. Journal of Applied Psychology, 65, 60-66.
- Bingham, W.V. (1939). Halo, invalid and valid. Journal of Applied Psychology, 23, 221-228.

- Bjorkman, M. (1972). Feedforward and feedback as determiners of knowledge and policy: Notes on a neglected issue. Scandinavian Journal of Psychology, 13, 152-158.
- Bogart, D.H. (1980). Feedback, feedforward, and feedwithin: Strategic information in systems. Behavioral Science, 25, 237-249.
- Borman, W.C. (1975). Effects of rating instructions to avoid halo errors on reliability and validity of performance evaluation ratings. Journal of Applied Psychology, 60, 556-560.
- Borman, W.C. (1977). Consistency of rating accuracy and rating errors in the judgment of human performance. Organizational Behavior and Human Performance, 20, 258-272.
- Borman, W.C. (1979). Format and training effects on rating accuracy and rater errors. Journal of Applied Psychology, 64, 410-421.
- Brunswick, B. (1943). Organismic achievement and environmental probability. Psychological Review, 50, 255-272.
- Brunswick, B. (1952). The conceptual framework of psychology. In the International Encyclopedia of Unified Science, Vol. 1, No. 10.
- Campbell, W. (1986). Construct validation of role playing exercises in an assessment center using BARS and behavioral checklist formats. Unpublished masters thesis, Old Dominion University.
- Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multi-trait-multimethod matrix. Psychological Bulletin, 56, 81-105.



- Cardy, R.L., & Kehoe, J.F. (1984). Rater selective attention ability and appraisal effectiveness: The effect of a cognitive style on the accuracy of differentiation among ratees. Journal of Applied Psychology, 69, 589-594.
- Cooper, W.H. (1981). Conceptual similarity as a source of illusory halo in job performance ratings. Journal of Applied Psychology, 66, 302-307.
- Cronbach, L.J. (1955). Processes affecting scores on understanding of others and assumed "similarity." Psychological Bulletin, 52, 177-193.
- DeNisi, A., Meglino, B., & Cafferty, T.P. (1984). A cognitive view of the performance appraisal process: A model and research propositions. Organizational Behavior and Human Performance, 33, 360-396.
- Dickinson, T.L. (1986). Accuracy of work performance measurement: Target scores, statistics, and research issues. AFHRL Technical Paper, Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Dickinson, T.L. (1987). Designs for evaluating the validity and accuracy of performance ratings. Organizational Behavior and Human Decision Making, 40, 1-21.
- Dickinson, T.L., & Silverhart, T.A. (1986). Training to improve the accuracy and validity of performance ratings. Paper presented at the 94th annual convention of the American Psychological Association, Washington, D.C.

- Dudycha, A.L., Dudycha, L.W., & Schmitt, N.M. (1973). The learning of negative cue validities as a function of statistical sophistication and instruction. Paper presented at the 45th annual meeting of the Midwestern Psychological Association, Chicago.
- Feldman, J.M. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. Journal of Applied Psychology, 66, 127-148.
- Fisher, R.A. (1935). The fiducial agreement in statistical inference. Annals of Eugenics, 6, 391-398.
- Goldstein, I.L. (1980). Training in work organizations. Annual Review of Psychology, 30, 229-272.
- Hammond, K.R. (1971). Computer graphics as an aid to learning. Science, 172, 903-908.
- Hammond, K.R., & Summers, D.A. (1965). Cognitive dependence on linear and nonlinear cues. Psychological Review, 72, 215-224.
- Hammond, K.R., & Summers, D.A. (1972). Cognitive control. Psychological Review, 79, 58-67.
- Hammond, K.R., Summers, D.A., & Deane, D.H. (1973). Negative effects of outcome-feedback in multiple-cue probability learning. Organizational Behavior and Human Performance, 9, 30-34.
- Hassett, C.E. (1987). The influence of rater training and purpose of rating on the accuracy of ratings. Unpublished manuscript.
- Hedge, J.W. (1982). Improving the accuracy of performance evaluations: A comparison of the methods of performance appraiser training. Unpublished doctoral dissertation, Old Dominion University.

- Hendrix, W.H., & Dudycha, A.L. (1981). Feedforward and feedback in multiple cue probability learning - facilitating or debilitating? Journal of Experimental Education, 49, 137-146.
- Hoffman, C.C., & Dossett, D.L. (1984, August). Test of the assumed superiority of the expert rater. Paper presented at the annual meeting of the American Psychological Association, Toronto.
- Holt, R.R. (1958). Clinical and statistical prediction: A reformation and some new data. Journal of Abnormal and Social Psychology, 56, 1-12.
- Hursch, C.J., Hammond, K.R., & Hursch, J.L. (1964). Some methodological considerations in multiple-cue probability studies. Psychological Review, 71, 42-60.
- Ilgen, D.R., Fisher, C.D., & Taylor, M.S. (1979). Consequences of individual feedback on behavior in organizations. Journal of Applied Psychology, 64, 349-371.
- Johnson, R.L., Jr. (1987). The influence of direct versus indirect observation, candidate report format, and assessor training on the accuracy of assessor ratings. Doctoral Dissertation, Old Dominion University.
- Jones, M.B. (1968). Correlation as a dependent variable. Psychological Bulletin, 70, 69-72.
- Landy, F.J., & Farr, J.L. (1980). Performance rating. Psychological Bulletin, 87, 72-107.
- Landy, F.J., & Farr, J.L. (1983). The measurement of work performance: Methods, theory, and applications. New York: Academic Press.

- Landy, F.J., & Trumbo, D.A. (1980). The psychology of work behavior (rev. ed.). Homewood, Ill.: Dorsey Press.
- Latham, G.P., Wexley, K.N., & Pursell, E.D. (1975). Training managers to minimize rating errors in the observation of behavior. Journal of Applied Psychology, 60, 550-555.
- Lindell, M.K. (1976). Cognitive and outcome feedback in multiple-cue probability learning tasks. Journal of Experimental Psychology: Human Learning and Memory, 2, 739-745.
- Lord, R.G. (1985). Accuracy in behavioral measurement: An alternative definition based on raters' cognitive schema and signal detection theory. Journal of Applied Psychology, 70, 66-71.
- Magnusson, D., & Nystedt, L. (1971). Cue relevance and feedback in a clinical prediction task. Report No. 272 from the Psychological Laboratories, University of Stockholm, cited by P. Slovic and S. Lichtenstein, Comparison of bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 6, 649-744.
- McIntyre, R.M., & Athey, T. (1985, August). Rater training research: Group size and other methodological considerations. Paper presented at the annual meeting of the Academy of Management, San Diego.
- McIntyre, R.M., Smith, D.E., & Hassett, C.E. (1984). Accuracy of performance ratings as affected by rater training and perceived purpose of rating. Journal of Applied Psychology, 69, 147-156.

- Murphy, K.R., Garcia, M., Kerkar, S., Martin, C., & Balzar, W.K. (1982). Relationship between observational accuracy and accuracy in evaluating performance. Journal of Applied Psychology, 67, 320-325.
- Newton, J.R. (1965). Judgment and feedback in a quasi-clinical situation. Journal of Personality and Social Psychology, 1, 336-342.
- Nystedt, L., & Magnusson, D. (1973). Cue relevance and feedback in a clinical prediction task. Organizational Behavior and Human Performance, 9, 100-109.
- Pulakos, E.D. (1984). A comparison of training programs: Error training and accuracy training. Journal of Applied Psychology, 69, 581-588.
- Pulakos, E.D. (1986). The development of training programs to increase accuracy with different rating tasks. Organizational Behavior and Human Decision Making, 38, 76-91.
- Silverhart, T.A., & Dickinson, T.L. (1985a). The effects of training and rating formats on the accuracy of performance ratings. Paper presented at the meeting on the Southeastern Psychological Association, Atlanta.
- Silverhart, T.A., & Dickinson, T.L. (1985b). Training to improve ratings with the mixed standard scale: An investigation of cognitive strategy and rater accuracy training. Paper presented at the meeting of the Eastern Psychological Association, Boston.

- Smith, D.E. (1984). Increasing rater accuracy through rater training: A look at the performance appraisal process. (Doctoral Dissertation, Colorado State University). Dissertation Abstracts International, 45, 1614-B.
- Smith, D.E. (1986). Training programs for performance appraisal: A review. Academy of Management Review, 11, 22-40.
- Smith, D.E., Hassett, C.E., & McIntyre, R.M. (1982). Using student ratings for administrative decisions: Are ratings contaminated by perceived uses of the information? Paper presented at the meeting of the Western Academy of Management, Colorado Springs, CO.
- Spool, M.D. (1978). Training programs for observers of behavior: A review. Personnel Psychology, 31, 853-888.
- Summers, D.A., & Hammond, K.R. (1966). Inference behavior in multiple-cue tasks involving both linear and non-linear relations. Journal of Experimental Psychology, 71, 751-757.
- Thornton, G.C. III, & Zorich, S. (1980). Training to improve observer accuracy. Journal of Applied Psychology, 65, 351-354.
- Todd, F.J., & Hammond, K.R. (1965). Differential feedback in two multiple-cue probability learning tasks. Behavioral Science, 10, 429-435.
- Tucker, L.R. (1964). A suggested alternative formulation in the developments by Hursch, Hammond and Hursch, and by Hammond, Hursch and Todd. Psychological Review, 71, 528-530.
- Vance, R.J., Kuhnert, K.W., & Farr, J.L. (1978). Interview judgments: Using external criteria to compare behavioral and graphic rating scale ratings. Organizational Behavior and Human Performance, 22, 279-294.

- Vaughan, G.M., & Corballis, M.C. (1969). Beyond tests of significance: Estimating strength of effects in selected ANOVA designs. Psychological Bulletin, 72, 204-213.
- Wexley, K.N., & Latham G.P. (1981). Develping and training human resources in organizations. Glenview, Illinois: Scott Foresman and Company.
- Wherry, R.J., & Bartlett, C.J. (1982). The control of bias in ratings: A theory of rating. Personnel Psychology, 35, 521-551.
- Zellinger, P.M. (1981). Effects of discussion and type of feedback on group judgments in multiple cue probability learning. Unpublished doctoral dissertation, Colorado State University.

Appendix A:  
Scripts for Interview Simulations



## Script for Training Simulation #1

T: So how do you like working here at our store here?

D: Good. It's a lot busier than what I'm used to; but, generally I like it pretty well.

T: How have you adjusted to the big city life?

D: It's good. Again, it is alot more crowded, but it's, it's fine.

T: Ok, let's talk about a few things here. Overall, you have done a pretty decent job, but there is some room for improvement. That is why we are here, not to criticize or anything, but what we're trying to do here is talk about a few things and hope that we can build for the future, to improve on everyone's performance, not just yours or mine but everybody's. Everybody needs to open up the lines of communication. The first thing I noticed is that you need to delegate some of your responsibilities a little more thoroughly. You seem to have trouble delegating. You seem to want to have a hands-on approach to accomplishing the tasks in your department. I'm sure that at your last job your department was a lot smaller and you had to take a hands-on approach and assume alot of these responsibilities. Here we would like you to take the role of supervisor. What we would like you to do is delegate and let the others do the work, and just guide them along in their duties, not so much to do them yourself and assume the responsibilities.

D: Well, I try to do that.

T: Ok, well, what we would like to see in the future is for you to expand on that role. Delegate some of the decision-making. The lesser decisions should go to some people in your department. Um, that way we can see how they do. Sort of groom them along, and the only way that you are going to move up is to groom someone who can take your position.

D: Well, I'll try to do that. I just hope that they do the work when I tell them to.

T: Well, that's it. It's not so much telling them as it is teaching them. You know how to do the job. You do it very well. The trick now is for you to teach someone else, your subordinates. Delegate the responsibility to them. Let them make the decisions and teach them so you can move up in the organization.

D: I will try.

T: OK, good. I noticed when I observed you that sometimes you need a

little more patience in dealing with your employees. A lot of times they don't know as much as you, and it is frustrating. I know with myself one of my biggest problems is trying to teach people things because I don't have a lot of patience, but it is something that we're all going to have to work on. We have to try. What you need to do is give them the benefit of your years of experience and training and then you can impart that on the people that work for you. That way they will be better workers, so when you are away from the job the person you leave in charge, you'll know can handle the job so when you come back after the weekend you know that everything will be in order.

D: Well, I've been trying to do that. I just have some people who don't want to work.

T: Ok, well, do you think there are some people in your department who don't belong there?

D: Yes. I think there are a couple of people who shouldn't be in that department.

T: Ok, well, do you think that those people are destructing your department?

D: Well, I told them things to do and they don't always do it.

T: And what happens when they don't do it?

D: Then I do it.

T: Oh, Ok. If that job is going to get done you need to sit down with that person, not yelling or screaming or anything, but sit down with them and teach them how to do it. In a patient manner explain it to them and tell them what needs to be done and sort of set a goal. Give them a task, set a goal and let them accomplish that.

D: Ok.

T: Does that sound reasonable?

D: Yes, I'll try to do that.

T: Ok. Now the other thing - Um, how are you handling scheduling at your department?

D: Pretty much the same way I did at my other store.

T: Ok, and how was that?

D: Well, I had a schedule set for my full-timers to work on weekends.

T: Ok. The way we try to do things around here is we try to rotate the weekend schedule, that way it gives everyone a chance to have the weekends off, as well as giving everyone a chance to work with everyone else on the weekends. That way everyone has a weekend off,

and that's good because everyone likes to have a weekend off, as I'm sure you do, to spend with their children.

D: Well, I wish people would tell me that. I mean no one has mentioned this to me at all. I feel like they're coming to you with all their problems, and I told them that if they had things they were concerned about they could come to me.

T: Ok. They should come to you. You are perfectly right. I am not saying that people come here, I just heard a few things and I just want to get things out into the open so we can talk about them. Um, maybe you need to have a meeting with your employees to bring some of these problems out in the open. Just have a meeting, maybe even away from the office so that they'll feel more comfortable speaking with you. Now, that way we can open the lines of communication. It's nothing personal. If they're not bringing the problem to you then you can't read their minds. I know that. We need to open up the communications, I think this is the most important thing we have to try and do. Ok, now, the job rating I'm going to give you for this first period here is just an average rating. Now, I know you are used to higher ratings, but I think that with coming to a new store, and the new employees and adjusting to the big city life, I think that's the major part of that. Um, I expect you to be receiving higher ratings in the future as you have in the past.

D: Well, I think I'll be alright, it's the people I have.

T: Well, the problem is though, that you're just one person and however many people there are in your department, 15 or 20, um, we can't just wipe out all of those people when we bring a new manager in. We have to work with what we have. The labor pool here is a little different than what you're used to back home, and a lot of the people you'll be working with won't be what you're used to. Sometimes you'll tell them to do something and they won't always do it. So what we have to do here is have a little more patience. I know it's tough, that's why I'm saying to you use the hands-off approach. Don't assume the responsibilities but delegate the responsibilities to your employees and be with them. Show them how to do it and be with them until they've done it a few times, until they feel very comfortable with it. Ok, now, it takes a lot of patience, I know it does because that is one of my major problems, so I can sympathize with you. Now, if you need any help or advice in the future don't hesitate to come to me because I know it is frustrating, and I can empathize with you because I've been through it all myself.

D: Ok.

T: Um, like I said, I don't see any problem with things improving. I think you have all the right qualifications. You have done a good job in the past, and I expect you'll do a good job in the future.

D: Ok.

T: Alright, well, thanks very much for dropping by, and in the next

six months I hope to give you a higher rating.

## Script for Training Simulation #2

K: How are things going?

D: Pretty well. I have been busy, but I think things are going OK.

K: How's the family?

D: Fine.

K: Kids doing alright?

D: Yeah, they're doing okay.

K: It's quite an adjustment moving from a smaller store to one quite as large as we are.

D: Yeah. It has been an adjustment. I mean there are a lot more customers to deal with but I think I have done pretty well. I mean I like the higher volume. I like keeping busy.

K: Good. Well, I know you are working really hard.

D: Yeah.

K: So, how are things going in your department?

D: About as well as could be expected, I guess.

K: Any problems?

D: There are just problems that you would normally expect, I guess.

K: Like what kind of problems?

D: Well, I don't think I'm always getting the support I am asking for.

K: Okay. So, you're having problems with the responsibility that you delegate. Do you think that your employees are not handling this responsibility?

D: That's it for the most part.

K: You feel that your employees are not handling this responsibility?

D: That's it.

K: Ok. There have been some problems in your department with things not getting done and hasty decisions being made. What can we do to help you with your scheduling and overcome some of these personnel

problems?

D: Well we can get rid of some of the people or give them more money.

K: You feel that giving them more money...

D: Well, I think that's to some extent part of it. They are not motivated to work if we are just paying them \$3.50 hour.

K: Well, some of your employees have complained that they are not given responsibility and they feel...

D: I have tried to give them responsibility.

K: Yes, okay.

D: I have tried to do that.

K: Maybe we can work together and set up some real goals and layout how we can delegate some of that responsibility and hold your employees more accountable.

D: That will be fine with me.

K: Ok. Some of your employees have also expressed that you sometimes show a lack of concern on occasion.

D: They said that to you?

K: Yes.

D: See when I came here I told these people that I had a open door policy. If they had problems or had things on their mind that they could come and see me. No one has approached me yet.

K: Ok. I think your employees are maybe feeling that they can't communicate with you, that you are not receptive to their problems.

D: They haven't given me a chance to be.

K: So you don't think your employees are giving you a chance? Do you think there is a personality conflict between you and your employees?

D: I don't think so. Not for the most part. I mean, there are a couple of people that I'll tell them to do something and they don't do it. But for the most part, no, I don't think there is any conflict at all. From my eyes there is not.

K: Okay. If you are responsible for the employees in your department then it is up to you to take action when the employees are not performing their duties. Are you dealing with on a regular basis and giving them feedback for their performance of the job?

D: Probably not every single time because I don't have time to

babysit these people. I mean, they have been here a lot longer than I have and they should know how to do the job. Now, do you agree with me or not?

K: Oh, sure.

D: Then in that case I am doing the best I can. I try to tell them what to do and there are so many other things I have to get done that I don't always have time to go back and follow up.

K: Ok. How can we relieve some of that work that you have daily that seems to get you so bogged down? Can we help you in any way?

D: You can get me some more help.

K: Get you some more help? And yet you have employees in your department that sometimes feel that they don't have things to do to keep them busy. How can we delegate some more work to them and keep them motivated and challenged in their job?

D: I thought I was.

K: Okay. Maybe these are some of the things that you can look at try to work on. Specifically, set up job descriptions for your employees or let them know what you expect of them and how it is going to be measured when the job gets done. Now I know that takes time in the beginning, but I think that you'll find that it will save you time in the long run, and will give you a chance to manage instead of doing the job yourself.

D: I can try, I guess.

K: Okay. Well let's see how things go in about a month, and let's get back together. What do you think about that?

D: That's fine with me.

K: Okay. I appreciate your attitude in trying to work with them.

## Script for Training Simulation #3

A: Pat, um, I see that you've been transferred from Kendall 66 to Kendall 15, and you have favorable recommendations, so it looks like your doing a pretty good job. There are a few small incidents that I've been informed about.

D: Problems?

A: Yes. One is that I've been told that you have poor decision-making judgments.

D: Who told you that?

A: Um, (pauses and looks up) I uh, I uh have been informed that, you know. Well, I have a specific incident here where you ordered picnic tables without checking last year's inventory records.

D: Well, that wasn't my fault. We had such a crowd rush that day there was no way we could have had enough picnic tables ordered.

A: And this resulted in underordering of merchandise that was needed.

D: Again, that was because of the crowd rush.

A: Crowd rush that we weren't expecting?

D: No, I mean there's no way we could have been able to tell that.

A: And you've been scheduling the same full time employees to work on weekend nights.

D: Right, I thought that's the way they wanted it.

A: Well maybe, you could uh, you know, move them around and have other employees working on weekend nights.

D: You see, I feel like they're telling you all their problems and not saying anything to me.

A: Well, I'll, uh, talk to them about that. Maybe they should be talking to you instead of to me but I am talking to you about this now. So uh, why don't you go back to your subordinates and talk to them about it. Maybe some people that have been working long week end hours would rather not work on weekends. And it says that you do alot of work that you could delegate to other people, that you do some jobs that a staffer could be doing.

D: Well, I'm ultimately responsible for how this department goes, so its, you know, I want to make sure things are done correctly. I feel like you're saying that I'm the problem in all of this and I don't



agree with that.

A: (pauses while looking down at paper) Well it is important to, um you know, rely on the help of others and not do all the work yourself.

D: Oh, I agree. I've told some of my people to do things.

A: You have been working 60 hours a week and (pause while looking down at paper) it says here that you yelled at a staffer ...

D: Well, I did because I'm sure there was reason to.

A: What, what exactly happened ?

D: I told John to set the display up front.

A: Uh huh.

D: And when I went up there he had it all screwed up. So I had to do it myself.

A: Um, well, I'll talk to John about that. (pause) Well, I'll talk to some of your people that you work with and uh, we'll see.

D: OK I appreciate that, I mean, am I going to get a bad review ?

A: Um, maybe, a mixed review.

D: See I think I'm taking the blame for alot of things that are my people's fault.

A: Uhm, I didn't consider that. I'll talk to uh, I'll talk to some people.

D: OK, I appreciate it.

A: OK

D: Is that all?

A: Yes

D: OK, thanks.

## Script for Experimental Simulation #1

O: How's it been going since you started here at Kendall 66?

D: It is going pretty good, a lot busier than the other store, but generally I like it.

O: Good, well as I indicated in your first week when you came in to start working here, that periodically what I like to do is sit down with new people and to talk about their performance, to talk about some of those things which you are doing well and areas that perhaps need a little improvement in them, and ways I can help you to work on those...

D: Ok.

O:...Set up a development plan, and then come back at a later date and see how we are doing. One of the things that I've certainly observed in your work since you've been here is the amount of enthusiasm and the amount of time you spend in working. You seem to put a lot of effort into your work.

D: Well I feel like its my department and I want to make sure that it runs well.

O: Do you tend to be satisfied with how your employees are doing?

D: They're ok.

O: What sort of employee relationships did you have in your previous job?

D: We were close. I mean all the people would, if they had problems, I felt like they could talk to me, and visa versa. If I told them something to do they would do it and those type of things. But I thought we were a real good group.

O: Good, Good. A couple of the areas that I've observed that I'm a little concerned with is perhaps in making some of your decisions. Sometimes I get the impression that you might be a little bit hasty and not thinking them through.

D: Why's that?

O: Well, Um, sometimes in scheduling some of your employees, in that you had some of them working on weekends, full time employees, and uh, that's not the best utilization of them. (Dolph interrupts while he continues to talk, "some of them have complained")

D: That's our busiest time. They've complained to you?

O: Well, I've heard complaints that have come from other people.

D: Well, see that's something I don't understand. I told these people when I came to work here that if they have problems they can come to me and and they're already not doing it.

O: Are you getting any feedback from them at all?

D: No, I mean that's the first I've heard about that situation.

O: Ok, it certainly is appropriate for them to do that. Another one of the concerns that I have is in the area of time management. I'm a little concerned you may burn yourself out in the number of hours that you're working (Dolph interrupts at number of hours "Well I'm working a lot of hours"). You seem to be working 60 hours in a week and all, you know in a short period of time probably, in special situations ...

D: (interrupts) But again I'm doing it because I feel like I've got to do it. I'm ultimately responsible for how well this department is run, you know, and I've got to be here.

O: Sure, well sometimes and it certainly is a difficult thing for people to learn how to do. I certainly had difficulty with it in my first management position, in learning how to let things go and delegate them.

D: Well, I've tried to do that.

O: That takes a long, long time to get comfortable with that and to expect other people to do it and feel comfortable with that. Um, the last area that concerns me a little bit is perhaps in being impatient with some of your employees and their doing things, perhaps maybe not being clear in your instructions to them of what you want them to do.

D: I've tried to tell them what they need to get done. I expect them to do it. They've been here a lot longer than I have.

O: Um hum, well sometimes it helps to define for people so that they will know what your expectations are rather than sort of just ,you know, demanding, sometimes it helps, it helps to clarify for them what your performance standards are. You know all managers operate a little bit differently. It will take them some adjustment period for them to get used to you.

D: I'll try to that. I've tried to do that a couple of other times, and it doesn't always seem to work.

O: Yes, well, I think that if you keep at it over a period of time as they adjust to you they'll get used to that and your expectations of them. Um, perhaps I should ask you if there are any particular areas that you would, that you feel you need help on, that you would like, you know, to put into the development plan that we are going to put together.

D: Well, just that I've, you know, I've tried to tell some things to

some people and its not always done very well. That's...I never had that at my other store.

O: Do you, uh, can you identify any of the reasons for that difference?

D: No, I don't know what the reason is. I mean, the people here just don't seem to be motivated to do the job. I mean I've told them things to do, and I've gone to check behind them and its either not done very well or not all. And I have to do it myself.

O: Um hum.

D: I think one thing is the money. I think we're not paying these people enough. I would think that for the type of work that they are doing we could pay them more. I would like to give all these people a raise or just get them out and get some people in that want to work.

O: Well, that's certainly something that we can look at, and talk with our personnel people to look at our salary scales and see what we can do about that.

D: I think that's something we need to do.

O: Ok, that's a good suggestion. We'll certainly look into that. What I'd like to do is to meet again with you in another month and to sit down and talk to see how you are doing. You know, and talk again about what areas are working well for you, and what areas still may need a little more work on. It certainly takes, takes time to get up to speed in working in a different place.

D: Yes, it takes time to adjust.

## Script for Experimental Simulation #2

J: I'm Chris Palmer, I don't know if we've met previous to this or not. How do you like it here, working here, compared to the other store?

D: I like it pretty well. It's alot busier. There's more volume so there's alot more customers and alot more staff. But I like it pretty well. I mean, its a nice store, I like keeping busy.

J: Yes. I can tell. You've been putting alot of hours in so...uh... is it uh...if its busier and you're staying busier, I mean, how's it, and the volume's more...

D: Yes, we just have more customer traffic so I'm here alot more.

J: We want you to work out well here at the store, we've done an evaluation. We do evaluations twice a year on people. I don't know how the other stores have been doing them. We do them twice a year. We want to make sure everyone understands what their responsibilities are and they're doing all right. I was worried about...the only...I see some good things. Sixty hours, that's alot of hours you can put into a week. I know you've just been bushed and all that. I wanted to encourage you to...um...put your people to work as much as you can.

D: (interrupts) Well, I mean I try to do that.

J: (continues) So we might take some of this load off you, rather than overworking you. You're not going to do us any good when you're worn out.

D: Well I've tried. I've tried to give my people more work.

J: How many people you got working under you right now?

D: I have about 16.

J: 16, OK. Are you keeping them plenty busy so you can take...

D: Well, I mean I've tried to give them work to do.

J: Are you work...What I...I think where I'm mostly concerned is I see how many hours you are putting in and I want you to be able to figure out a way so that you can cut down your hours and put your people to work as much as possible.

D: I, well, I feel like I am ultimately responsible for the success of the department and if things need to be done I need to make sure it is done and that's why I'm working so much.

J: Have you got particular work categories for leaving people so that

work...so that they know automatically...so they know where their assignments are, where their responsibilities are. Do you have a clear cut...so that they know and you know where...for each situation rather than having them come in and watching all the time they pretty much know where your categories are.

D: I thought they did. I mean they've been here longer than I have and I just assumed they knew what their jobs were.

J: But you're not too sure?

D: Apparently not. I mean I...

J: (interrupts) Well I think, I think it would be beneficial for you, again 60...I think you are working as much as 60 hours a week and um...I know it is a big jump from the store you were at to this store so um...I'm wondering if you might want to get together with your people to work underneath you to have maybe a meeting to define some of the responsibilities that have been going...Because before you got here the man that you took...whosever place you took probably had policies established and I think we need to reinforce how you want to have your people function what...what capacity you want them to function in and that will take some of the load off of you so you don't have to put in those long hours like you've been doing. And it does the company no good to have you worn out all the time and having to extend yourself so far. Um, I suggest, and I trust your judgment on this, and I suggest, that it would be good to get together with the people that work for you and just clarify for your own sake, and for their sake how your responsibilities are going to flow. And, uh...you've got some good people working under you and I'm pretty sure that from what all I've heard are pretty responsible, and they probably want that responsibility assigned to them if you can get comfortable with that. It's hard sometimes to turn over...uh...turn over responsibility because it's hard to...because sometimes it feels like you are losing some control.

D: I've tried to give them some responsibility and they haven't really, haven't always taken it.

J: Can you give me an example?

D: Well I told John the other day to fix the display in front and it really wasn't done very well or done, you know, a halfway job...

J: (interrupts) To your expectations?

D: So I had to do it myself.

J: Could you have had John redo it? Would he have made improvements...next time you're going to have that same go round with him next time you ask him to do a display. Either you're going to have to do it yourself or you're going to have to get John...or you're going to redo what John did. And that's...in essence that's going to make it harder on you, number one because you're going to have to go

behind him all the time, and number two, I think it's going to make him feel bad about himself because he can see what you're doing, that you're following behind him and doing that so it's going to demoralize him and it's going to wear you out. Um...it's a hard...I think it's hard telling...keeping other people in line is a hard job but from your own work load you don't have the time to be redoing any work for him.

D: No.

J: You can work with them a little bit...uh...and tell him in other words in that case tell him what he did wrong because you've got that expertise. He's calling on what information and knowledge he's got on his...in his background and you see a bigger overall picture, probably see more of the business all the way around because you're the manager so go ahead and call on your expertise and tell him what he did wrong that time. Uh...and he'll probably be a little miffed at first, but he will probably think about it and see that what you're saying is right...better that than do it over and him seeing that his work is changed and that we didn't give him the benefit of learning from it. Go ahead and see if we can't...you know go ahead and delegate that job to him to do again. Show him what points you want improved and the next time he will be a better man and you won't be so worn out. For 60 hours you can't keep this up.

D: Yeah, I'll try.

J: OK. I appreciate the hours that...that's a lot of hard work and that means you have a lot of loyalty there. I don't think anyone is going to put in 60 hours and not have loyalty to the company. I appreciate that. We just don't want to wear our workhorses out, because we want you to be around for awhile in the company. We want you to practice getting those people underneath you to do what they've been paid to do and we want you to show them how to do it, not be spending your hours doing it for them. Other than that everything looks good. I think you can slow down on the hours and increase the number of things that you can...put these other people to good work. It will keep them out of trouble that way, and I think it will work out alright. Is there any question that you have of what responsibilities or obligations or um...that you are having that we can work on now, and find some objectives to reach before we have our next performance evaluation?

D: No, not really.

J: Not really...because I'm sure it's a two-way street. Sometimes these situations get kind of locked into the manager. Upper management sort of cracks a whip and the other person doesn't have much input. I hope that we can get you off this 60-hour a week routine as much as possible because, like I said, it doesn't do us any good if you're so worn out that you can't do what you need to do. Well, I've sure enjoyed seeing you again.

D: (nods his head)

J: And we will meet here after 6 months and go over again to see how well you can get these other people underneath you to work, and you come back with me with what you think can be done.



## Script for Experimental Simulation #3

R: Okay, it is performance evaluation time.

D: Yes.

R: Pat, it seems you came from store #15 with a favorable recommendation. Your performance evaluations in the past have been good. I don't understand what has been happening since you came to our store.

D: Why do you say that?

R: Well, it seems you have been making some very hasty decisions. This is just one example, but there was a time you ordered picnic tables without even checking last year's inventory.

D: Well, I don't really consider that my fault.

R: Whose fault would it be?

D: Well, you saw how busy we were that day. Nobody could have guessed right on the picnic tables.

R: Why didn't you just look in last year's records? You never, ever, you don't have to have a degree to know that you never order anything without checking the stock to see what we already have.

D: I thought I ordered the right amount.

R: Well that cost us because there were alot of things we couldn't order because we ordered those. It also seems you have been repeatedly scheduling the same employees for week-end nights. Do you have a grudge against these people or what?

D: No, I just thought they wanted to do that. At my other store the full-timers loved the weekends because they could make their most money.

R: Well, have you talked to your employees here? I mean not everybody is alike. Maybe the employees at your other store needed the money, but with the system here we are getting complaints about it. Your employees obviously don't want it that way. I think maybe you need to talk to them.

D: Well, I told these people when I came here that I had an open door policy. If they had problems they could talk to me about them, no one

has approached me about anything.

R: I think they might be scared of you Pat.

D: Scared of me?

R: Yeah. It seems you are very demanding. You yell at them.

D: Well, I have high standards.

R: I'm sure, I'm sure, but the way it gets through to me, it just seems like you are not patient with them. You need to sit down and listen to what they have to say. You can't sit there and yell at them for not knowing something, yell at them for not remembering something. I have an example here. Someone actually heard you say...two of your staffers had asked you to explain how the inventory systems worked.

D: I remember that. Those two guys had been slacking off all day long. They had not done anything.

R: Maybe they did not know how. Did you think of that?

D: Well, they have been here a lot longer than I have. They should know how.

R: That's what it says you said. How are they going to know if there is no one to turn to to tell them?

D: Well, they wanted me to go back and do their stock inventory for them and I wasn't going to do that.

R: Well, I think you just need to sit down with your employees and find out exactly where the problem is lying.

D: I can tell you that.

R: So, you obviously think it is in your employees.

D: Well, if I tell someone to do something and they don't do it or they don't do it very well then I have to go back behind them to do it.

R: That's not the way it should be.

D: I know it's not. It was never like that in my other store.

R: Something's just not right here. Somewhere down the line you are not clicking with your employees. I've got six months before another performance evaluation and I want to see something done. I don't care how you do it, but somehow you've got to start communicating with your employees.

D: I will try.

R: If it takes discipline, if it takes a reward...

D: Well, see I agree with both of those. There are people here that don't want to work. I think we should either get rid of those people or get some people in that want to work, or give them more money. We pay these people minimum wage and that's the type of help we get.

R: Yes. I agree. But you know, try the system it takes to get respect. I want your employees to be able to respect you, but I also don't want them to be scared of you. I want them to be able to come to you with a problem, and I want you to solve it with no conflict.

D: I'll try.

R: Okay.

## Script for Experimental Simulation #4

R: How are you doing today?

D: Pretty good.

R: Okay. I, uh, just wanted to talk to you about your performance evaluation.

D: OK.

R: Ok, I noticed you're a real hard worker.

D: Yeah, I like to make sure things are done right.

R: Yes, I noticed from, ah, the other store, that looking at your recommendation, they said you work really hard.

D: Yeah, I try to. I had a good store over there.

R: Ok, then compared to the other store, we're a bigger store over here, so that we have more staff. And it seems that you need to delegate more responsibility. You understand what I'm talking about when I say that?

D: Well, I'm working a lot of hours and I've tried to delegate.

R: Yeah, I see that it says you're working up to 60 hours a week.

D: Yeah.

R: Now, I think what you need to do is that we have a pretty good staff. You need to let them help you out more, do things in the stock room and stuff.

D: Yeah, I tried all that.

R: It seems that you need to take more time to schedule their hours better and work on the inventory and stuff like that. Let them do more of work, so you can manage them.

D: Yeah, I tried to do that.

R: Is there a problem with your workers?

D: Yeah, well I told a couple of people to do something and they didn't do it quite right. It wasn't up to my standards so I had to do it.

R: Can you give me an example?

D: Well, I told John to move some furniture for a display and it wasn't done very well. He moved a couple pieces here and a couple pieces there. So, I ended up doing it.

R: Well, did you tell him exactly what you wanted done?

D: John's been here longer than I have. He should know how to do that.

R: (pause) Well, have you had any other problems with him?

D: Not him specifically, no.

R: Well, what would you say in general with the staff you have here? Are you pretty satisfied with them?

D: Ah, they're average.

R: Average. You think you give them enough responsibility?

D: Ah, like I said, I tried on more than one occasion. But you know paying people \$3.45 an hour, it's the type of help we get.

R: How do you think we can get them to show more responsibility?

D: I don't know. I mean, I would think that if you give them anything, let them do it. But, apparently that's not the case.

R: (pause) Well, maybe, maybe you should give them more responsibility? See how that works for awhile.

D: I can try.

R: I think that would be a real good idea.

D: OK.

R: I think, um, if you try that, you, ah, you should have more time to attend to your other responsibilities, see that things get done.

D: Yeah.

R: So, you know, so you can have more time, to, ah, schedule, to do the inventory, stuff like that...ah, take care of the problems.

D: What problems?

R: Well, like the ordering.

D: Oh, that wasn't my fault. I mean that we just had a big rush that day. I thought I ordered the right amount.

R: Okay, well I think to avoid, you say you had a rush that day? How

come?

D: That's a good question. I think because we had such a good sale.

R: Well, I think in that kind of environment, I think what you need to do is if you let your, ah, staff do more for you...let them do the stuff more.

D: OK, that's fine with me. I'll be glad to do that.

R: Good, I think you give them more responsibility, you can spend more time, ah, doing the inventory...ah, checking the stock, seeing what we need to order.

D: OK.

R: Well, let me see what else I have.

D: Something else? I feel like you're saying I'm doing a bad job, and I think things are going pretty well.

R: Well, I don't know. I see how much you work and I know you're working very hard, but I think we can utilize you better, ah, if your not doing so much of the routine stuff, the day to day stuff, if your doing more managing.

D: OK.

R: Yeah, I mean I, I see you out there 7 days a week, 10 hours a day, ah, moving furniture around, and that's not what we paid you for. We, ah, we don't really need that. We need a good manager.

D: Well, like I said, I tried that. I mean I tried to give them more responsibility.

R: Okay. Yeah, because we, ah, we didn't hire you, ah, because we're paying you more than them. We're paying you more than we pay them. We expect you to help out more in the running of the operation, okay?

D: OK, I can try.

R: Okay, well I think that should, should do it. Ah, if you have any more problems, feel free to come back and see me.

D: OK, fine.

R: OK. Well, thank you.

## Script for Experimental Simulation #5

B: How are you doing Pat?

D: Pretty good.

B: Good, glad to see you. Glad you could come in. So how is the new job going?

D: Good.

B: That's good.

D: It's a lot busier than I'm used to. But I think things are going pretty well.

B: Good, alright, looks good. So, OK, as you know, this is your performance evaluation, the first one that you're going to get from me. And I'd like a little feedback before I start. How are you rating your performance right now on the job that you're doing?

D: Well, I think I'm doing a pretty good job. I'm putting in a lot of hours making sure that the department runs well, and I think things are going pretty smoothly.

B: Any particular troubles you've had so far?

D: A couple of areas, but....

B: OK, alright, well I've got a few problem areas that have come to my light. And I'd like to discuss them with you as well. OK, as you know, you came with very high recommendations from number 15, Kendall 15. And I was curious on a few areas. It's come to my attention that you have, you know, have on occasion made some hasty decisions without, without checking your records, things like that, and things along those lines. You know, making decisions before you've really thought them out.

D: I'm not sure I know what you mean?

B: OK, well, for example, I had ahh...ahh, oh, remember those picnic tables came in and you ordered...

D: Well, I don't really consider the picnic tables my fault. I mean we had such a crowd rush that day that I don't think anyone could have guessed the right amount.

B: Well, ok. But if you had checked the schedule, then, if you had checked the past orders, you'd notice that we always stack heavy for that season because there's a big order for it. And it's going to be getting used to the job, I'm sure, but you have to, you know, you have to think in those terms. And look at our old records, our past

performances, and because that'll tell you a lot of insights to things like that. Because ahh, we've got to be prepared, and we did lose quite a few customers on on that thing. OK?

D: Ok.

B: Alright, another area, is the full-time workers. You know. I've had some complaints.

D: My workers?

B: Yes.

D: They've complained to you?

B: Well, no, no I hear it from other sources. No, they didn't complain to me.

D: Well, see, I told these people when I came here that I had an open door policy.

B: OK.

D: If they had problems, they could approach me. And no one has said anything to me about those problems.

B: No one has approached you about it?

D: No!

B: OK, alright, well that's good. If you told them that then, they didn't tell you...Well let me make a note of that. Because a lot of the full-time workers here at the store, they tend to think, well, that their weekends are, well that they've earned the right not to work on weekends, stuff like that. And you see, you know, it's things like that.

D: Well at my other store, full-timers loved the weekends because that's when they could make their most money.

B: Really? Ok, alright. Then, it's the nights, especially, that they, really don't like. They want that time to do other things. I'm saying well it's probably that, it's from you changing from a bigger format here that it would it would make it more difficult. It's larger, and some of the things will be different and you just have to go along. All I'm doing here...is, and don't take it negatively, ok? I'm just pointing out to you areas that I have seen as weaknesses in the changing, ok, and a lot of it is probably coming from from a smaller to a larger format. You know and, and things like that. And alot of the policies are going to be a little different, ok?, you know, not much. We try to fill you in as much as possible when you took over the job. But this is just my way of pointing out areas that I see. And you know, I don't want you just to sit here and think I'm cutting down everything that you're doing. These are just weaknesses



that have shown up. Let's see....another thing, well, the people around here, they like to have, they like to have the trust of their, you know, their superiors.

D: Right.

B: And, they like, prefer to, for their superiors to tell them what to do and what...and then to have it done. I've noticed that you're a real hands-on type of manager, and really getting in there. And I've seen you occasionally doing things that you can assign someone else to do.

D: Well, I've tried to delegate.

B: You have tried?

D: I've tried.

B: And they're giving you a hard time?

D: Well, there are some times when I tell people to do some things and they're not done very well, or not done at all.

B: Uh, Ummm.

D: So, I end up having to do it.

B: Well, OK. Well, those situations, then that's good. A manager has to do that. That's true.

D: Well see, I feel like I'm ultimately responsible for the success of this department.

B: That's good. Well, that's true. You know as the manager that is very true. However, you gotta, on the same token, I mean, you are responsible. So, I won't interfere with that at all. All you have, you just have to allow the employees the chance. Like if they get a bit, a little behind, instead of you going and doing it, how about talking to them about it? Say, "you know, look, you know, you don't have to leave it behind and, and I've given you this responsibility, so, so, why don't you stick with it? You know you don't, don't leave it behind so that I have to come back and get it." And you know, they'll probably appreciate that more than, than having you go, you know going and doing it. Because, then, they feel like they're not being trusted to do the work. And, and if they don't deserve that trust you know, then we have to do something.

D: Well, see, that's what I was going to suggest.

B: You feel like they don't deserve the trust?

D: Well, I think there are some people we should get rid of, or give them some more money.

B: Ok, well, have you sat down and talked with these people?

D: No, not really. I mean, these people have been here a lot longer than I have. They should know what the situation is.

B: Right, ok, well, that's true. But on the same token, well they are your subordinates. So, maybe, well firing is an extreme. You know we don't like to do that. You know? It's just not good policy. It really isn't. If you can work with them, you know, these people have shown us in the past, you know, they've shown us that they've all been average or above average in their performance. Ok? And you...have you had any personality conflicts?

D: No, those things are going to pop up though if they're not doing the work I'm delegating to them.

B: Uhm Uhmm...

D: But, I don't think anything major.

B: OK, alright, that's good, that's good. I'd recommend that you ask them, you know that you sit down and talk, especially if you've got one or two that seem to be slacking off, and slacking off the responsibilities that you're giving them. Sit down with them and discuss it. You'll find that most of them, you know, that they're pretty open-minded. You know, they'll listen to you. And, and if that doesn't work, then come to me and we'll sit down again, and we'll work it out. OK? I mean if we got to fire them, then that's the way it is. I mean that's something that, that's ultimately your decision as, you know it's your decision as well. I mean it's your department. Let's see, the only other thing I can really think of, that I've had problems with, or that other people you know have been working with you, is that you gotta watch all of these people as far as demanding too much. OK, now keep them working. You know, you gotta be patient with them though. Because if you're not patient with them, they take an attitude.

D: Well, I have high standards.

B: Well, OK. That's fine. There's nothing wrong with high standards you know. But, you gotta realize that they might not have standards as high as yours. OK? And if you want them to reach these then you have to express it and don't, say...if you berate them about it, they're not going to do it. Alright? And if you go in there and and, you know jump on them, it's kind of like, like you know, the cart pulling the mule. They're not going to be able to work for you, by jumping on them. It's better to try and sit down and say, "you know look, you know, this is what I expect. This is what I want you to do." And don't go in there and say you know, you didn't do this, and tell them they're doing a bad job. Say, "this is what I expected of you and, and you're not, I'm not getting quite as much out of you as I expect from you. I'm not getting what I want out of you, and you're going to have to put, perform a little more, and put a little more effort into it. Try to, you know try to talk to them a little more on

their level than as a subordinate-superior. You know try to say I understand the job. Show them that you understand the job and and not demand that they do it. You follow that all?

D: Yes, I'll try to do that.

B: Ok. Well, if it works out, you know alright. You're filling in someone else's shoes. Ok, you know that's always going to be a problem. Alright, whenever there's a management change, and people are more accustomed to his way than yours, and you've probably...back...I'm sure that the guy who came into your place is having the same amount of trouble. I'm just saying that you have to work with them, because, because people have shown in the past that they're dependable. And they are good workers. And I don't want any animosity between you and your workers, because of of your differing management styles. And these are the areas that I feel you have to work on. All I'm trying to do is resolve these problems. OK? That's what I want to do. These are the problems that have been brought to my attention. And I just want to resolve them. Now is there any other ways I can help you resolve them?

D: Well, I just hope you don't think that I'm the problem. I'm not trying to be a problem here.

B: Yeah, right, right, I know. I realize that. I'm not blaming you, per se. OK? You know. I'm not saying it's your fault. I'm not. I'm just saying that these are areas where something is going to have to be worked on, or they will become problems. You've only been here four months, and you're still getting your feet wet with everyone involved. I'm not trying to tell you that you are the problem. But I'm saying, you know that these are problem areas, and if we don't do something about it, there is going to be a definite problem. And it's going to be either, you know, we're going to fire all of them, or move them around, or move you around. And I'd rather have it work out with you, with you to work out with them so that everybody can stay here. And, and you all work together as a team. That's, what I think is the best way to get things accomplished. That's my personal philosophy. And as your superior, you know I prefer people to work it out, you know, than to have to move people around. You know, uhm, you know and fire them, you know. And these are the areas that I feel need some work. You know, just take your time. And we don't, well unless it's something right there on the floor that needs your immediate attention, take a little extra time, and think about your decisions before you make them, you know? Do a little research if you have the chance. You know, watch out for the full-time employees, and give them a break, because they've been here for a while. Give them a weekend off every now and then, or like a weekend night... something like that. Give them a day on the weekend. That's the way we've always done it here, and they've come to expect that. Ok?

D: Ok.

B: And, you know there's no sense in changing that. You should be able to change instead of them, because we've pushed for that over a

period of time. Oh, and they've come to expect that. Ok, well that's something to watch for, you know. Give them a little time. And, and give them a little rein. If they're not doing their job, then you got to sit down and talk to them, you know. Like, you don't, should, be down there. You know, you shouldn't be down doing the staffer's job. You've put in a lot of hours, and it's just that, you know, you're following up behind your employees, and you're really not accomplishing anything. Because they should you know, because they should be doing this. They should be getting the work done. And it shouldn't be in your lap. And, if they're not getting the work done then you gotta talk to them. And if that still didn't, doesn't work, then the three of us will come up here and we'll sit down together, you, me and the employee. And if that still doesn't work out, then you know, there's something wrong and he's going to have to go, or she's going to have to go. But give them a chance, because they really, well, in the past, they've been able to do the work. OK? And like I said, it's a rough period and right now, so give them the chance and have a little patience with them.

D: Ok, I'll try to do that.

B: Ok. That that's about all I can say. Those are the only... overall your performance hasn't been bad. OK, so I guess there's a few problem areas. I know you're changing to a new situation, which always causes problems for you and your subordinates. But you know, you gotta work them out. OK?

D: Ok.

B: Ok. Good you could come in. Glad we could get things straightened out.

## Script for Experimental Simulation #6

S: Hello Pat, and how are you doing?

D: Fine, thanks.

S: Good. Well, as you know, this is the semi-annual evaluation. First of all I'd like to ask how you feel the job's been going for you since you moved over?

D: It is going pretty well. It is a lot different here. It's a larger volume store, more customers, larger staff. But I think I've adapted pretty well.

S: Do you feel you keep pretty busy, busier than you were before?

D: Yeah. I've been working a lot of hours.

S: Yes, you have been putting in quite a few. Well, have you run into any specific problems that you need to ask us about or anything I can help you with? From your experience any major problems?

D: No, not really.

S: OK. Well, a few complaints have been made to me and I think we need to discuss them. One comes from several employees and it involves scheduling difficulties. Scheduling the same full-time employees to work week-end nights. They have expressed some complaints about this to us.

D: They complained to you?

S: Yes. I am not sure exactly how it got to me but it got to me. I don't know who complained first.

D: See, I thought that's the way they wanted it. At my other store the full timers loved the weekends because that's when they could make their most money.

S: Really? So they haven't told you about this?

D: No.

S: Okay. Do you have regular meetings with them or is there any way they can get in touch with you to find out how serious a matter it is? It may have been a casual remark, but when it got to my level it was a complaint.

D: Well, I haven't talked to them about it, not as of yet. I told these people when I came here that I have an open door policy, and that if they have problems they could come to see me about them, and

haven't approached me about anything.

S: Maybe it would be a good idea to find out how these people feel. Do you have regular staff meetings?

D: No.

S: Maybe you might want to have one. We'll wait and see on that. Something else here, several staff members have expressed dissatisfaction about having very little responsibility. Can you shed some light on that ?

D: I have tried to give them some responsibility, they just don't accept it.

S: OK. Do you have a hard time getting the performance that you want from them?

D: Sometimes I do.

S: Is that frustrating?

D: Sure it is.

S: Well, I guess at this point I'd like to talk to you a little bit about delegation versus supervising. I'm not sure how much, you worked with a smaller staff before and maybe that's part of the difficulty of adjusting to a larger staff, but with delegation, how much training have you had in delegation?

D: I know what delegation is and I've tried to give my people things to do, but again, it's just not getting done.

S: OK. When I think of delegation I think of assigning a certain task to someone who is responsible and someone who has the capability to handle that task. You get them started on whatever it is and you make yourself open. You said you have an open door policy. Are you showing them this open door policy when they approach you?

D: I think so. I think that's part of the problem. People just haven't approached me about anything.

S: Maybe there is a problem because they are complaining, but they're not complaining to you. I know that you put in more hours and you're more frustrated. Maybe you are losing a bit of your sensitivity. It happens when you put in more hours.

D: Well, I put in more hours because I want to make sure things get done. If I tell someone to do something and it's not done very well, or not at all, then I have to do it. I'm ultimately responsible for the success of this department.

S: True.

D: So I've got to make sure things get done and that is why I work so many hours.

S: So you don't want to do it yourself if it's not done?

D: Right.

S: OK, Well I would suggest to you, instead of taking that course, that if you have delegated it to someone and given them proper assistance, and told them that you have an open-door policy, and you've told them to come to you if they have any problems and they still have not done it; maybe instead of going and doing it yourself you should pull that person back and say, "what's the problem here?" Follow-up on them because they're not always going to come to you, and say, "Look, I've got a problem here," especially you're new, 3 or 4 months. And they might feel intimidated or whatever, and if it's wrong say, "what can I do to help you with this?" Because you have certain responsibilities that are in your job description that aren't being done, and you can't afford the time to do all of their things and if they're saying they don't feel that they have enough responsibility something has got to be worked out so that you are not doing it.

D: See, these people have been here a lot longer than I have.

S: True. Do you think they need more training after observing them for a certain length of time?

D: I think there are a couple of people who need something. Whether we should get rid of them or give them some more money.

S: So, you don't think training would help?

D: I don't know. These people have been here a long time. If they haven't learned their jobs by now I'm not sure exactly what training would do for them.

S: Maybe we need to do some evaluations with them like what I'm doing with you right now. Maybe we need to evaluate them to see if they're trainable, and if they're not trainable then fine, we can get rid of them. But maybe they need another chance. This is another point. Several employees have informed me that sometimes they feel you're too demanding with them, that you don't show enough concern with them. From what I have seen, talking to you now, you do seem frustrated. When you are working so many hours and you have this extra responsibility I can see how this can happen.

D: Well I have high standards. I thought that is how you got ahead in this company, and I expect my people to have high standards as well. This problem never occurred at my other store.

S: You can set high standards, but you've got to help these people live up to those high standards. You can't say, "I'm not going to give you any more training. I don't care what kind of person you are,

you don't meet up to my standards, you're fired." That's not good personal relationships. I'm sure you understand the value of good personal relationships because it directly affects the performance of a person. Maybe we need to talk about communication. How well do you communicate with these people? Do you have any specific personal problems with any of the employees?

D: Well, I think there are some people that don't give me the respect that I deserve. But, see, I feel like you're saying that I am the problem, when I think its because the people are not doing what I tell them to do.

S: You're saying that is the root of the problem? Well if the root of the problem is that they are not doing what you are telling them to do, maybe we should look at how you're telling them to do it. If you are demanding, then you are not going to get their respect. If I commanded you to do this, this is my standard and you don't meet up to it, then I am firing you. That doesn't come across well and you are not going to respect me. So maybe we should try seeing it from their point of view. How do you come across?

D: Well I think I come across ok.

S: Do you see what I am saying?

D: I see what your point is.

S: I know you put in a lot of hours, and there are some things in your job description that are not being done, and I realize again that you're new and it's hard adjusting to new things sometimes when you're new, but I think that you should try to work on your relations and how you communicate with them. You've got to play a game with them, you've got to get them to want to do a good job, and if you can communicate that to them, and have them start doing their own responsibilities and you don't always have to go in and always clean up behind them, then pretty soon, maybe you can turn your attention to your own job description. Do you understand? These things do tend to slack off if you're always going in and doing other people's work and you can't do your own. Do you understand?

D: Yeah, I understand.

S: Do you know how you can do anything about...what are you going to do?

D: Well, if it were up to me again, I would either give these people more money or get rid of some of these people.

S: Do you think that they deserve more money?

D: Well, I think that if you pay people \$3.50 an hour, that's the type of help you get.

S: Maybe we should try a little more personal relations. Work on



communication and we can meet in a few weeks and see how it's going. But, I don't think that's an option for us right now to just fire them, that would be too expensive to fire them. To just totally fire them and hire new people would cost us in other ways I think. So what are you going to do?

D: I will talk to them now, I guess.

S: How? Talk to them as in this is not what you're doing or....

D: Well I feel like if those people need some help I'll be there to help them.

S: Ok. The problem, as I understood it, they don't feel that you're patient enough with them or that you're not concerned with their needs. Do you understand what their needs are? Do you have a perspective of what their needs are, like what equipment they need, or how much time they have to get something done?

D: I think so.

S: Maybe you should ask them. Instead of always saying I'm here if you need me, say you need me because you're not getting this work done and what can I do to help you. I think that would be a better approach.

D: Okay.

S: Okay. We'll get back together in a couple of weeks and see how things are going.

D: Okay, fine.

S: Well, in the meantime, as that area starts to improve you will find that you have less work to do and there will be less frustration, and I think you will be able to get your work done as well.

D: I hope so.

S: I hope so too. How are there any questions or anything that I can help you with?

D: No.

S: Ok. Well, thanks Pat for coming in and talking with me today.

D: OK. Thank you.

## Script for Experimental Simulation #7

C: Hi Pat. How do you like your job here so far?

D: Fine.

C: Good. Good. Glad to hear that. Mr. Randolph from store 15 spoke highly about you. Are you familiar with the performance evaluation meeting? Have you ever been to one?

D: Yes, I have had a couple of them.

C: What we want to do today is talk about your performance here and do what we can to work out any problems that you might have or find out any points I might have noticed. We want to do what we can to work better together. What comes to my attention first off: a couple of complaints that employees have come to me with...

D: My employees?

C: Your employees have come to me with...have you had any trouble with your employees that you feel they might direct at you, complaints?

D: I'm not sure what they are complaining about. If anybody should be complaining it's me.

C: What are some of your complaints with them right now?

D: Well, I've had better workers before.

C: So you feel like they are not as dedicated as they should be.

D: No, not as much as my other store.

C: If you could improve anything with your employees what would it be?

D: I think we should get rid of some of them and get some people in here that want to work, either that or give them more money. I think we pay these people minimum wage and that's the type of help we get.

C: OK. You do have a point there. Let me, I don't want to be too abrupt with you but let me talk to you about a couple of critical incidents that people have spoken to me about and things that I have noticed. I don't want to put you on the defensive side but I...we need to work at this and figure how...if any, what the problem is that exists. I sometimes wonder if you are paying attention to detail. You are definitely dedicated. As far as I come in I see you are working 60-hour work weeks. You are also willing to come in on off hours, and I appreciate that. I am wondering maybe if your time could just be better spent if you would manage it better and possibly delegate some responsibilities.

D: Well, I try to delegate.

C: Did...what type ...did you work out any specific system of delegating responsibilities?

D: No. If something needed to be done I would just tell someone do it.

C: And as related to that I had 2 staffers...I overheard 2 staffers ask you how the inventory system worked. You told them that you hope they found out soon. I was not sure of what your meaning was behind that.

D: Well those 2 people had been slacking off all day long. I had been doing their work most of the day and this was just another chance for them to get me to do the inventory for them.

C: And then yelling at a staffer...I guess that was for the same reason.

D: Same reason. Same reason.

C: They were aggravating you...also I noticed that the weekly inventory has not been being taken so we'll know what to order and we'll have some accuracy in the department. I know you work long hours. Some things I consider critical just aren't being done- the inventory and cleaning behind the back ledge which always is getting dirty from all the plants we have back there. I'm just thinking that possibly you could assign some specific tasks, or maybe give them notecards with their specific responsibilities on it. Uh...sit down and talk to the employees you know, if necessary decide they are going to be your friends even if...they are not going to have an agreeable basis between you, even if they seem to be contrary. You know, just do the best you can. And get them to agree to the task that you want to assign them and maybe work out a few little things. Maybe if they feel that they would be better at one little thing or another...

D: Do you think that'll work?

C: I was thinking, sit down and go over the task with them, and then write them a notecard for what they are supposed to do.

D: These people have been here alot longer than I have. They should know what their jobs are.

C: OK. Well it seems that they could be in need of some direction and they could want some more direction I feel. I think it is good that you are willing to do the work, but alot of mundane tasks that I've seen you doing and I just really feel like one of them could be doing and you could be using your time more wisely doing the things that requires your experience. Mr. Randolph from your other store has told me about your expertise, things that require skills that these workers don't have. But to go back to assigning the tasks, you can pitch in

occasionally and let them know that you are not afraid to work, that you have delegated these tasks to them. "I can do anything that I assign you to do." But I don't want you doing other people's jobs. I don't care how bad they moan and groan...

D: Well, I am ultimately responsible for how this department is run.

C: OK. Well let's try this...no matter how bad they do moan and groan just tell them that it's their job and in a nice way follow up and find out if they are having problems rather than just leaving them and telling them they are going to have to do it. Try to listen to what they have to say and get some feedback and if it absolutely doesn't work out, then come to see me and then we'll get rid of them. It seems to me that you are possibly doing alot of things that you could be delegating to some other people. And if they knew what their direction is, possibly they would be more inclined to do the work if they knew what it is that they have to do. Unfortunately, minimum wage, which is all we could pay, attracts only a certain mentality you might say, or a certain type of person...uh...and alot of times that type of person responds well to just lists, and mundane orders...just...they want to know what they have to do. They are people with low initiative alot of times. What is your overall opinion on this?

D: Well, I guess my opinion is if we want go-getters why do we pay these people minimum wage?

C: You can be the go-getter that runs the department and delegates these tasks-mundane, the mindless things, cleaning the back ledge, counting items for the inventory. You can use the inventory and order things so that we don't have the trouble like we had with the picnic tables.

D: Well, I don't really consider the picnic tables my fault.

C: OK, what was the problem?

D: We had such a crowd that day, there was no way I could have ordered the right amount.

C: So maybe we ought to even start planning ahead on that too. Maybe you can give the deadlines of the tasks, on the tasks that aren't daily tasks like, you know, "as you get time, I need the front windows cleaned by Tuesday- "I'll need a count on aisles 2 and 3 of all the merchandise on aisles 2 and 3 on Wednesday afternoon so I can get it in here by Friday." Try to realize that they have been here for a long time and I've gotten to know some of them-not on a personal basis-but I've seen them. I feel like their intentions are good and they feel sort of misdirected and without direction sometimes possibly. How do you feel? Do you think this will work for you?

D: I'll try. That's all I can say.

C: Did you have any type of system like that at the other place?

D: No. Those people there wanted to work.

C: Just everybody pitched in and you never had to tell anybody to do their particular task?

D: Well, sometimes I did, but generally people knew what they had to do.

C: How about if we say...I think you need to spend more time at home...at least away from work. I know you would like to work and you are a good worker and I appreciate it, and I'd like to see you strive for a 40-hour work week delegating as much responsibility as you can. Some nights you'll have to work late maybe...if a truck comes in and you want to make sure things get put in their proper place. Whenever possible just give a little bit of responsibility to the people working for you and let them know that you trust them, generally. I hope this will work for you.

D: Well, I'll try to do that.

C: If not just come back and we can try to work something else out. I almost feel you might be overworked.

D: I feel that way too.

C: I know you're frustrated too. I appreciate it. Maybe delegating these tasks and making sure the employees agree with what they are going to have to do, like I say even giving them a notecard with what they're going to do...uh...pitch in just occasionally to show them that you're not afraid to get your hands dirty and listen for feedback. And is there something I can do to improve you job or your working condition?

D: No. I don't think so.

C: Is there anything you feel like you need to talk about, or explain or...any gripes?

D: No. I just hope you don't think that I'm the problem.

C: Well, I see problems and I trust you're going to do what you can to work out the problems in your department.

D: Well, I will try.

C: I want you to be aware of them, that I notice things...that I realize its not your fault that all these matters are coming up. But I'd appreciate it if you would give it a try.

D: I will.

C: Well, I'm going to have to put you in for...what do you feel like your performance rating should be for the last 6 months ?

D: I think it should be at least a six. I'm working a 60-hour week and I think this department has been one of the best departments.

C: OK..do you feel like...what do you feel like is the highest level you could get to?

D: Well, 7 is top of the scale.

C: Well, how about if we let you work on these areas and would you feel like...a 5 is not a low rating compared to many managers who get less than 5. Would you feel like a 5 would be a tarnish on your reputation?

D: I would just feel like you're taking my people's performance more into consideration than my own actual performance.

C: I feel like the employee problems as far as...I know the fact that you are overworked is probably why you yelled at the staffer across the store that day which sort of embarrassed me...and the 2 people who asked you about the inventory...something you're in a position of respect and you have to constantly realize that you are looked up to. I feel like if you improve on that you could easily get a 6 or maybe even a 7 next time. How about if we just put you in for a 5 today and hope for some improvement?

D: OK.

C: OK. It's not a personal thing but I think you can look at the personal items as far as how you treated people when you're overworked and upset with them. Maybe you can put in not as many hours. Delegate responsibility and not therefore be so irritable at them, you know, because of their lack of performance.

D: OK.

C: OK. Is there anything else you want to add?

D: No, I don't think so.

C: Ok, that will be it Pat. Thanks.

D: Thanks.

Appendix B:

Definitions of the Performance Dimensions

Performance Dimensions

- Problem Analysis    The assessee asks questions to uncover unknown aspects of the problem or states how different parts of a problem are related.
- Problem Solution    The assessee suggests, recommends, or outlines one or more specific ways to resolve the problems.
- Sensitivity            The assessee shows concern for the individual and the individual's problems.



Appendix C:  
The BARS Rating Forms

### Problem Analysis

**Problem Analysis** - The assessee asks questions to uncover unknown aspects of the problem or states how different parts of the problem were related.

Assessee could be expected to relate the employee's lack of patience in the dealings with his subordinates to his long hours.	5
---	---

In response to the employee's comment about the poor quality of his subordinates' work, the assessee could be expected to ask the employee whether he told his subordinates what his standards were.	4
--	---

Assessee could be expected to ask the employee what he thinks could be done to improve his relations with his subordinates.	3
---	---

Assessee could be expected to inquire whether the employee has any questions about his responsibilities.	2
--	---

Assessee could be expected to inquire whether the employee had ever received any complaints from his subordinates but goes no further with this information or fail to engage in problem analysis.	1
--	---

Problem Solution

Problem solution - The assessee suggests, recommends, or outlines one or more specific ways to resolve the problems.

Assessee could be expected to outline what the employee should have done when discussing problem areas.	5
---	---

Assessee could be expected to suggest that the employee show his subordinates what he wants them to do rather than doing it himself.	4
--	---

Assessee could be expected to suggest that the employee sit down with his subordinates and attempt to develop a better working relationship with them.	3
--	---

Assessee could be expected to recommend that the employee try delegating more responsibility to his subordinates without explaining how.	2
--	---

Assessee could be expected to suggest that a goal could be obtained without specifying the manner in which it could be accomplished or fails to propose solutions to the problems.	1
--	---

Sensitivity

**Sensitivity** - The assessee shows concern for the individual and the individual's problems.

Assessee could be expected to express the desire to work with the employee to remedy the problems.	5
Assessee could be expected to compliment the employee on the responsibility he feels for his position.	4
Assessee could be expected to acknowledge that the employee's past performance appraisals were good.	3
Assessee could be expected to acknowledge that a lot of employees are apprehensive about the appraisal process.	2
In asking questions, the assessee could be expected to convey the impression that the employee was guilty until proven innocent.	1

Appendix D:

Checklist of Behaviors Used by Expert Raters

## Interview #1

### PROBLEM ANALYSIS

- \_\_\_\_\_ Asks whether there are workers in the department that do not belong there.
- \_\_\_\_\_ Asks what happens when workers do not do work that is assigned to them.
- \_\_\_\_\_ Asks how the employee is scheduling his workers.
- \_\_\_\_\_ Asks how the employee handled scheduling at his old store.
- \_\_\_\_\_ Does not ask questions concerning why the employee's subordinates are not coming to the employee with their problems.
- \_\_\_\_\_ Notes that the change to the new store may be related to the employee's lower performance ratings.

### PROBLEM SOLUTION

- \_\_\_\_\_ Suggests that the employee rotate the weekend schedule.
- \_\_\_\_\_ Suggests that the employee show his workers what to do and to remain with them until they are comfortable with the task.
- \_\_\_\_\_ Recommends that the employee teach his subordinates by sharing his experience and training.
- \_\_\_\_\_ Suggests that the employee give his subordinates a task to perform and to set a goal for them.

SENSITIVITY

- \_\_\_\_\_ Asks the employee how he likes being at the new store.
- \_\_\_\_\_ Asks how the employee has adjusted to the "big city life."
- \_\_\_\_\_ Clarifies that the purpose of the appraisal interview is to build for the future and to improve on everyone's performance.
- \_\_\_\_\_ States that it takes patience to show subordinates how to do a task and then to stay with them until they are comfortable with it.
- \_\_\_\_\_ States that he also has a problem being patient with his workers.
- \_\_\_\_\_ Tells the employee that if he needs help or needs any advice not to hesitate to come to him.
- \_\_\_\_\_ States that the employee has all of the right qualifications, and based on his past performance, he is sure that the employee will do a good job in the future.
- \_\_\_\_\_ Thanks the employee for coming in to see him.

## Interview #2

### PROBLEM ANALYSIS

- \_\_\_\_\_ Asks how things are going in the employee's department.
- \_\_\_\_\_ Asks if there are any problems.
- \_\_\_\_\_ Asks what kind of problems the employee is experiencing.
- \_\_\_\_\_ Asks if the employee is having problems with the responsibility that he delegates.
- \_\_\_\_\_ Asks if the employee feels that his subordinates are not handling their responsibilities.
- \_\_\_\_\_ Asks if there is a personality conflict between the employee and his subordinates.
- \_\_\_\_\_ Asks if the employee is dealing with his subordinates on a regular basis and giving them feedback.
- \_\_\_\_\_ Does not identify any relationships between problems.

### PROBLEM SOLUTION

- \_\_\_\_\_ Suggests that the employee let his subordinates know what is expected of them by setting up job descriptions and telling them how their performance will be measured.
- \_\_\_\_\_ Asks the employee what he thinks would be an appropriate solution to help with the scheduling and personnel problems, but does not use this information to suggest a solution.
- \_\_\_\_\_ States that they could work on establishing goals and means for delegating responsibility but does not discuss how they will go about it.



SENSITIVITY

- \_\_\_\_\_ Solicits the employee's input on the issues being discussed.
- \_\_\_\_\_ Acknowledges the difficulty in moving to a larger store.
- \_\_\_\_\_ Recognizes that the employee is putting forth a great deal of effort.
- \_\_\_\_\_ States that she appreciates the employee's attitude in trying to work with his subordinates.
- \_\_\_\_\_ Asks the employee if there is some way to relieve him of some of his work so that he is not bogged down.
- \_\_\_\_\_ Does not thank the employee for his time at the conclusion of the interview.

### Interview #3

#### PROBLEM ANALYSIS

- \_\_\_\_\_ Asks what happened when discussing why the employee yelled at on of his subordinates.
- \_\_\_\_\_ Asks if the crowd was unexpected when discussing the underordering of the picnic tables.
- \_\_\_\_\_ Does not identify any relationships between problems.

#### PROBLEM SOLUTION

- \_\_\_\_\_ Suggests that the employee move his subordinates around and have others work weekend nights.
- \_\_\_\_\_ Suggests that the employee talk to his subordinates about their schedules.
- \_\_\_\_\_ Suggests that the employee delegate more.
- \_\_\_\_\_ Tells the employee she will talk to some of his subordinates.

#### SENSITIVITY

- \_\_\_\_\_ Acknowledges that the employee had favorable recommendations in the past.
- \_\_\_\_\_ Does not greet the employee at the beginning of the interview.
- \_\_\_\_\_ Makes no effort to find out how the employee is adjusting to the new store.
- \_\_\_\_\_ Does not recognize the efforts of the employee.
- \_\_\_\_\_ Does not solicit the employee's input.
- \_\_\_\_\_ Does not express her willingness to help the employee to become a better manager or resolve the problems.
- \_\_\_\_\_ Does not thank the employee for his time at the conclusion of the interview.

**PLEASE NOTE:**

**This page not included with  
original material. Filmed as  
received.**

**University Microfilms International**

### Training Simulation One

What I would like to do now is provide you with the information that the expert raters used from the interview when evaluating the assessee's performance on the three dimensions.

First, lets look at the specific behaviors that were relevant to the problem analysis dimension. For two of the problem areas, the assessee asked some basic questions to obtain an overview of the problems. These problem areas were the performance of the subordinates and the scheduling complaints. When the employee complained that there were some people who should not be in the department, the assessee responded by asking if these people were destructive. The assessee also asked the employee what happened when the employee's subordinates did not complete their responsibilities, and how the employee handled scheduling at his previous store. In both of these instances, however, the assessee failed to pursue these problem areas to fully understand the problems. The questions that were asked were a good beginning but were broadly focused and did not funnel down to uncover specific information.

At another point in the interview the assessee did not probe into a discussion area that was initiated by the employee. The assessee did not pursue why the employee's subordinates were not coming to the employee with their problems.

Remember that noting relationships between problems is also a part of the problem analysis dimension. In this interview, the assessee did note that the change to the new store may have been a factor in the employee's lower performance ratings.

To re-emphasize these behaviors which were taken into account by

the expert raters when evaluating the assessee on the problem analysis dimension, I'd like to show you what these behaviors looked like again.

[show problem analysis behaviors]

Now lets look at the behaviors that the assessee engaged in which are relevant to the problem solution dimension. Throughout the interview numerous solutions to the problems were suggested. These included suggesting that the employee rotate the weekend schedule to give everyone a chance to have weekends off, and suggesting that the employee show his subordinates what to do and remain with them until they were comfortable with the task. In effect, the assessee was recommending that the employee teach his subordinates by sharing his experience and training. Some of the assessee's solutions, however, could have been more specific. This was illustrated when the assessee suggested that the employee give his subordinates a task to perform and to set a goal for them. This solution would have been more effective had the assessee outlined specifically the approach that the employee should take in doing so.

Again, lets take another look at this assessee as he engages in these problem solution behaviors.

[show problem solution behaviors]

Now that we've talked about and taken another look at the behaviors which were considered by the expert raters as being relevant to the problem analysis and problem solution dimensions, I would like to point out specific behaviors which were considered when evaluating the sensitivity dimension. The first behavior contributing to the sensitivity dimension was engaged in right at the beginning of the

interview when the assessee asked the employee how he liked being at the new store and how he had adjusted to the "big city life". This was interpreted by the expert raters as an attempt to set the employee at ease before discussing the employee's performance. The assessee also clarified that the purpose of the appraisal interview was to build for the future and to improve on everyone's performance. Later in the interview, the assessee empathized with the employee by stating that it takes patience to show subordinates how to do a task and then to stay with them until they are comfortable with it. The assessee noted that he had problems with this as well. The assessee also stated that if the employee needed help or advice that he should not hesitate to seek help. In addition, he showed confidence in the employee by stating that the employee had all of the right qualifications, and based on his past performance, he was sure that the employee would do a good job in the future. Also, in closing the interview, the assessee thanked the employee for coming in to see him.

Lets take another look at these "sensitivity" behaviors.

[show sensitivity behaviors]

### Training Simulation Two

Again, what I would like to do now is to look at some of the behaviors for each of the dimensions that the expert raters focused on when making their evaluations.

First lets look at the problem analysis dimension. The assessee initiated a line of questioning regarding the problems the employee was having with his subordinates. She began this line of questioning by asking how things were going in the employee's department and if there were any problems. She followed this with a question that asked

what the problems were, and then two questions that sought clarification whether the subordinates were not handling the delegated responsibility. The assessee, however, did not ask specific questions to gather additional information. Beyond the problems the employee was having with his subordinates, the assessee did not ask specific questions that would help her to understand the problems presented. The assessee asked if there was a personality conflict between the employee and his subordinates, and how the employee was dealing with his subordinates, but she did not continue with this line of questioning to identify all possible sources of the problems. The assessee also did not identify any of the relationships that may have helped to explain the problems.

Lets take another look at the assessee as she engages in problem analysis.

[show problem analysis behaviors]

Now lets look at the behaviors the assessee engaged in that are relevant to the problem solution dimension. During the interview, the assessee generated one solution. She suggested that the employee let his subordinates know what is expected of them by setting up job descriptions and informing subordinates as to how performance would be measured. While this is an appropriate solution, it did not include a complete plan of action to insure it would be accomplished. The assessee also asked the employee what he thought would be an appropriate solution to help him with his scheduling and personnel problems, but did not use the information to suggest a solution. Also, the assessee stated that they could work on establishing goals, and means for delegating responsibility and holding the subordinates

accountable, but never pursued this suggestion or discussed how they would go about it.

Here is another look at these problem solution behaviors.

[show problem solution behaviors]

Now lets focus on the sensitivity dimension. Throughout the interview, the assessee solicited the employee's input on the issues that were being discussed. She also acknowledged the difficulty in moving to a much larger store, and recognized that the employee was putting forth a great deal of effort. The assessee stated that she appreciated the employee's attitude in trying to work with the subordinates. In asking the employee if there was some way to relieve him of some of his work so that he is not bogged down, the assessee was expressing a desire to help the employee remedy his problems. It also should be noted that the assessee did not thank the employee for his time at the conclusion of the interview.

Here are some of the sensitivity behaviors that were considered in evaluating this dimension.

[show sensitivity behaviors]

### Training Simulation Three

As we have done for the other interviews, I'd like to take some time and look at the behaviors that the assessee engaged in during this interview that were taken into consideration by the expert raters when the performance of the assessee was evaluated.

First lets look at the problem analysis dimension. Throughout the role play, the assessee failed to pursue any of the problem areas with questions in order to understand them more completely. In fact, only two questions were asked during the course of the interview. In



discussing why the employee yelled at one of his subordinates, the assessee asked what happened. And later on during the interview, when discussing the ordering problem, she responded to the employee's comment that the picnic tables were underordered because of the crowd rush by asking if the crowd was unexpected. In both of these instances, the assessee did not pursue these areas any further. In addition, the assessee did not identify either of the two relationships that may have partially explained the problems.

Lets take a look at the questions that she did ask.

[show problem analysis behaviors]

With regards to problem solution, the assessee suggested a few solutions during the course of the interview. First she recommended that the employee move his subordinates around and have others work weekend nights, and to talk to the subordinates about their schedules. In addition, she told the employee to delegate more, and that she would talk to some of his subordinates. For these problem areas that the assessee did suggest solutions, the expert raters felt the solutions were shallow and did not include outlines of action plans that would be followed to resolve the problems.

Lets take a look at these behaviors.

[show problem solution behaviors]

In terms of sensitivity, the expert raters felt that the assessee engaged in one sensitive behavior during the interview. The assessee did acknowledge that the employee hade favorable recommendations in the past. The assessee made no effort to find out how the employee was adjusting to the new store, did not recognize the efforts of the employee, and did not solicit the employee's input regarding the

problems that were being discussed. The assessee did state that she would talk to some of the employee's subordinates, however, she did not express her willingness to help the employee become a better manager or resolve the problems that he was experiencing. In addition, the assessee did not greet the employee at the beginning of the interview, and did not thank the employee for his time at the interview's conclusion.

Lets look at some of these behaviors that are relevant to the sensitivity dimension.

[show sensitivity behaviors]

Appendix F:

Target Scores for the Training and Experimental Interview Simulations

	<u>Problem Analysis</u>	<u>Problem Solution</u>	<u>Sensitivity</u>
Training Simulation 1	2.0 (0.0)	4.0 (0.0)	4.2 (0.45)
Training Simulation 2	3.0 (0.0)	2.2 (0.45)	4.4 (0.55)
Training Simulation 3	1.0 (0.0)	2.0 (0.0)	1.4 (0.55)
Experimental Simulation 1	1.0 (0.0)	1.0 (0.0)	3.8 (0.84)
Experimental Simulation 2	2.0 (0.0)	2.4 (0.55)	4.0 (0.0)
Experimental Simulation 3	1.0 (0.0)	1.8 (0.45)	1.2 (0.45)
Experimental Simulation 4	3.0 (0.0)	2.0 (0.0)	3.0 (0.0)
Experimental Simulation 5	2.8 (0.45)	4.0 (0.0)	4.0 (0.0)
Experimental Simulation 6	4.0 (0.0)	3.2 (0.45)	4.0 (0.0)
Experimental Simulation 7	3.0 (0.0)	2.8 (0.45)	5.0 (0.0)

Note. Standard deviations are included in parentheses.

Appendix G:  
Informed Consent Form

INFORMED CONSENT FORM

You are invited to participate in a study which will examine how accurately people can observe and evaluate performance. If you decide to participate in this study, you will be asked to view and rate several videotaped performance evaluation interviews between a store manager and a new department manager in that store. During the first phase of the study you will be familiarized with the details of the performance evaluation interview and the materials that will be used in observing and evaluating the performance of the store manager. In the second phase, you will actually observe and rate the interviews.

Any information obtained from you in connection with this study will remain confidential. This informed consent form is the only form on which you will be asked to record your name. You will be assigned an anonymous identification number to record on all other materials. Only the experimenter and the principal investigator will have access to individual data. Results of the study, if published or presented at scientific meetings, will be presented in group form so that there is no way that individual participants can be identified. You are free to withdraw your consent and terminate your participation at any time, without penalty.

If you decide to participate you will be compensated \$40. Please note that you will only receive this compensation if you complete all phases of the study.

You are making a decision whether or not to participate in this study. Your signature indicates that you have decided to participate. In addition, it indicates that you have read and understand the information provided above. If you have any questions about this study and what is required of you, please ask the experimenter prior to signing this form.

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature of Participant

\_\_\_\_\_  
Signature of Witness

\_\_\_\_\_  
Signature of Principal Investigator

Appendix H:  
Script of Videotaped Presentation

I'd like to begin today by giving you a brief explanation of the nature of this research project. The main focus of this project is to examine the manner in which people observe and evaluate performance. More specifically, the project is concerned with how the performance of managers or potential managers is observed and evaluated in management assessment centers.

The assessment center is a method that is used to evaluate individuals' management skills. Information provided by an assessment center is useful either for selecting people that should be hired or promoted to some management position, or evaluating the strengths and weaknesses of current managers in an organization.

Individuals that are being evaluated in an assessment center, which we will be referring to as assessee's, participate in a number of tests and exercises which are developed to measure specific management skills that have been shown to be important parts of the management positions in question. As an example, if oral presentation was an important part of the job, an exercise may be included in the assessment center in which the assessee's are provided with some information and then asked to actually make a presentation.

Performance on the various exercises is observed and evaluated by a group of assessors. These assessors are typically managers from the organization and are familiar with the job and the specific dimensions of managerial performance that are of interest. This process is the focus of this research project. That is, we are interested in how people observe performance on an assessment center exercise and then use that information to make evaluations on specific dimensions of



performance.

In addition to providing us with the information that we need regarding the manner in which performance in an assessment center exercise is observed and evaluated, we hope your participation will be a valuable learning experience. It is our intention that you will come away from this research with a better understanding of the complexity involved in observing and evaluating performance, and to improve your skills in doing so. Through your participation, you also should develop a meaningful insight to the assessment center process which should prove to be useful at some time in your careers.

## II. Exercise

The task that you will be asked to complete in this project will be to observe and evaluate the performance of several students who participated in an assessment center that was conducted as one of the first phases of this research project. This assessment center actually consisted of several exercises, however, for the purposes of this study we will only be interested in one of those exercises.

The performance on this exercise of each of the assessee's was evaluated by a group of psychologists who are experienced with assessment center methodology, and have considerable expertise with the processes involved in performance evaluation. The means of these expert raters' scores have been taken to define a set of Target Scores. The ratings that you make during the study will be compared with these Target Scores.

What I'd like to do now is describe the exercise that you will be observing and evaluating. The exercise is an interview simulation in which the assessee's were asked to play the role of a store manager

for a department store. For clarity, you can think of the store as something like a Sears or J.C. Penneys. As the store manager, the assessee's were responsible for conducting regular performance evaluations of their department managers. For the exercise, the assessee's had to conduct a performance evaluation interview with a new department manager who was having some performance problems. The objective for the assessee's was to identify the true nature of the problems and develop plans for resolving them.

What you will be observing are videotaped re-enactments of the assessee's conducting this performance evaluation interview with one of their department managers. In the videotapes, the same person will always play the role of the department manager.

The procedure that was followed in the assessment center was to provide the assessee's with information about the role that they were to play, give them time to prepare what they were going to say or the approach that they wanted to take, and then instruct them to conduct the performance evaluation interview in whatever fashion they felt was most appropriate. What I'd like to do now is to show you the actual information about the situation that the assessee's were given.

[pause to handout and read roleplay description]

### III. Dimensions

As I mentioned earlier, assessment center exercises are developed to measure certain skills, or dimensions of managerial behavior that have been shown to be important for being a good manager. We will be using the performance evaluation interview simulation to measure how well the assessee exhibits three dimensions of the manager's job.

The first dimension that we will be evaluating is problem analysis. Please look at the definition for problem analysis that is given on the "Dimension Definitions" handout. The definition reads, "the assessee asks questions to uncover unknown aspects of the problem or states how different parts of a problem are related."

Please note that this dimension involves two types of behaviors. First is the behavior of asking questions to uncover unknown aspects of the problem, while the second is stating how different parts of a problem are related.

The second dimension is Problem Solution. Please look at the definition for Problem Solution on the handout. The definition reads "The assessee suggests, recommends, or outlines one or more specific ways to resolve the problems."

The third dimension is Sensitivity. As you can see on the handout, the definition of the sensitivity dimension is "the assessee shows concern for the individual and the individual's problems."

[BARS]

After you have reviewed an interview we would like you to evaluate how well the person conducting the interview, the assessee, performed on the Problem Analysis, Problem Solution, and Sensitivity dimensions. The forms that will be used to evaluate the performance will be handed out now.

[PAUSE WHILE BARS HANDOUTS ARE PASSED OUT]

Now that you have the evaluation forms note that there is one form for each dimension and the definition for the dimensions are printed on the top. Underneath the dimensions are five statements which are meant to represent five different levels of performance with the number 5 statement being the highest level of performance and the number one statement being the lowest. In using this form to evaluate the performance of the assessee we would like you to circle the number corresponding to the statement which most accurately reflects the level of performance that the assessee demonstrates on that dimension. Please use only the whole numbers 1, 2, 3, 4, and 5 when making your ratings. That is, please don't select ratings that are in between these numbers.

In order to circle the number which corresponds to the statement which most accurately reflects the level of performance demonstrated in the interview, it is necessary to take into account all of the behaviors relevant to a dimension that are demonstrated in an interview. If you look at the statements, you will notice that each of them includes the phrase "Assessee could be expected to ...". This means that when you take into account all of the behaviors that the assessee demonstrated which are relevant to the dimension in question,

you would expect the overall level of effectiveness on the dimension to be the same as the level of effectiveness reflected in the statement. It is important to keep in mind that the statements on the form are meant to represent different levels of effective performance. Sometimes the actual behaviors that occur in an interview will also be in one of the statements, and sometimes they won't. Since the statements are meant to represent different levels of overall performance on a dimension, just because a behavior which is listed in a statement occurs in an interview does not necessarily mean that the number corresponding to that statement should be circled. The statement that should be circled is the one which reflects the level at which you would expect the assessee to consistently perform. In order to determine this you would have to take into consideration all of the relevant behaviors that were exhibited. If, for example, an assessee engaged in a behavior which was listed as a level four behavior but also exhibited behaviors which could be considered level 1 or level 2 behaviors when all of the behaviors were taken into account the level which you would expect the assessee to consistently perform at would most likely be somewhere lower than the 4.

I would also like to point out that in addition to taking all of the behaviors into account when making a rating, it is important to carefully read each of the statements. With this evaluation form you cannot look only at the numbers and assume that a five is excellent, a three is average, and a one is poor. The number that is circled should correspond with the statement which best represents the level of performance at which, based on the behaviors exhibited in the interview, you would expect the assessee to consistently perform.

At this point lets pause for a minute to see if there are any questions about how to evaluate the assessee's performance.

[PAUSE FOR QUESTIONS]

Ok, what I'd like to do next is familiarize you with the actual statements that are listed on the evaluation forms. First lets look at the problem analysis form. Note that on the top of the form the definition of problem anlysis is listed. Underneath the definition are the five statements with number 5 representing the highest level of behavior and number 1 representing the lowest level.

The number 5 statement would be circled if the level of performance demonstrated in the interview was such that the assessee could be expected to relate the employee's lack of patience in his dealings with his subordinates to his long hours.

The number 4 statement reads, in response to the employee's comments about the poor quality of his subordinates' work, the assessee could be expected to ask the employee whether he told his subordinates what his standards were.

The number 3 statement would be circled if the level of performance demonstrated in the interview was such that the assessee could be expected to ask the employee what he thinks could be done to improve his relations with his subordinates.

The number 2 statement reads that the assessee could be expected to inquire whether the employee has any questions about his responsibilities.

The number 1 statement says that the assessee could be expected to inquire whether the employee had ever received any complaints from his subordinates but goes no further with this information or fails to

engage in problem analysis.

Before we go on to the next dimension, I'd like you to take a minute and read over these statements to familiarize yourself with them. This also is a good time to ask any questions that you might have about these statements or anything else to do with the Problem Analysis dimension.

[BRIEF PAUSE]

Ok, now lets look at Problem Solution. Again, please note that the definition for Problem Solution is listed at the top and the five statements are written underneath the definition.

Lets look at the number 5 statement. It reads that the assessee could be expected to outline what the employee should have done when discussing problem areas.

Looking at number 4, this statement would be circled if the level of performance observed on the tape was such that the assessee could be expected to suggest that the employee show his subordinates what he wants them to do rather than doing it himself.

The number 3 statement reads that the assessee could be expected to suggest that the employee sit down with his subordinates and attempt to develop a better working relationship with them.

Statement number 2 says that the assessee could be expected to recommend that the employee try delegating more responsibility to his subordinates without explaining how.

The number 1 statement would be circled if the level of performance were such that the assessee could be expected to suggest that a goal could be obtained without specifying the manner in which it could be accomplished or fail's to propose solutions to the

problems.

Again, I'd like to pause for a minute to give you a chance to read over the statements and ask any questions that you might have about these statements or anything else about the Problem Solution dimension.

[BRIEF PAUSE]

Ok, now lets look at the Sensitivity dimension. The number 5 statement would be circled if the level of performance observed was such that the employee could be expected to express the desire to work with the employee to remedy the problems.

The number 4 statement reads that the assessee could be expected to compliment the employee on the responsibility he feels for his position.

The number three statement says that the assessee could be expected to acknowledge that the employee's past performance appraisals were good.

The number two statement would be circled if the level of performance observed was such that the assessee could be expected to acknowledge that a lot of employees are apprehensive about the appraisal process.

And last of all, the number one statement reads that in asking questions, the assessee could be expected to convey the impression that the employee was guilty until proven innocent.

Again, at this time I would like to pause briefly to give you a chance to read over these statements and ask any questions that you might have about the Sensitivity dimension.

[BRIEF PAUSE]



Appendix I:

Background Information for the Interview Simulation Exercise

## INTERVIEW SIMULATION

In this exercise you are Chris Harmon, store manager for KENDALL #66. KENDALL is a large chain of retail department stores. You have been the store manager for three years. There are 12 department managers who report directly to you. One of the standard policies of KENDALL #66 is to conduct semi-annual performance evaluation meetings with each of the department managers. One of the department managers is Pat Winchell.

Pat is the manager of the Lawn Furniture department. Pat was recently transferred to KENDALL #66 from KENDALL #15, which is a smaller volume store. Pat comes to KENDALL #66 with favorable recommendations from KENDALL #15 store manager. In the past Pat has received especially good performance evaluation ratings. This is your first performance evaluation meeting with Pat, since Pat first joined KENDALL #66 four months ago.

It has come to your attention that at certain times Pat has shown poor decision-making judgments. Pat has frequently made hasty decisions, based on assumptions and emotions, instead of relevant information. For example, there was the time that Pat ordered picnic tables without checking last year's inventory records. This resulted in the under-ordering of much needed merchandise. Also, Pat has repeatedly scheduled the same full-time employees to work weekend nights. This has led to several employee complaints.

You have also noticed that there are a number of things in the department that don't get done, even though Pat works nearly 60 hours per week. Pat even comes in at off hours to supervise the department. On one occasion you have observed that Pat does the work that a staffer should be doing. Some of the staffers in Pat's department have expressed their dissatisfaction with having so little responsibility, and you suspect that Pat is one of those people who has to do everything, rather than relying on the help of others.

In addition, you have been informed that Pat is often too demanding and does not display the patience and concern for others that the staffers desire. Pat, on at least one occasion, yelled at a staffer who did not remember if a piece of merchandise was still in stock. Moreover, two staffers have asked Pat to explain how the inventory system works, and Pat only replied, "I suggest you find out soon".

Today is November 19, 1985, the day of your meeting with Pat. Your goal is to discuss Pat's performance evaluation and to resolve any problems. You may handle the situation any way that you feel is appropriate. Act as if the situation were real.

AT THIS POINT, IF YOU ARE UNCLEAR ABOUT YOUR ROLE, ASK FOR CLARIFICATION.

Appendix J:  
Script for Training Role Play

Training Script for Interview Simulation

A: Hello Pat. How are things going?

D: Not bad. I have been pretty busy, but I think things are going pretty well.

A: How do you like being here at this store?

D: It's OK. It's taking a little time to get comfortable with all the changes, but basically I really like it here.

A: What kind of changes are you referring to?

D: Well, there are alot more customers with the higher volume, and I have alot more staffers here than at my other store.

A: Yes. I realize it is difficult to get used to things when you move to a bigger store. But based on your past performance, I'm sure you'll do fine here. Looking at the recommendations you had, I can see why you were promoted to this store.

D: I really like this company and would like to move up.

A: OK. Well that's really what I wanted to talk you about. I can see that you're really putting in alot of time effort and that tells me alot about how serious you are about your job. What we want to do here is to take a look at your performance and see how we can improve upon it because I consider it my job to help you move up. It's important for you to let me know what I can do to help you resolve any problems you may have.

D: I appreciate that.

A: There are some things that have been brought to my attention that I would like to discuss with you. You mentioned earlier that it was taking you some time to get comfortable with the changes here. Are you having any problems with that?

D: No, I don't think so.

A: OK. Before we get started, is there anything that you would like to discuss?

D: Well there have been a couple of things, but I don't think anything that more time here won't resolve.

A: One of the things that concerns me is that you've made some questionable decisions?

D: I'm not sure I know what you mean.

A: Let me give you an example. Whenever you place an order for merchandise, it is important that you go back and check previous inventory records to give you some idea of how much to order.

D: Are you talking about the picnic tables?

A: Yes. You underordered on those because you didn't check the inventory. That cost us.

D: I don't consider that my fault.

A: Whose fault was it?

D: We had such a crowd rush that no one could have guessed how many we needed. I thought I ordered the right amount.

A: OK. But in the future, I think it's important that you check the inventory records because we lost a lot of customers by not having the tables. You've got to pay attention to little details like that. Another area that I think we need to talk about is scheduling. I assume that you were responsible for the scheduling at the other store.

D: Right.

A: Did you have any problems with that?

D: None that I was aware of.

A: Well it has come to my attention that some of your subordinates are not happy with the way you have been writing the schedule. Has anyone voiced any concerns about this to you?

D: They have complained to you? No, no one has said anything to me about this.

A: It seems that you have been scheduling your full-time employees to work weekend nights.

D: At my other store that was what my full-timers wanted. They could make their most money then in commission.

A: Have you talked to your employees to see if that were true here?

D: No.

A: OK. I think you need to sit down and talk with your employees and see if they have particular preferences as to when they would want to work, especially the full-timers. It's not safe to assume that people here will prefer the same schedule as those at your other store. It's important that you consider these preferences because it shows that you are concerned about them.

D: That's fine. I just wish that if these people had problems they would talk to me first. I told these people when I came here that I had an open door policy but no one has approached me with any problems.

A: Let's talk about that. There may be a reason why your people are not coming to you. I have noticed that you seem to be a little impatient when responding to your employees.

D: I'm not sure I know what you mean.

A: Well I know of one incident where you snapped at a staffer who had asked you about the inventory. Can you tell me about that?

D: I think you are referring to the incident with John. He had been slacking off all day and he saw this as another opportunity for me to do his work.

A: Have you had other problems with John?

D: No, not really.

A: Have you had problems with any of the others in your department?

D: There have been a few.

A: Can you explain?

D: I may tell some individuals to do something, but, I don't know, they don't seem to do it very well or sometimes they don't do it at all.

A: Can you give me an example?

D: Well, for example, I told someone to set up a display in the front of the store. Later I went up there to check it and it was a mess. I had to redo it myself which took an hour that I didn't have.

A: Do you have any thoughts on why these people are not performing their jobs?

D: I just think there are some people here that don't want to work. I think we need to get rid of some of these people and replace them with people who want to work. Either that or give them more money. I feel like that you get what you pay for. We pay these people minimum wage and that is the type of help we get.

A: OK. Let's think about that. First of all, to simply fire everyone and replace them is not very cost effective. We would have to go through the whole process of selection and retraining if we did. It's also difficult to pay them much more than minimum wage because the profit margin of the store is so small. Let me ask you, do you think more training would help those people?

D: I'm not sure.

A: Do you think they know what you expect of them and how you like the work to get done?

D: They should know. These people have been here a lot longer than I have.

A: Yes, but you have to remember that these people may be much different than people you worked with at your other store. Because you're new here, people may not know what you expect of them so it's important that you make this clear up front. How do you feel about this?

D: I agree. But I have high standards and I expect people to give me their best work.

A: And you should expect that. But I think it's also important to try to look at this from the employee's perspective. Some have complained that they don't feel you are giving them enough responsibility. I have noticed the number of hours that you are working and I appreciate your dedication. But it seems you do a lot of the work that your employees should do.

D: Well, I'm ultimately responsible for how well this department is run. If they don't get it done, then I have to do it. That's why I'm working 60-hour weeks.

A: Are you saying then that you work so many hours because your employees aren't doing the work you delegate to them?

D: That's exactly what I'm saying. I never had this problem at my other store.

A: Well I think you are ultimately responsible for this department, and you have the authority to take the appropriate action when needed. You have to let them know that you are in charge of the department.

A: It seems your working so many hours may account for some of the other problems you've had like losing your patience. Do you have any suggestions, other than firing them or giving them raise, that would improve how you work with your employees and how they work with you?

D: No. I wish I did.

A: OK, let me suggest something. Perhaps you could sit down with your employees and get a feel for some of the things that concern them. For example, the scheduling. Find out if they understand what you expect of them, and give them more responsibility. Perhaps you could have shown John how you wanted the display to be set up for example. You have to remember that these people may want to move up in the company just like you. If that's the case, you need to use your expertise so that they understand the importance of doing their jobs properly for themselves and for the business.

D: I have tried to delegate on more than one occasion.

A: And what happened when it wasn't done?

D: I did it myself.

A: Do you know what we expect of you?

D: Yes I think so.

A: Then you know that your primary responsibility is to manage and not just be another employee, and it is as a manager that we can best use you. You can't always do their work and get the things done that you need to as department manager. Probably in your other store you could do those things more often because it was a smaller store. But the size of this store makes it almost impossible to operate that way and I'm sure it is difficult to get used to that. But you can't do their work for them.

D: Things would be much easier if they completed what I delegated to them.

A: Let's do this. Meet with your employees just as I am doing with you. It doesn't have to be a formal meeting. It is probably better that you do it informally. Discuss their concerns, let them know what you expect, and how you will evaluate their performance, and reinforce them when they do the job correctly. But it's important that you give them more responsibility so that you can spend more time with your responsibilities. You need to do it now so that the problem gets no larger and you don't burn yourself out working so many hours. What do you think about this?

D: I will certainly try.

A: If this doesn't work then you have the responsibility to take the appropriate steps even if that means cutting their hours back or letting them go. But show them first what you expect and let them do it. That way if they know what you expect and that they will be held accountable you won't have to be concerned when you're doing what you need to do or when you're not there. In fact, you may want to train someone to step into your role so that you can move up.

D: I understand.

A: I'd like to see you reduce the number of hours that you're working to maybe about 45 or so in the next three weeks but it all starts with you communicating with your employees. Let's try this for a couple of weeks and see what happens. Then we can get back together to evaluate how this is working out.

D: Fine

A: Ok. Thanks for coming in and if there is anything I can help you



with in the meantime, just let me know.

Appendix K:  
Post-Experimental Questionnaire

## PARTICIPANT REACTION QUESTIONNAIRE

1. RATER #: \_\_\_\_\_

2. Sex: Male      Female      (circle one)

3. Age: \_\_\_\_\_

4. Ethnic Background: White Black Hispanic Asian Other (circle one)

5. Class: FR SO JR SR Grad Student      (circle one)

PLEASE CIRCLE THE NUMBER THAT MOST CLOSELY REFLECTS YOUR ANSWER TO EACH OF THE FOLLOWING QUESTIONS.

6. To what extent did the information presented in Session 1 help you to accurately evaluate the assessee?

5.....4.....3.....2.....1  
 completely                      to some                      not at all  
    degree

7. To what extent was the information presented in Session 1 understandable?

5.....4.....3.....2.....1  
 completely                      to some                      not at all  
    degree

8. To what extent was the information presented in Session 1 useful for making accurate ratings?

5.....4.....3.....2.....1  
 completely                      to some                      not at all  
    degree

9. To what extent did the information presented in Session 1 improve your ability to make accurate ratings?

5.....4.....3.....2.....1  
 completely                      to some                      not at all  
    degree

10. To what extent do you feel confident that your ratings are accurate measures of the assessee's performance?

5.....4.....3.....2.....1  
 completely                      to some                      not at all  
    degree

11. To what extent was the information presented in Session 1 logical?

5.....4.....3.....2.....1  
 completely to some not at all  
 degree

12. To what extent was the information presented in Session 1 confusing?

5.....4.....3.....2.....1  
 completely to some not at all  
 degree

13. To what extent did you enjoy Session 1?

5.....4.....3.....2.....1  
 completely to some not at all  
 degree

14. To what extent did you enjoy Session 2?

5.....4.....3.....2.....1  
 completely to some not at all  
 degree

15. To what extent was the experiment a learning experience for you?

5.....4.....3.....2.....1  
 completely to some not at all  
 degree

16. To what extent would you be interested in participating in another research study similar to this one?

5.....4.....3.....2.....1  
 completely to some not at all  
 degree

17. Please list other research projects that you have participated in this year.

18. Please use this space to make any comments that you would like concerning the research study.

## AUTOBIOGRAPHICAL STATEMENT

Todd A. Silverhart was born on October 7, 1959 in Glens Falls, New York. After graduating from Saratoga High School (Saratoga Springs, New York) in 1977, he attended Middlebury College (Middlebury, Vermont). In May 1981, the author graduated Cum Laude from Middlebury College with a Bachelor of Arts degree and High Honors in Psychology. Before pursuing his graduate education, the author served one year on a pre-professional mental health internship with The Devereux Foundation in Devon, Pennsylvania.

Todd A. Silverhart's graduate studies were conducted at Old Dominion University (Norfolk, Virginia). A Master of Science degree in Psychology was earned from that institution in August, 1984. The title of his master's thesis was The Effects of Training and Rating Formats on the Accuracy of Performance Ratings. While at Old Dominion University, the author worked as an instructor of psychology for courses in quantitative methods and personnel psychology, as a teaching assistant, and as a research assistant. During his tenure as a graduate student, the author was inducted into the Phi Kappa Phi Honor Society, named to the National Dean's List for Graduate Students, and awarded a University Fellowship.

The author's professional affiliations include membership in the following organizations: Academy of Management, American Psychological Association, Personnel Testing Council of Metropolitan Washington, Society for Industrial-Organizational Psychologists, and the Southeastern Psychological Association.