The Effects of Training, Sex of the Rater, and Contrast in the Evaluation of LGD Behavior

1973

William Richard Garrett
University of Central Florida

Find similar works at: http://stars.library.ucf.edu/rtd

University of Central Florida Libraries http://library.ucf.edu

Part of the Psychology Commons

STARS Citation

http://stars.library.ucf.edu/rtd/52

This Masters Thesis (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of STARS. For more information, please contact lee.dotson@ucf.edu.
THE EFFECTS OF TRAINING, SEX OF THE RATER, AND CONTRAST IN THE EVALUATION OF LGD BEHAVIOR

BY

WILLIAM RICHARD GARRETT
B.A., Florida Technological University, 1971

THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Psychology in the Graduate Studies Program of Florida Technological University

Orlando, Florida 1973
The author would like to thank Dr. Cabot Jaffee, Dr. Wayne Burroughs, and Dr. Fredric Frank for assistance and guidance in the completion of this research. Thanks would also like to be extended to Dr. Ed Shirkey for assistance in statistical procedures.

A special thanks would like to be extended to Dennis Peache, lab technician, for a great deal of help in procuring equipment and technical assistance without which this research would have been impossible.
# Table of Contents

<table>
<thead>
<tr>
<th>Item</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgement</td>
<td>iii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>iv</td>
</tr>
<tr>
<td>List of Tables</td>
<td>v</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Method</td>
<td>8</td>
</tr>
<tr>
<td>Results</td>
<td>11</td>
</tr>
<tr>
<td>Discussion and Conclusions</td>
<td>23</td>
</tr>
<tr>
<td>Appendix A</td>
<td>29</td>
</tr>
<tr>
<td>Appendix B</td>
<td>32</td>
</tr>
<tr>
<td>References</td>
<td>34</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cell Means and Standard Deviations</td>
<td>12</td>
</tr>
<tr>
<td>2. Analysis of Variance</td>
<td>13</td>
</tr>
<tr>
<td>3. Mean Ratings</td>
<td>15</td>
</tr>
<tr>
<td>4. Comparisons for Contrast Effects Across Levels of Training</td>
<td>16</td>
</tr>
<tr>
<td>5. Mean Number of Positive and Negative Behaviors Checked</td>
<td>18</td>
</tr>
<tr>
<td>6. Analysis of Variance of Positive Behaviors Checked</td>
<td>20</td>
</tr>
<tr>
<td>7. Analysis of Variance of Negative Behaviors Checked</td>
<td>21</td>
</tr>
<tr>
<td>8. Analysis of Variance of Simple Effects of Levels of Training for Negative Behaviors Checked</td>
<td>22</td>
</tr>
</tbody>
</table>
In the past decade much research has been done in the area of
the importance of contrast effects in social perception. Generalizing
from the area of contrast in psychophysics, Holmes and Berkowitz
(1961) hypothesized that a person, when judging two stimulus individ-
uals, forms as a frame of reference his experience with the first
person. Contrast then occurs in the perception of the second person
when the difference between the two is great. The stimuli in the study
were taped conversations between a psychologist and a student. If the
psychologist, when presented first, was "belligerent" and was followed
by the student (non-belligerent, non-aggressive) the immediate differ-
ence between the two was perceived to be greater. The student, when
evaluated on a personality measure, would be evaluated as a much more
pleasant individual than under the opposite condition where the
student was preceded by a "friendly" psychologist.

Rowe (1967) found similar results using hypothetical individuals
whose descriptions were made up of six adjectives. Subjects were
asked to choose those individuals described by the adjectives that
they would like to have as a friend. The descriptions were composed
of different proportions of favorable (F) and unfavorable (U) adject-
ives. Each subject was shown 103 hypothetical descriptions. The
first 30 and the last 10 were the same type (i.e., 3F-3U type).
These depicted an average individual with equal proportions of favor-
able and unfavorable adjectives. The test descriptions were arranged
into three different orders; ascending order in which the descriptions
started with the CU type and progressed proportionally through to the
6F type (6U, 5U-1F, 4U-2F, ... 6F), descending order in which the descriptions started with the 6F type and progressed proportionally to the 6U type, and random order in which the descriptions were randomly placed. Results showed that more "like" responses were recorded in the ascending order than the descending or random order. The authors concluded that if one first encounters an unpleasant individual in a social situation he would be more likely to see an average person as more favorable than if he first encounters a very pleasant individual.

Hakel, Ohnesorge, and Dunnette (1970) used specially prepared sets of resumes to investigate the importance of contrast effects in the employment decision. Each subject was shown three resumes. The first two were always the same quality (either high, average, or low) and were used to establish a frame of reference for the subject. The third resume was used to test for contrast effects and was either high or low in quality. This resulted in six sequences; HHH, HHL, AAH, AAL, LLH, LLL. The authors found that the evaluation of the high resume in the third position were highest when compared with a contrasting sequence (LLH greater than HHH). The low resume received its lowest rating in the third position when compared with a contrasting qualification (HHL lower than LLL).

Similar results were found by Wexley, et al. (1972) using videotaped interviews of prospective job candidates. Each tape represented an applicant with different qualities (high, low, or average). Each subject was shown three videotapes in the experimental condition. The first two were always of the same quality condition and were used to establish a frame of reference for the subject. The third tape was
used as a measure of contrast and was either high, low, or average in quality which resulted in eight sequences: HHH, HHA, HHL, AAH, AAL, LLL, LLH, LLA. It was found that a high quality individual in the third tape position received their highest ratings when preceded by two low's as opposed to being preceded by two high's. Similarly, a low quality individual received his lowest ratings when preceded by two high's rather than two low's.

Another area of interest in social perception involves the effects of differential training and experience on the ability of individuals to evaluate others in social situations. One of the first studies done in this area was undertaken by Martin (1938) in an attempt to determine practice effects in judgments of traits. His procedure involved having judges view an individual and rate him on some objective traits (weight, height, age, etc.). The subjects then viewed the individual in an interview situation and rated him on some more subjective traits (introversion-extroversion, and intelligence). After ratings had been completed, the experimenter compared these with data gathered before the session began. Subjects' ratings were discussed and errors pointed out, after which more practice sessions were run. Martin found that typical trial and error were evident for each of the traits observed. He noted, however, that estimates of the more subjective traits demonstrated greater variability than did the more objective traits.

Carlson (1967) using hypothetical applicants described by written information accompanied by a photograph, found no significant differences in the inter and intrarater agreement among and within managers.
with different lengths of experience (less than 3 1/2 years, 3 1/2-5 years, more than 5 years) or type of experience (less than 140 interviews conducted, 141-300 interviews conducted, more than 300 interviews conducted) for either ratings or rankings of photographs or applications constructed of written information.

Through further analysis Carlson found that high tenure-low activity managers would offer more contracts on the basis of photographs while the low tenure-high activity managers offered more contracts based on written factual information. The degree then to which managers rely on factual information is directly related to the rate of actual interviewing rather than on the amount of time spent on the job. Greater experience in actual interviewing increases the ability to pick out and rely on relevant data in making employment decisions.

In a study designed to show that untrained persons have had enough experience with people to differentiate between a facilitative and a harmful therapist, Shapiro (1968) used tapes of therapy sessions to compare the ratings of naive, untrained undergraduates with those of 13 trained raters with 20-100 hours of experience. The results were tabulated over three categories; empathy, warmth, and genuineness. Results indicated that ratings made by untrained, naive subjects correlated .67 on empathy, .89 on warmth, and .78 on genuineness. It can be seen then that untrained raters evaluated therapy sessions in much the same way as did raters with a much greater degree of experience and training.

Greenwood and McNamara (1967) carried out a study to investigate the interrater reliability of managers in situational tests. Ratings were compared between professional (psychological and sociological
personnel) assessors and non-professional (high level management with a minimum of training in assessment techniques) assessors. Results indicated that reliabilities appeared equivalent and that neither group appeared to give consistently higher or lower ratings of participants, nor did either group have higher interrater reliability than the other.

Hakel, Ohnesorge, and Dunnette (1970) found that contrast effects (using written resumes) occurred for both experienced and inexperienced interviewers. Ratings were compared between 97 employment interviewers with experience ranging from 1 month to 23 years (median 5 years) and actual interviews conducted ranging from 3 to 2500 per year (median 450) with those of 102 male undergraduates with a median number of times being interviewed of 2. Results indicated that both groups were susceptible to the contrast error. Using two resumes to form a frame of reference for the subject and a third resume to measure the contrast, the experimenters found that the high quality resume seen in the third position received its highest ratings when preceded by two lows as opposed to being preceded by two highs. The reverse was true when the low quality resume was seen in the third position. It received its lowest ratings when it was preceded by two highs as opposed to being preceded by two lows. These results were found for both the experienced and inexperienced raters.

Wexley, et al. (1973) have demonstrated that training can reduce contrast effects. He conducted a series of four studies. Study I involved warning the subject about the occurrence of contrast effects when judging individuals. Results indicated that this warning did not significantly lower contrast errors. Study II involved the use of
written anchors depicting a high suitability applicant for a job and a low suitability applicant for a job. Subjects were shown both these descriptions and were instructed as to the reasons they were rated as high and low. These gave the subjects an anchor of the extreme qualifications reflected by applicants and a guide to rating subsequent applicants. Again, results showed this method to be ineffective in reducing contrast. Study III involved a combination of both methods from Studies I and II; i.e., a warning of the existence of contrast errors and the use of anchor descriptions. Results indicated that this method was also ineffective in reducing contrast. Study IV involved a two hour workshop which utilized discussions of the job description for the position for which the hypothetical candidate was applying, lists of the requirements needed to perform the job successfully, and the way in which these could be used to help evaluate the candidate. Subjects were also shown videotapes of hypothetical applicants depicting high, low, and average suitability for the job and were asked to rate these after which, these ratings were discussed with the subjects. Different types of rating errors were also discussed with the subjects. Results showed that this method was effective in reducing significantly the contrast errors. Contrast was measured by showing the subject three videotapes of an interview situation. The first two tapes were always the same (either high or low quality applicants) and were used to establish a frame of reference for the subject. The third resume was of an average suitability applicant. This was done due to earlier studies (Hakel, et al. 1970 and Wexley, et al. 1972) indicating that the greatest amounts of contrast occur when an average person is used
in the third position than either a low or high quality person. The measure for contrast then was the comparison between the rating given the average candidate in the LLA sequence and the rating given the average quality candidate in the HHA condition. The differences between these two had accounted for 64%, 62%, and 66% of the variance in the ratings in Studies I, II, and III. However, in Study IV the contrast error was reduced to 3% using the method of training selected by Wexley.

The intent of the present study is to investigate the possibility of contrast effects occurring under social conditions other than those already described. Previous studies have used basically written resumes, adjective descriptions, videotaped interviews, recorded conversations, and photographs to measure contrast effects. The present study utilized a fairly common selection technique to investigate the possibility of contrast effects occurring in a selection decision. The instrument used was the Leaderless Group Discussion (LGD). Essentially, this technique involves placing several prospective candidates to be evaluated in a group discussion in which their task is to solve some hypothetical problem. No leader is assigned to guide the progress of the group and participants are given a particular position to defend and background information to use in formulating their arguments. Several different skills are measured in an LGD so it gives an evaluator a chance to see the candidate in an interpersonal situation and how he uses certain skills to his best advantage. This situation has demonstrated high interrater reliability and validity in the business setting (Bass, 1954; Bray and Grant, 1966; and Greenwood and
McNamara, 1967). The three areas of interest involved in the study were contrast effects, differential levels of training of the raters, and sex of the rater. (Although not a primary object of the study, statistical analysis necessitated the control of the sex variable.)

**Method**

**Subjects**

Subjects consisted of 24 male and 24 female undergraduate students in Psychology at Florida Technological University.

**Stimuli**

Two 30 minute videotaped LGD's served as the stimuli for the subjects. Each LGD contained a panel of five males participating in a problem solving session. Each participant was assigned a particular area of the problem to address himself to and defend in his arguments. Each participant was given background information on his area and the areas of the other members of the group. Although there were 5 members in the group, the subjects were asked to rate the performance demonstrated by one of the members who the experimenter designated as the target person. The target person in each film demonstrated different level of performance, either high quality performance (H) or low quality performance (L). High quality performance was characterized by the target displaying behaviors such as; actively participating in the conversation, actively trying to establish himself as a leader, making good suggestions and arguments, readily defending his position but not to the point that he acts as a block to the progress of the group, and a willingness to give a little to achieve as much of his original goal
as possible. Low quality performance was characterized by the target person displaying behaviors such as: little active participation in the discussion, a tendency to follow the lead of others rather than initiate them, suggestions and arguments offered were of poorer quality, and being argumentative to the point that he blocked or hindered the progress of the group.

Procedure

Subjects were randomly assigned to one of four test conditions for viewing of the LGD's. The first condition required the subject to view the high quality LGD (H) and rate the target person's performance. The second condition required the subjects to view the low quality LGD (L) and rate the target person's performance. Conditions three and four required the subjects to watch a two film sequence. In condition three the subjects were asked to watch the low quality LGD and rate the target person. They were then shown the high quality LGD and asked to rate the target person in it (L-H condition). The measure used in this sequence is the rating given on the second film in the sequence. The first film is used to establish a frame of reference for the subject. The fourth condition is the same except that the subject is required to watch and rate the high quality LGD followed by the low quality LGD (H-L condition). Again, the measure of interest is the rating given on the second film in the sequence.

The following comparisons were made to determine the effects of contrast. When comparing the L condition with the H-L condition, contrast would predict that the L rating in the H-L sequence would be lower than the L condition alone due to the influence of the H condition.
preceding it. When comparing the H condition with the L-H condition contrast would predict that the H in the L-H sequence would be seen as higher than the H condition alone. Each of these four conditions were repeated over all levels of the other two variables of training (high and low) and sex of the rater (male and female). The resulting design was a 2 x 4 x 2 factorial design with 16 cells and an n=3 per cell.

Low training. Subjects in the low training condition were given approximately 30 minutes instruction on the purposes of the LGD in assessment setting, the use of the rating forms, brief description of the skills to be measured, and a brief explanation of the problem that the LGD members would be discussing. After this discussion the subjects were immediately placed in the experimental conditions.

High training. Subjects in the high training condition were given approximately 2 hours of instruction on the LGD as an assessment device. Subjects were given background information on the use of LGD's for assessment purposes, instruction on the skills to be measured and the rating forms to be used, and discussion of some common rating errors (halo, central tendency, constant errors, etc.). Subjects were also exposed to two short (15-20 minute) LGD's. The purpose of these were to demonstrate to the subjects some of the positive and negative behaviors often seen in an LGD and how they should be evaluated and to get some practice in the use of the rating forms. After training, subjects were run in the experimental conditions. The time lag between training and participation ranged from 1 to 6 days (mean = 3.8 days). (The resulting design was a 2 x 4 x 2 factorial design with 16 cells and an n=3 per cell.)
Rating forms. Basically two types of rating forms were used. The first form used was a behavioral checklist which listed the positive and negative behaviors associated with a particular skill to be measured. The subject was asked to simply place a check mark beside those behaviors which he thought that the target person displayed during the course of the LGD. The second form used was a 6 point rating scale with anchors at the extremes. The anchors were constructed of behaviors which readily depicted the skill to be measured or rated. The task of the subject was to mark the behavioral checklist during the course of the LGD and to use this as an aid in making his ratings of a particular target individual. These were scales for each individual trait and one for an overall rating. Basically, overall ratings were used in the statistical analysis. (For samples of both behavioral checklists and rating forms, see Appendix A and B.)

Results

Analysis of the data was carried out using a 2 x 4 x 2 fixed effects Analysis of Variance with n=3 per cell (Mincer, p. 246). Table 1 displays cell means and standard deviations for the 2 x 4 x 2 design.

Results of the Analysis of Variance are displayed in Table 2.
### Table 1

Cell Means and Standard Deviations

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th></th>
<th>Females</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>H</td>
<td>L-H</td>
<td>H-L</td>
</tr>
<tr>
<td>Low</td>
<td>2.0</td>
<td>4.66</td>
<td>5.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Training</td>
<td>.77</td>
<td>.53</td>
<td>0.00</td>
<td>1.41</td>
</tr>
<tr>
<td>High</td>
<td>1.66</td>
<td>4.66</td>
<td>4.66</td>
<td>1.66</td>
</tr>
<tr>
<td>Training</td>
<td>.50</td>
<td>.53</td>
<td>.53</td>
<td>.50</td>
</tr>
<tr>
<td>X</td>
<td>1.83</td>
<td>4.66</td>
<td>4.83</td>
<td>1.83</td>
</tr>
</tbody>
</table>
### TABLE 2

Analysis of Variance

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training(A)</td>
<td>.02</td>
<td>1</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td>Quality of Performance(B)</td>
<td>136.89</td>
<td>3</td>
<td>45.63</td>
<td>50.70</td>
</tr>
<tr>
<td>Sex of Rater(C)</td>
<td>.18</td>
<td>1</td>
<td>.18</td>
<td>.61</td>
</tr>
<tr>
<td>AB</td>
<td>.23</td>
<td>3</td>
<td>.07</td>
<td>.07</td>
</tr>
<tr>
<td>AC</td>
<td>.53</td>
<td>1</td>
<td>.53</td>
<td>.58</td>
</tr>
<tr>
<td>BC</td>
<td>3.07</td>
<td>3</td>
<td>1.02</td>
<td>1.13</td>
</tr>
<tr>
<td>ABC</td>
<td>.06</td>
<td>3</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td>Within Cell</td>
<td>29.00</td>
<td>32</td>
<td>.90</td>
<td></td>
</tr>
</tbody>
</table>

* p less than .001
Contrast effects. The only significant differences that occurred in the study were the main effects associated with the levels of performance of the target individuals in the LGD's. Since the measure of contrast would be determined by differences between these conditions, a series of post-hoc comparisons (Scheffe') were carried out. In the absence of interaction effects we can assume that the main effects are equal to the simple main effects such that the investigation of contrast can be carried out across both levels of sex of the rater and levels of training. Table 3 displays the cell and column means for each level of performance across both levels of training and sex of the rater. It can be seen from these tables that column means are equal for both levels of training and sex of the rater. Therefore, comparisons made on one table would be equal to the same comparisons made on the other table. Because of this, comparisons made across levels of training were generalized to levels of sex of the rater.

Table 4 represents the post-hoc comparisons made on the data to locate the differences established by the main effects due to levels of performance. Comparison A contrasts the difference between the L condition and the H-L condition. If contrast predictions held true we would expect the L rating in the H-L condition to be significantly lower than the L condition alone. As can be seen, the difference is not significant and is, in fact, in the opposite direction from what was expected. Comparison B contrasts the difference between the H and L-H conditions. The results indicate that the difference is also insignificant. In general, we can state that contrast effects did not occur across either levels of training or levels of sex of the rater.
### TABLE 3

Mean Ratings

Training X Quality of Performance

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>H</th>
<th>L-H</th>
<th>H-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Training</td>
<td>1.66</td>
<td>5.00</td>
<td>5.00</td>
<td>1.83</td>
</tr>
<tr>
<td>High Training</td>
<td>1.50</td>
<td>5.16</td>
<td>5.00</td>
<td>1.66</td>
</tr>
<tr>
<td>Column X</td>
<td>1.58</td>
<td>5.08</td>
<td>5.00</td>
<td>1.75</td>
</tr>
</tbody>
</table>

Sex of Rater X Quality of Performance

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>H</th>
<th>L-H</th>
<th>H-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>1.83</td>
<td>4.66</td>
<td>4.83</td>
<td>1.83</td>
</tr>
<tr>
<td>Female</td>
<td>1.33</td>
<td>5.50</td>
<td>5.16</td>
<td>1.66</td>
</tr>
<tr>
<td>Column X</td>
<td>1.58</td>
<td>5.08</td>
<td>4.99</td>
<td>1.75</td>
</tr>
</tbody>
</table>
TABLE 4
Comparisons for Contrast Effects
Across Levels of Training

<table>
<thead>
<tr>
<th>Comparisons</th>
<th>L</th>
<th>H</th>
<th>L-H</th>
<th>H-L</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>+1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>.166</td>
<td>.184</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>+1</td>
<td>-1</td>
<td>0</td>
<td>.042</td>
<td>.046</td>
</tr>
<tr>
<td>C</td>
<td>+1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>73.5</td>
<td>81.6*</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>+1</td>
<td>-1</td>
<td>63.4</td>
<td>70.4*</td>
</tr>
</tbody>
</table>

* p less than .05
On the basis of the results we can conclude that seeing a frame of reference, behaving in a prescribed manner, before rating a target individual did not effectively alter the evaluation made on that target person. As can be seen from the results of the post-hoc comparisons, the significant differences that accounted for the main effects associated with levels of quality of performance were those contrived for the study and designed to be significantly different so as to allow contrast to occur. In other words, the significant overall effect was accounted for in the differences between the H and L conditions and the L-H and H-L conditions.

Levels of training. Even though the Analysis of Variance indicated that no significant differences occurred as a result of different levels of training, a further analysis of the behavioral checklist was carried out. The measure used was the total number of positive and negative behaviors checked by each rater.

Even though ratings between different levels of training were essentially the same, differences between number of behaviors checked were compared to see if high training groups saw or took into consideration more behaviors than did low training groups when making their evaluations.

To evaluate the significance of training on the number of behaviors checked by raters, two 2 x 4 (fixed effects) Analysis of Variance (n=6 per cell) were carried out. One Analysis of Variance for positive behaviors checked and one for negative behaviors checked for low and high levels of training. Table 5 displays the mean number of behaviors (both negative and positive) checked by low and high training groups.
### TABLE 5

Mean Number of Positive and Negative Behaviors Checked

<table>
<thead>
<tr>
<th></th>
<th>L</th>
<th>H</th>
<th>L-H</th>
<th>H-L</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low Training</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>+</td>
<td>2.00</td>
<td>20.33</td>
<td>22.60</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>19.33</td>
<td>1.83</td>
<td>.50</td>
</tr>
<tr>
<td><strong>High Training</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>+</td>
<td>2.66</td>
<td>18.33</td>
<td>18.83</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>14.66</td>
<td>2.50</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 6 displays the results of the Analysis of Variance for positive behaviors checked for low and high training conditions. As can be seen there were no significant main effects for levels of training indicating that low and high training groups did not tend to see or utilize different numbers of positive behaviors when making their evaluations. There also appeared a main effect of levels of quality of performance of target persons, however, this would be expected due to contrived differences in the study.

Table 7 displays the results of the Analysis of Variance for negative behaviors checked for low and high training conditions. Results indicate that there is no main effect due to level of training present. However, the presence of the significant interaction effect indicates that differences might exist depending on which level of quality of performance of the target person is being observed. With this in mind, an Analysis of simple effects of levels of training was carried out. The results of which are displayed in Table 8.

As can be seen from the results of the Analysis significant differences between high and low training groups existed in two of the four quality of performance conditions. Results show that when subjects viewed the low quality of performance condition low training subjects checked more negative behaviors than did high training subjects. Significant differences also existed when subjects viewed the H-L condition. The results show that in this condition high training subjects checked more negative behaviors when making their evaluation than did low training subjects.
TABLE 6

Analysis of Variance of Positive Behaviors Checked

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Training(A)</td>
<td>18.75</td>
<td>1</td>
<td>18.75</td>
<td>0.80</td>
</tr>
<tr>
<td>Quality of Performance(B)</td>
<td>3498.4</td>
<td>3</td>
<td>1166.1</td>
<td>49.91</td>
</tr>
<tr>
<td>AB</td>
<td>38.4</td>
<td>3</td>
<td>12.8</td>
<td>0.547</td>
</tr>
<tr>
<td>Error</td>
<td>934.4</td>
<td>40</td>
<td>23.36</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4489.9</td>
<td>47</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p less than .05
TABLE 7
Analysis of Variance of Negative Behaviors Checked

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Training (A)</td>
<td>.2</td>
<td>1</td>
<td>.19</td>
<td>.034</td>
</tr>
<tr>
<td>Quality of Performance (B)</td>
<td>2347.7</td>
<td>3</td>
<td>782.5</td>
<td>142.2</td>
</tr>
<tr>
<td>AB</td>
<td>94.24</td>
<td>3</td>
<td>31.4</td>
<td>5.70</td>
</tr>
<tr>
<td>Error</td>
<td>220.17</td>
<td>40</td>
<td>5.50</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2662.32</td>
<td>47</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p less than .05
### TABLE 8

Analysis of Variance of Simple Effects of Levels of Training for Negative Behaviors Checked

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training for L Condition</td>
<td>65.33</td>
<td>1</td>
<td>65.33</td>
<td>11.87 *</td>
</tr>
<tr>
<td>Training for H Condition</td>
<td>1.33</td>
<td>1</td>
<td>1.33</td>
<td>.24</td>
</tr>
<tr>
<td>Training for L-H Condition</td>
<td>.75</td>
<td>1</td>
<td>.75</td>
<td>.13</td>
</tr>
<tr>
<td>Training for H-L Condition</td>
<td>27.03</td>
<td>1</td>
<td>27.03</td>
<td>4.91 *</td>
</tr>
<tr>
<td>Error</td>
<td>220.17</td>
<td>40</td>
<td>5.50</td>
<td></td>
</tr>
</tbody>
</table>

* p less than .05
The author expected that if differences occurred the high training groups would use significantly more behaviors than the low training groups due to more extensive training. Significant differences in the H-L condition were as expected but those in the L condition were not.

Sex of the rater. Results indicated that female raters did not rate levels of quality of performance any differently than did the male raters.

Discussion and Conclusions

In general the following conclusions can be drawn from the results of this study:

1. Contrast effects did not occur across either levels of training or levels of sex of the rater. Seeing a frame of reference opposite in value to that of the target person did not alter significantly the ratings given that target person.

2. Different levels of training had no differential effect on how a target person was rated across different qualities of performance or levels of sex of the rater. Different levels of training did not affect the amount of positive behaviors associated with a rating made on a particular target person. Level of training had a significant effect on the number of negative behaviors associated with a rating made on two (L and H-L) of the four quality of performance conditions. Only one of the two significant differences was in the expected direction (H-L condition) and as such these results must be interpreted with caution.
3. Male and female raters did not tend to rate target persons differently across levels of training or levels of quality of performance.

Results of this study must be viewed in light of several limitations. First, there was a much greater time lag between training and participation in the task for the high training groups than the low training groups (low training groups were given the task immediately following training whereas the high training groups had a mean time lag of 3.8 days between training and participation in the task). This was not a design of the study and time lag resulted from scheduling voluntary subjects on their availability. Even though subjects were rebriefed upon participation, the time lapse possibly affected their ratings. Secondly, part of the reason that contrast effects did not occur might be due to a possible ceiling-floor effect on the scale used to evaluate the target persons. A 1 (low) to 6 (high) scale was used to evaluate the target persons. A look at the mean ratings given the L and H target persons across both levels of training and sex of the rater (L = 1.58, H = 5.08) indicated that the subjects rated each person at the extreme ends of the scale thus leaving little room for the ratings to vary in the L-H and H-L conditions and contrast to occur. High and low quality performances were used in the study to ensure that differences would be evident and contrast might occur. However, the extreme positions left little room for the contrasts to occur. These results tend to support Hakel's, et al. (1970) and Wexley's, et al. (1972) notion that contrast effects are not as important a phenomena for extreme stimuli as for average stimuli. Hakel, et al. (1970)
using extreme stimuli found that contrast effects accounted for only 1-2% of the total variance even though the differences were significant. Wexley, et al. (1972) found that even though contrast effects were significant for extreme stimuli they accounted for only 2% of the total variance while average stimuli producing contrast accounted for 80% of the total variance. Average stimuli in these cases are less subject to a ceiling-floor effect. This question cannot be fully answered until a system of measuring extreme stimuli that removes the ceiling-floor effect can be developed and the true effect of contrast on extreme stimuli can be measured. The idea, however, that contrast effects might have a greater meaning for average stimuli rather than extreme stimuli makes a great deal of sense when looked at in the light of the selection process. If the applicant population can be considered a normal distribution then the extremes of this population are easily discernable (best and worst applicants). The greatest percentage of the errors made in the selection process would occur in the middle range of applicants where skills possessed would overlap somewhat and differences might be small.

It might also be considered that in previous studies more than one stimuli was used to establish a frame of reference before testing for contrast (Hake!, Ohnesorge, and Dunnette, 1970; Wexley, et al., 1972). These were usually of a very brief nature and this author felt that a 30 minute viewing of a stimuli would be sufficient to establish a frame of reference. A recent study by E. L. Lipscomb (unpublished Master's thesis, 1973) attempted to reduce contrast effects using a sequence of three, three-minute LSD's. The paradigm used was similar to that of
Wexley's, et al. (1972) in that the subjects were shown the first two LGD's to establish a frame of reference (either HH, LL, or AA) and a third LGD to establish contrast (average). The study attempted to reduce contrast through the use of training. The training involved showing the subjects a written description of standards of performance (either high, low, or average) followed by a videotaped LGD showing the behaviors associated with that standard of performance and a discussion of how it should be rated. Results showed that contrast was reduced in the LLA sequence for trained subjects but not the NHA sequence. This along with Wexley's, et al. (1973) study indicate that training can reduce the contrast effect. It should also be pointed out the parallel between the types of training given the subjects in the aforementioned studies and the training in the present study even though Wexley's study was unavailable to the author at that time.

An additional consideration when evaluating the LGD as an instrument in investigating contrast effects is the contrast that occurs between members within the same group as well as the contrast that occurs between LGD's. In the present study and the Lipscomb study subjects were asked to evaluate a target person displaying a prescribed set of behaviors. The other members of the group, however, were not controlled as to their behaviors and were allowed to act freely. The problem that arises is that when a subject evaluates the target person, he does so by comparing him with other members of the group. Expectancies of good and poor behaviors are thus based on all members of the group and not just the target person. This within group contrast might then act to diminish the contrast between groups if members of a subsequent LGD
exhibit different degrees of positive or negative behaviors than the preceding group. The task becomes one of judging groups relative to their collective behaviors rather than judging on one person alone. A better control over all individuals in the LGD and the behaviors they elicit would perhaps lead to a better investigation of the effects of contrast.
APPENDICES
APPENDIX A
Check List.

Instructions: Place a check mark by those behaviors listed that you feel the target person exhibited during the course of the Leaderless Group Discussion.

I. Interpersonal Skills:

Attitude Toward Others:

1. Cooperated well with others in the group.
2. Was polite and congenial in dealing with others.
3. Failed to cooperate well with others.
4. Appeared difficult to deal with during interactions.

Oral Communication Skills:

1. Spoke clearly and distinctly.
2. Appeared to have little trouble expressing himself.
3. Used good eye contact when speaking to others.
4. Used good voice inflection when stressing points of argument.
5. Speech was unclear and indistinct.
6. Seemed to have some trouble expressing himself.
7. Eye contact was poor when speaking to others.
8. Speech tended toward a monotone.

Reaction to Conflict:

1. Took criticism of his views in stride, did not seem overly distressed.
2. Was able to formulate good counter-arguments in reaction to criticism.
3. Seemed somewhat distressed or stressed by criticism of his arguments.
4. Seemed unable to counter criticism with good arguments.

II. Leadership Skills:

Reaction from Others:

1. Was able to persuade others to act upon his suggestions.
2. Tended to be looked upon by others to carry the conversation.
3. Generally was looked upon as a source of suggestion or information.
4. Generally failed in his efforts to persuade others to use his suggestions.
5. Tended to look to others to carry the conversation.
6. Generally looked to others for suggestion or information.
Motivation to Lead:

1. Usually made his presentations or suggestions first.
2. Acted as an organizer of group activity.
3. More or less guided the group so that a solution could be reached.
4. Was usually quiet unless called upon.
5. Tended to go along with decisions of others rather than initiating them.
6. Initiated little or no suggestions to the solution of the problem.

Effectiveness:

1. Effective in keeping the discussion going toward a solution to the problem.
2. Generally presented the opinion that prevailed.
3. Consistently and effectively tried to form coalitions to strengthen his position.
4. Used effective arguments to present his position or opinion.
5. Suggested compromises when necessary so as to reach a solution to the problem.
6. Discussion did not tend to be in favor of the position he favored.
7. His ideas or suggestions did not prevail.
8. Failed in his efforts to form coalitions to strengthen his position.
9. Failed to suggest compromises that would have led to the solution of the problem.
10. Arguments used tended to be ineffective.

Forcefulness:

1. Persistent in his point of view without becoming a barrier to the solution of the problem.
2. Resisted attempts by others to present their point of view as final.
3. Insisted on a final decision before time limit was up.
4. Persistent in his attempts to run the group without becoming a barrier to the solution of the problem.
5. Readily defended his position against criticism.
6. So persistent in his point of view that he acted as a block to the progress of the group in solving the problem.
7. Tended to yield to the arguments of others.
8. Little or no attempts at leading the group.
9. Failed to make effective counter-proposals to other people's suggestions.
10. Not forceful enough in pursuit of his point of view.
11. Argumentative to the point that he blocked the progress of the group.
III. Individual Work Characteristics:

General Activity Level:

1. Appeared alert and energetic.
2. Participated freely in the discussion.
3. Seemed enthusiastic about the task at hand.
4. Participated little in the general activity of the group.
5. Lacked enthusiasm for the task.

Thoroughness:

1. Arguments were good and showed good reasoning.
2. Suggestions were sound; directed at solving the problem as expeditiously as possible.
3. Arguments tended to lack sound reasoning.
4. Suggestions lacked quality; were not expeditious to the solution of the problem.

Adaptability:

1. Recognized the need to be flexible in his requests in order to solve the task.
2. Was willing to alter his request in order to get as much of his original suggestion as possible.
3. Failed to recognize the need for flexibility in order to complete the task.
4. Was not willing to alter his original request; stubborn to the point that he blocked progress of the group.
APPENDIX B
Rating Scales

Instructions: Based on the behaviors that you have checked on the preceding list, circle the area on the following continuaums that you feel best represents the performance of the target person.

I. Interpersonal Skills:

Altitude Toward Others:

Cooperated well with others; not hard to get along with. 6...5...4...3...2...1

Oral Communication Skills:

Communicated well orally. 6...5...4...3...2...1

Reaction to Conflict:

Handled criticism well; did not seem stressed. 6...5...4...3...2...1

II. Leadership Skills:

Reaction from Others:

Was generally successful in influencing others. 6...5...4...3...2...1

Motivation to Lead:

Generally acted as the focal point of the group. 6...5...4...3...2...1

Effectiveness:

Tended to be effective in getting things done. 6...5...4...3...2...1

Forcefulness:

Sufficiently forceful in presenting his point of view. 6...5...4...3...2...1
III. General Work Characteristics:

General Activity Level:

Participated freely in the group. 6...5...4...3...2...1 Failed to participate; remained in the background.

Thoroughness:

Points made were sound and showed good reasoning-6...5...4...3...2...1 Reasoning generally lacking; points made not well taken.

Adaptability:

Recognized the need for flexibility in order to complete the task. 6...5...4...3...2...1 Generally failed to recognize the need for flexibility; acted as a block to the progress of the group.

Overall Ratings:

Based on the ratings that you have just made of different specific skills, rate the target person on the following overall traits.

I. Interpersonal Skills:

   Outstanding 6...5...4...3...2...1 Poor

II. Leadership Skills:

   Outstanding 6...5...4...3...2...1 Poor

III. Individual Work Characteristics:

   Outstanding 6...5...4...3...2...1 Poor

Based on the ratings given on both the specific and overall skills, rate the target person as to his overall potential as a leader.

Overall Rating:

   Outstanding 6...5...4...3...2...1 Poor


