

Winter 1961

Municipal Police Performance Rating

George N. Beck

Follow this and additional works at: <https://scholarlycommons.law.northwestern.edu/jclc>

 Part of the [Criminal Law Commons](#), [Criminology Commons](#), and the [Criminology and Criminal Justice Commons](#)

Recommended Citation

George N. Beck, *Municipal Police Performance Rating*, 51 *J. Crim. L. Criminology & Police Sci.* 567 (1960-1961)

This Criminology is brought to you for free and open access by Northwestern University School of Law Scholarly Commons. It has been accepted for inclusion in *Journal of Criminal Law and Criminology* by an authorized editor of Northwestern University School of Law Scholarly Commons.

MUNICIPAL POLICE PERFORMANCE RATING

GEORGE N. BECK

Lieut. George N. Beck has been a member of the Public Information Division of the Los Angeles Police Department since 1959. He joined the Los Angeles Department in 1950 serving in various capacities including Accident Investigation, Planning and Research, and as an administrative analyst in the Personnel and Training Bureau. Upon promotion to sergeant, he was assigned for a time to a field patrol division and subsequently, as an investigator in the Detective Bureau. Lieut. Beck received a degree in Public Administration, University of Southern California in 1950 and his Master's degree from the same institution in 1958. This paper is a resume of his master thesis.—EDITOR.

One of the constant problems facing police administrators is that of adequately appraising employee performance. This problem is obviously not peculiar to police administrators. Whisler has termed it the "unavoidable responsibility" of management.¹ Coupled with the mandatory nature of this managerial responsibility is the agreement among writers in the field that it involves grave difficulties. Employees are evaluated, whether formally or informally, casually or periodically. The problem is to devise a method whereby those measures will be unprejudiced, objective, and uniform as possible.

The difficulty in devising a method of obtaining objective conclusions from performance ratings lies in the fact that such ratings are essentially a personal audit of one man's "conduct" by another. The attempt to obtain objective data from personnel ratings has been a major problem in the development of the rating process.

Recent trends have been toward recognizing the ineluctable subjectivity of the rating process. Best states, "while ratings may be systemized they cannot be objectified because the rating process is inherently subjective."² Rather than attempt to compensate for the subjective nature of ratings, observers in the field have begun to recognize that the main justification for ratings resides in this quality that does not lend itself to testing.

There follows a brief summary of the scope of this study, the means used in collecting the data presented, and a review of some findings of the study insofar as the extent to which ratings are used, the kinds of ratings used, and the purposes for which ratings are made.

¹ THOMAS L. WHISLER, "A Realistic Role for Merit Rating," *THE JOURNAL OF BUSINESS*, 28:39, January, 1955.

² WALLACE H. BEST, "Some New Directions in Personnel Appraisal," *PERSONNEL*, 37:46, September-October, 1957.

COLLECTION OF DATA

Information was collected by means of a survey questionnaire directed to municipal police departments in cities having in excess of 50,000 population as of the 1950 census. Police agencies were requested to furnish information regarding the means used in evaluating the performance of their employees. Departments using rating forms were asked for the types of ratings used and the objectives of their rating program. They were also asked for detailed information regarding the frequency of ratings, by whom ratings were made, whether ratings were subject to review and/or appeal, and what changes were anticipated in their rating forms.

Additional information was collected, recorded, and evaluated from an analysis of the rating forms furnished by reporting departments. From these forms data were gathered on the kinds of forms used, the nature of traits rated, and the means used to summarize ratings.

EXTENT OF RATING USAGE AND PURPOSE OF RATINGS

The extent of rating usage. One hundred of the 232 cities queried indicated the use of performance rating forms. It may be tentatively concluded that at least 43 per cent of all police departments in cities of over 50,000 population use performance ratings. The 178 cities responding represented 76.7 per cent of the cities contacted.

It was determined that a direct correlation existed between the size of the police departments contacted and their inclination to use employee ratings. Most of the largest departments reported the use of ratings, but the smaller departments were not so likely to express a recognition of the need for a formal rating program.

The responses of all surveyed departments were

plotted geographically in order to ascertain whether there was any relation between geographical location of police departments and their use or non-use of performance rating forms. It was found that departments in the Northeast, Mid-Atlantic, and Midwestern regions were less prone to use ratings than were departments in the South, Southwest, Rocky Mountain, and Pacific Coastal regions. The causes for this variance in the use of rating forms were not apparent in this study. It was hypothesized that further study might reveal a predisposition on the part of departments in cities having progressive administrative policies to utilize a formal system for the periodic collecting and recording of supervisory evaluations of employee performance.

Information received from police departments not using performance rating forms indicated that there is no alternate method of employee evaluation in widespread usage other than the reliance upon the memory of the supervisor as reflected in oral or written replies made upon the conduct of subordinates. Of the seventy-eight departments reporting non-use of rating forms, only thirty-seven stated that other means were used to evaluate employee performance. All of these indicated the use of oral and/or written reports from supervisors as the sole means of employee evaluation.

Comments from departments that have either abandoned rating forms as a means of employee evaluation or who have never tried them express a general reluctance to rely upon the judgment of supervisors as contained in the confined media of rating forms. It has been said, however, that properly conceived performance rating is a procedure in which the rating form plays a relatively minor part in its success or failure. It is not the rating form itself but failure to take into account other important considerations that leads to the failure of, or dissatisfaction with, the ordinary performance rating procedure.³

Purpose of ratings. Police departments contacted in this study were asked to indicate the uses made of their ratings. From responses by ninety-seven departments it appears that the great majority of police departments use their ratings for both personnel transactions: e.g., placement, promotion, and in the evaluation of training techniques, and as a means of improving employee performance. This practice has resulted in compromises made

by raters and in the content of rating forms which have seriously hampered the effectiveness of rating programs. The difficulties encountered in attempting to devise one common form for the two divergent purposes are one of the primary reasons for the dissatisfaction with ratings which is prevalent at the operating level in police departments throughout the United States. The solution recommended by authorities in the field is to confine the purpose of employee ratings to the improvement of employee performance.⁴

KINDS OF RATINGS AND TRAIT SCALE CHARACTERISTICS

Kinds of ratings in use. Performance evaluation by formal ratings are almost exclusively effected by means of either check-lists or rating scales. A few police departments have, in the past, used rank order ratings, paired comparisons, and forced distribution scales, but the trend has been toward the graphic rating scale. In this study 88 per cent of the forms received were within the broad category of graphic ratings. Of the remaining 12 per cent all but one were check-lists, the exception being the field review type rating used by the Police Department of Glendale, California. The evidence available indicates a trend away from the check-list type of rating. A survey made in 1932 by the United States Conference of Mayors showed 46 per cent of the departments using ratings were using a check-list. An unpublished survey made in 1952 by Lieutenant Sinclair of the Los Angeles Police Department indicated that departments using check-lists had dropped to 35 per cent of the number reporting the use of rating forms. The present survey shows less than 12 per cent of all departments using ratings are now using the check-list type of rating.

Police departments now exhibit a growing tendency to rely on the graphic rating scale technique in the evaluation of employee performance. It should be noted that there are a number of variations of the graphic rating scale. Traditionally, a graphic scale has been thought of as one on which traits are rated by placing a check mark at the approximate place on a line indicating the various degrees of possession of, or proficiency in, the trait. One variation substitutes categories identifying the rating of a trait by means of alphabetical or numerical symbols in place of the usual linear scale. Another variation presents the trait

³ E. A. RUNDQUIST AND REIGN H. BITNER, *RATING EMPLOYEE AND SUPERVISORY PERFORMANCE*, ed. M. J. Dooher and V. Marquis. (New York: American Management Association, 1950), pp. 69-71.

⁴ Details are set forth in the author's master thesis which is available at the University of Southern California in book form.

range in the form of adjective options for the rater to check. All of these variations are presently in use by police departments.

Much of the distrust and disappointment with rating scales has resulted from the traditional attempt to secure unnecessary differentiation among individuals on the part of the rater. Scales employing numerous subdivisions or steps encourage and promote the tendency to rate employees against each other. One means that has been suggested to combat this tendency is to confine the rating steps in employee evaluation ratings to three alternatives. Thus, rating steps would be reduced to some expression of unsatisfactory, satisfactory, or superior. This technique focuses the intent of the rater on the evaluation of the ratee in terms of job performance.

The effect of reducing the rating steps available on a trait scale to three is to compress all ratings formerly assigned to categories other than the two extremes into one classification representing acceptable performance or possession of the trait in question. The reasoning behind the three category scale is that the differences within the range representing satisfactory performance are not worthy of the great amount of trouble the distinctions cause. On the other hand, it is worth the effort to distinguish between the inadequate and the definitely superior. There appears to be one problem worthy of concern in connection with the three-step trait scale. Errors of central tendency would seem to be subject to increase where the rater must either rate inferior, satisfactory, or superior.

Verbal anchoring of trait scale steps. It has been demonstrated in experimental study that both the reliability and the amount of information transmitted by a performance rating increase with the identification of steps in a rating scale by means of descriptive phrases.⁵ Of the seventy-four departments reporting the use of graphic type rating forms, twenty-seven use forms which have no verbal anchoring of the trait scale steps or definitions of the traits to be rated. It is quite possible that these departments provide definitions of the rating traits for their raters in a location apart from the rating form, or in some other manner not known to the writer.

Rating traits should be introduced with a definition following the trait name. The trait name is

⁵ A. W. BENDIG AND S. B. HUGHES, "Effect of Amount of Verbal Anchoring and Number of Rating-Scale Categories Upon Transmitted Information," *JOURNAL OF EXPERIMENTAL PSYCHOLOGY*, 46:67, August, 1953.

only a label unless it is identified by verbal cues which further distinguish trait steps. In the absence of these sign-posts for the rater, results are very likely to be unreliable and misleading. Trait definitions should be formed in language compatible with the job; therefore, operational terms should be used as much as possible. Guilford states that rating-scale cues should have the double purpose of supplementing and reinforcing the definition and of providing "anchors or mileposts" to guide the rater in making quantitative judgments. He lists six requirements for rating-scale trait cues: clarity, relevance, precision, variety, objectivity, and uniqueness.⁶

Number of rating-scale traits. There is great variation among reporting police departments in regard to the number of traits included on their rating scales. The minimum number of traits reported was three, and a maximum of thirty-five rating traits was reported, with a median of 11.05 and a mean of 11.6.

The number of rating traits used by police departments consulted in this survey is in approximate agreement with the recommendations of some authorities. One recommendation from within the law enforcement field is to the effect that rating 200 factors should be limited to from four to ten on the grounds that less tend to be meaningless and more tend to confuse the rater.⁷ Another writer observes that there is an almost universal tendency to use an excessive number of traits and recommends that seven to ten is enough.⁸ However, it has been argued that simplicity or brevity in a rating form is not necessarily a virtue. If ratings are to be reliable, they should contain a sufficient number of traits to give an adequate profile of the employee's abilities.

An analogy has been drawn between a rating scale and an objective test in relation to the accepted principle that increasing the number of test items will increase the reliability of the test. Similarly, increasing the number of pertinent factors in a rating scale will secure a wider sampling of work factors and should increase both its reliability and validity. The key to the problem of how many traits to include in a rating form lies in the concept expressed by the word "pertinent"

⁶ J. P. GUILFORD, *PSYCHOMETRIC METHODS* (Second Edition, New York: McGraw-Hill Book Company, 1954), pp. 292-94.

⁷ MUNICIPAL POLICE ADMINISTRATION (Chicago: International City Managers' Association, 1954), p. 166.

⁸ WALTER MANLER, "Some Common Errors in Employee Merit Rating Practices," *PERSONNEL JOURNAL*, 26:68, May, 1947.

in the above reference. No arbitrary statement can be made concerning the ideal number of traits to be rated. Certainly only those traits essential to good police work, but not amenable to more objective evaluation, should be considered as proper material for rating form factors.

The nature of rating-scale traits. In considering what factors to include in a rating scale there are certain considerations and criteria that should be met. First, it is important to choose traits that are either demonstrated indexes of characteristics related to success or failure in the performance of the particular job to be done or which are significant in themselves.⁹ A decision must also be made as to whether the ratee should be rated as a whole or whether specific traits of behavior should be evaluated. Those in favor of the first alternative will say that it is impossible for supervisors to segment behavior for rating purposes, that the whole man is always judged. This tendency is a well recognized one but does not negate the fact that both the objectivity and reliability of reports increase as the factors are fairly presented. In organizations conditioned by appropriate training, the ratings obtained will be more valuable than overall judgments.

Another decision must be made in respect to the nature of traits to be evaluated. Should direct ratings be made of personality factors, such as integrity, tact, and initiative; or should the emphasis be on specific job behavior or activities, such as the extent to which work is done without immediate direction, the display of ingenuity in particular situations, and the ability to work with fellow employees without irritation?

It has been found that the rating traits most accurately judged are those for which there is objective evidence, those which are simple, and those which are carefully defined. There is less agreement about what makes one trait objective and another trait subjective. Perhaps the best conclusion is that the traits which are most reliably judged are those "which leave their mark on things or influence external events."¹⁰

Prior to developing a tentative list of rating traits, it is necessary to complete an analysis of the jobs to be covered by the rating procedure. It is imperative to know what is required of police officers before measurements can be made to see

whether they meet the requirements. Since the duties of police officers are numerous, it is quite laborious to develop a separate form for each duty assignment. It is therefore necessary to select for the general rating scale the important requirements that are common to many assignments. Bittner states that traits should be selected on the basis of the criteria of observability, universality, and distinguishability.¹¹

RATING TRAITS AND RESULTS OF RATINGS

Rating traits used by police departments. While a total of sixty-six different trait names were counted on the rating forms submitted by seventy-four departments, there was considerable overlapping present, both intra-departmentally and inter-departmentally. There were nine rating traits which appeared on half or more of all seventy-four rating forms. These were: initiative, quality of work, judgment, quantity of work, appearance and neatness, cooperation and tact, attitude towards work, knowledge of duties, and inter-personal relations.

A majority of the rating traits currently used by police departments violate the basic principles discussed previously in this study regarding the criteria for trait selection. Several items which have been rated frequently could have been evaluated more accurately from other sources; e.g., intelligence, punctuality, physical fitness, quantity of work, firearms ability, and safety mindedness. Several other traits fall into the nebulous and intangible category; e.g., initiative, loyalty, morale, integrity, force, and resourcefulness. These factors should not be rated, because either the raters do not have enough information to justify an opinion on these subjects or the factor itself does not permit recognition of individual differences in employees.

How results of ratings are summarized. There is a strong tendency on the part of organizations using performance rating forms to assess an overall or summary rating to a rating form. Police departments are no exception to this practice. Sixty-three departments utilized rating forms which required the rater to gather his judgments regarding the ratings on individual traits into one summary expression.

It has been said that the problem of the measuring instrument is important largely to the extent

⁹ O. GLENN STAHL, *PUBLIC PERSONNEL ADMINISTRATION* (fourth edition; New York: Harper and Brothers Company, 1956), pp. 321-22.

¹⁰ EDWIN E. GHIESELLI AND CLARENCE W. BROWN, *PERSONNEL AND INDUSTRIAL PSYCHOLOGY*, (New York: McGraw-Hill Book Company, 1948), p. 118.

¹¹ REIGN BITTNER, "Developing an Employee Merit Rating Procedure," *RATING EMPLOYEE AND SUPERVISORY PERFORMANCE*, ed. M. Joseph Doohar and V. Marguis (New York: American Management Association, 1954), p. 26.

that one insists upon comparing one person's performance with that of another. By evaluating against standards of performance only, the problem is reduced to the measurement of individual factors or job duties.¹² The elimination of summary or overall ratings makes it feasible to refrain from putting the conclusions on individual factors in terms permitting comparison with the performance of others. Thus, conclusions can be stated in factual terms for use in guiding the particular employee in making the most of his abilities. The major contributing factor of the resistance to ratings by employees has been the emphasis on comparison between employees brought to focus, usually through summary ratings which categorize employee performance into some kind of rank order.

Another injustice wrought by summary ratings is that when numerical or alphabetical scores are totaled to give a concrete score, it is very likely to be grossly unfair and inaccurate because subjective evaluations do not lend themselves to accurate mathematical reflections. Also, two identical total scores may be arrived at by completely divergent rating scores in each of the component rating traits. Again, all jobs or duties, if done well, would not necessarily produce similar totals where some jobs are rated on a smaller number of traits than others.

INHERENT PROBLEMS IN PERFORMANCE RATING PROGRAMS

Various obstacles stand in the way of effective police performance rating programs. The preliminary study made by the author as his thesis for a graduate degree (see footnote 4), points up the need for further research with respect to performance ratings. Four areas, in particular, stand out as in need of future research. The four topics may be considered as aspects of the environment in which ratings are made. For that reason, prior to discussing them, it is pertinent to consider standards which may be regarded as prerequisites to the installation of an employee rating program within a police department.

Standards

The following list of standards required of an organization prior to the installation of an employee evaluation plan was offered by Overholt, who concludes that only when the management of an organization can say that these five principles are operative can they expect an efficiency rating

plan to have a reasonable chance of success. The five principles are:

1. Employees must be fully informed of their duties and responsibilities and of the purpose of every task and must be trained in their duties.
2. Employees must be informed specifically of the quantity and quality requirements of their jobs and what constitutes minimum acceptable standards.
3. Employees must be told promptly when they are failing to perform acceptably and must be told when they are doing commendable work.
4. Employees must be rated on performance as related to standards in effect at the time the work was performed.
5. Employees have the right to protest ratings before a fair and impartial board.¹³

A further consideration vital to the success of ratings is wholehearted support and endorsement of the program by management. In the case of police departments, this means the upper level of administrators, including the chief. The chief of police must, in turn, have the support of his superiors.

Confusion as to Purposes of Ratings

Perhaps the greatest single obstacle to effective performance evaluation lies in the fact that there is much confusion of purposes for which ratings are made. All stated purposes reported by police departments contacted in this study fall into two widely divergent classifications: (1) to provide a basis for administrative actions, such as promotion, pay raises, assignment, and termination; and (2) to help supervisors in their job of striving for optimum employee performance. The current practice of employing a common form and rating standards for both purposes at once tends to produce poor results for each.

There is an urgent need for additional research to determine a feasible plan for police administrators to obtain separate supervisory judgments for use in both areas. A possible solution is to utilize different forms for each of the above listed tasks. Even if management agreed to the added costs implied, this policy would probably encounter crippling resistance from first line supervisors because of the increase it would bring in paper work and time away from supervisory duties.

¹³ JOHN OVERHOLT, "Appraising Employee Performance," PUBLIC PERSONNEL REVIEW, 10:19, January, 1948.

¹² STAHL, op. cit., p. 323.

A more practical program would be to design one rating for use as a basis for all administrative action and another for use in employee counseling and guidance. If alternate ratings were made semi-annually for these two purposes, there would be no increase in rating time for the majority of departments since six months ratings are the most favored. It is true that this policy would result in only one rating of each type annually, but it is possible that annual ratings would suffice if enough controls were present to ensure ratings from supervisors in the event of assignment changes by rater or ratee. Such a rating program would do much to overcome present weaknesses caused by the confusion of rating purposes.

Many writers feel that the best method of eliminating the problems brought about by multipurpose ratings is to de-emphasize personnel purposes. One writer states:

We are moving from the historical concept of a multipurpose formal rating system for personnel actions to the more dynamic concept of evaluation as an integral part of productive work relationships.¹⁴

This philosophy does not, however, obviate the necessity, in personnel actions, for a formal means of securing the aid of supervisory judgments in evaluating aspects of employee performance not susceptible of more objective measurement.

Resistance to Ratings

Resistance on the part of both raters and ratees to employee ratings appears to be quite prevalent in police departments. It may not find overt expression among raters of some departments, but the results are scarcely less devastating to the rating program. There is a need for research to determine practical means of removing the causes for resistance to ratings. There are many causes for this resistance; the following list contains those most frequently discussed by writers:

1. Confusion of purposes.
2. Confusion about what is to be measured.
3. Raters are not consulted or allowed to participate in their development.
4. Rating procedures frequently do not help the rater achieve the basic purpose of the scheme.
5. Inadequate training of raters.

¹⁴ JACK POCKRASS, "Performance Evaluation—Forms or Substance?" *PERSONNEL ADMINISTRATION*, 17:5, September, 1954.

6. Lack of management understanding of the operating problems entailed.

7. Raters are not convinced of the value of ratings.¹⁵

The above shortcomings lead to an added cause for resistance to ratings, especially on the part of ratees. Because of these inadequacies, there is a lack of communication between rater and ratee which further jeopardizes the rating program by creating ratee resistance. The causes of resistance to ratings will not be erased, even though they are identifiable, until operational research demonstrates techniques that will overcome resistance.

Barriers in communication have a tendency to produce low ratings. This attacks the validity of ratings since it becomes problematical whether low ratings are a result of deficiencies in performance or inadequate communication between rater and ratee.¹⁶

The recent trend in employee evaluation is away from elaborate schemes attempting to categorize and compare employee performance toward simpler performance reports with emphasis on improvement of employee performance rather than administrative actions. Probably the most outstanding factor responsible for this change, which has not as yet been very evident in police ratings, is that it has not been possible to secure adequate employee understanding, participation, or acceptance under traditional rating programs. The inability of organizations to secure the active support of employees in rating programs has been said to be the most common source of failure of rating programs.¹⁷ The need for action research to determine more acceptable rating methods is evident.

In view of the extent of resistance to ratings present among police organizations, and of the serious consequences upon the rating program, the possibilities of utilizing the democratic principles present in mutual ratings should be given thoughtful consideration.

It may be argued that the quasi-military nature of police departments precludes the use of the democratic processes of mutual ratings. In respect to this criticism, two points must be considered.

¹⁵ RUNDQUIST AND BITTNER, *op. cit.*, pp. 69-70.

¹⁶ AARON J. SPECTOR, "Influences on Merit Rating," *JOURNAL OF APPLIED PSYCHOLOGY*, 38:393, December, 1954.

¹⁷ WM. E. MOSHER, J. DONALD KINGSLEY, AND O. GLENN STAHL, *PUBLIC PERSONNEL ADMINISTRATION*, (New York: Harper and Brothers Publishers, 1950), p. 365.

First, American military organizations have been, and still are, among the foremost contributors to achieve research and experimentation in peer ratings. Secondly, it is questionable whether police organizations are necessarily more authoritarian in nature than other organizations that have increased efficiency through the utilization of democratic principles of organization and personnel policies. Certainly, the claims made for mutual rating are deserving of a complete examination and testing in the police field. Mutual rating affords unusual opportunity for experimentation concerning the effects participation of ratees in the rating process has upon resistance to ratings.

Supporters of mutual rating do not regard it as a panacea for all employee evaluation problems. Like other rating systems, it produces judgments rather than measurements. What is claimed for mutual rating is that it does apply the principles of group dynamics and provides ego-recognition for individuals. Unlike other rating systems, ratings are thus tailored to the needs of ratees as they see them. The essence of mutual rating is that ratings are gathered from each member of an organizational unit to form a profile of each contributing member, regardless of rank. The fact that the adoption of the mutual rating process and the composition of the rating form are decided upon by group action makes it a democratic program as opposed to the authoritarian nature of conventional rating systems.¹⁸ The impact of these principles upon the rating environment might well produce results of such a magnitude in improved morale and performance that employee ratings would be of great value to police departments.

Rater Training

Rater training in police departments should, in the opinion of the investigator, be one of the factors in in-service training programs, with a formal course for new supervisors and periodic refresher courses for all supervisors. The training should include lectures covering the objectives of the rating program, the need for ratings, explanation of the traits in the rating form, and recommendations concerning techniques and rating procedures. The importance of the rating to the ratee should be stressed. In order to be effective,

¹⁸ WALLACE H. BEST, "Mutual Rating Research Project (MRRP)," (unpublished mimeographed report, School of Commerce, University of Southern California, 1957), p. A-1.

raters must be led to believe in the basic purpose of the ratings. Special attention should be given to the importance of follow-up rating interviews.

In addition to formal training, raters should be provided with a pamphlet covering the subject of ratings, with particular emphasis on the purpose of ratings and detailed explanations of the meanings of rating traits. Raters should be encouraged to review the rating pamphlet prior to each rating session.

Rater training among police departments is at present limited to the preparation of rating instruction material in printed form and this is found in only a few cities.¹⁹ There are some signs that more extensive rater training is contemplated by police departments. Departments in three cities—Los Angeles; Peoria, Illinois; and Greensboro, North Carolina—are presently planning rater training programs for the near future.

The focus of employee rating by police agencies has been on the evaluation of individual performance. Communication to the person involved, while practiced in varying degrees by ninety-eight of one hundred departments reporting the use of ratings in the present survey, has received only secondary attention. It has been argued that for improvement in performance to occur as a result of ratings, the rated employee must have a knowledge of the results of his rating. Approval by implication is not sufficient. Failure to train the raters in communication techniques will also adversely affect the success of the rating program. Some problems arise requiring interim attention, and thus supervisors need communications skills daily. These skills are adaptable to a wide variety of supervisory activities involving interpersonal contact.

Design and Content of Rating Forms

In addition to the three areas cited as presenting problems there remains a fourth problem area in employee rating as practiced by police departments. This problem area is the design and content of rating forms.

While it appears that the techniques of rating are not as important as either rater skill or the

¹⁹ Eleven of the onehundred departments reporting the use of ratings in this study are known to supply raters with rater instruction booklets; others may utilize such rater guides. Those departments reporting the use of written rating instructions are: Chicago; Philadelphia; Detroit; New Orleans; Minneapolis; Kansas City, Mo.; Phoenix, Arizona; Glendale, California; Quincy, Mass.; and Oak Park, Ill.

rating environment, still the rating program is handicapped when the rating instrument is so imperfect as to prevent raters from presenting a fair and effective reflection of their judgments regarding the strengths and weaknesses of an employee's performance.

The analysis of rating forms currently utilized by police departments as presented in this study indicated many faults, both in form and content. The most serious design faults are: (1) the lack of adequate description of performance standards and anchoring phrases for rating guides; (2) the lack of an open end device or space for additional comments; and (3) the practice of weighting traits and/or providing a summary score for the rating which forces the rater to convert his judgments into an arithmetical or alphabetical sum.

The most serious faults present in current rating form contents are as follows:

1. Traits are not described in definite, simple terms.
2. Traits do not refer to a single activity.
3. Many traits are included which could be rated more objectively from other sources, such as intelligence, punctuality, physical fitness, etc.
4. Many traits are not observable in work performance.
5. There is a tendency to rate an excessive number of traits.

The errors present in current rating form design and content indicate a need for study to determine what is required in a rating form to assist the rater in his task. The fact that at least sixty-six different traits are rated now among police departments indicates a need for research to determine what traits are appropriate material for evaluation by employee ratings.

CONCLUSION

It is the belief of this investigator that the four problems discussed above will be resolved when enough attention has been focused upon them. If employee ratings are to serve as an effective means of promoting efficiency and positive personnel administration in police departments, these problems must be solved.

The primary inherent problems in police performance ratings point up the need for further research. Four areas for research are set forth hereunder followed in each instance with a summary as to why such research is needed:

1. *Rating purposes.* There is a need for the clarification of the purposes for which ratings are made. The present confusion of purposes constitutes a major stumbling block in the progress of employee ratings. There is strong evidence that the purpose of rating should be primarily one of employee counseling and guidance to effect the improvement of employee performance.

2. *Resistance to ratings.* There exists a strong resistance on the part of both raters and ratees to performance ratings. Optimum effectiveness of rating programs cannot be realized until means are found to alleviate this resistance.

3. *Rater training.* The skill of raters in evaluating employee performance and in utilizing such evaluations is a major factor in effective rating programs. At the present time, little effort is made to train raters. There is a need for research to aid in the development of practical techniques of rater training and in the development of formal courses for in-service rater training.

4. *Design and content of rating forms.* Rating forms now in use leave much to be desired as efficient tools for the reflection of rater judgments. Research can produce valuable information for police administrators now seeking better rating tools.

The intent of this study was to provide a useful beginning in research relative to the problem of employee performance ratings in municipal agencies. Its values are dependent upon the use made of the information concerning the current status of employee ratings among police departments. Operational conclusions made in this study were warranted only in a highly tentative sense and were made primarily as a means of identifying and focusing attention upon problems present in the police employee evaluation field.