INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

76-27,511

KING, Donald Murray, 1939-
A STUDY OF FORMAL PERFORMANCE APPRAISAL BY SUPERVISORS IN AN INDUSTRIAL ENVIRONMENT.

The University of Michigan, Ph.D., 1976
Education, business

Xerox University Microfilms, Ann Arbor, Michigan 48106

A STUDY OF FORMAL PERFORMANCE APPRAISAL BY SUPERVISORS IN AN INDUSTRIAL ENVIRONMENT

by
Donald Murray King

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Education)
in The University of Michigan
1976

Doctoral Committee:

Professor Curtis Van Voorhees, Chairman
Professor Frederick W. Bertolaet
Associate Professor Howard V. Brabson
Assistant Professor William S. Dunifon

DEDICATION

To Marilyn, Dayn
and the Eagle

ACKNOWLEDGEMENTS

To Dr. Curtis Van Voorhees, doctoral committee chairman, my appreciation for his encouragement and assistance. To the members of my doctoral committee, Frederick W. Bertolaet, William S. Dunifon, and Howard V. Brabson, for their time and effort.

To my parents, for giving me the opportunity to make this possible; and to John Murray, a special thank you for his assistance.

To Dayn, for his understanding of a difficult task.

To Marilyn, for her love, inspiration, support and patience in making this a reality.

TABLE OF CONTENTS

Page

DEDICATION. ii

ACKNOWLEDGEMENTS. iii

LIST OF CHARTS vi

LIST OF FIGURES vii

LIST OF TABLES viii

LIST OF APPENDICES. xiii

CHAPTER Page

I INTRODUCTION TO THE STUDY 1

Purpose 1
The Problem 1
Background Information 3
General 3
Foreman Selection System 4
Objectives 7
Organization. 8
Timing 9
Job Analysis 9
Predictor Development 20
Summary 25
Need for the Study 26
Hypothesis 28
Methodology 30
Sample 30
Instrumentation 30
Statistical Analysis 36
Definition of Terms 36
Limitations 44
Summary 45

II REVIEW OF LITERATURE 47

Introduction 47
Historical Overview 49
Philosophy 55

TABLE OF CONTENTS (CONTINUED)

CHAPTER		Page
II	Need	65
	Objectives	68
	Planning, Implementation and Utilization	70
	Rating Systems	92
	Appraisal Techniques	98
	Rating Scales	99
	Graphic Rating Scale Techniques	99
	Forced-Choice Rating Techniques	102
	Behavioral-Anchored Rating Technique	106
	Essay Appraisal Technique	107
	Critical-Incident Technique	110
	Ranking Methods	111
	Alternation Ranking Technique	111
	Paired Comparison Ranking Technique	112
	Group Assessment Technique	113
	Other Techniques	120
	Summary	120
III	ANALYSIS OF THE DATA	124
	Introduction	124
	Demographic Information	125
	Sample Population	125
	Instrumentation	136
	Methodology	137
	Responses	139
	Hard-Core Criteria	153
	Findings	157
	Hypotheses	157
	Summary	188
IV	SUMMARY, FINDINGS, IMPLICATIONS AND RECOMMENDATIONS	190
	Summary	190
	Findings	191
	Rater Comparisons	191
	Rater Reliability	192
	Implications	193
	Recommendations	201
APPENDICES		203
BIBLIOGRAPHY		245

LIST OF CHARTS

CHART		Page
"A"	ORGANIZATION CHART	5
"B"	ORGANIZATION CHART	6
"C"	HIGH/LOW CORRELATION SUMMARY	181
"D"	RANK ORDER OF CRITERIA, RATERS, AND TEST AREAS	182

LIST OF FIGURES

FIGURE		Page
I	LEVELS OF MANAGEMENT (LINE)	37
II	FLOW CHART OF PARTICIPATIVE-ORIENTED APPRAISAL PROGRAM	63

LIST OF TABLES

TABLE		Page
1.1	FOREMAN SELECTION SYSTEM REPRESENTATIVE TIME PLAN	10
1.2	JOB CATEGORY IMPORTANCE RANKINGS BY DIVISION	16
1.3	CLUSTER OF JOB ANALYSIS TASK CATEGORIES	17
1.4	ALTERNATIVE SELECTION PROCEDURES	22
1.5	NUMBER AND PERCENTAGE OF SAMPLE POPULATION BY LOCATION, SEX AND RACE . . .	31
1.6	PERFORMANCE DATA AVAILABILITY BY PERFORMANCE INFORMATION BY LOCATION. . . .	34
2.1	ORGANIZATION CROSS-REFERENCE MANAGERIAL LEVEL	50
2.2	HOW 141 COMPANIES APPRAISE MANAGERS . . .	64
2.3	PREVALENCE OF APPRAISAL PROGRAMS AT VARIOUS ORGANIZATIONAL LEVELS	75
2.4	SURVEY OF RESULTS: USE OF RATING IN BUSINESS FIRMS	76
2.5	FREQUENCY OF USE OF VARIOUS SOURCES OF INFORMATION IN THE EVALUATION OF TEACHING EFFECTIVENESS	82
2.6	PURPOSES FOR WHICH APPRAISAL IS USED . . .	87
2.7	GRAPHIC RATING SCALE	103
2.8	FORCED-CHOICE RATING SCALE	105
2.9	BEHAVIORAL-ANCHORED RATING SCALE	108

LIST OF TABLES (CONTINUED)

TABLE		Page
2.10	ESSAY APPRAISAL	109
2.11	PAIRED-COMPARISON RANKING	114
2.12	ASSESSMENT CENTER DIMENSIONS ASSESSED . .	116
3.1	NUMBER AND PERCENTAGE OF FOREMEN AVAILABLE FOR SAMPLE ON THE VARIABLES OF RACE AND SEX	126
3.2	NUMBER AND PERCENTAGE OF FOREMAN SAMPLE ON THE VARIABLES OF RACE AND SEX .	127
3.3	NUMBER AND PERCENTAGE OF FOREMAN SAMPLE ON THE VARIABLES OF LOCATION, RANKING, SEX AND RACE	130
3.4	NUMBER AND PERCENTAGE OF FOREMAN SAMPLE BY LOCATION AND SHIFT	131
3.5	NUMBER AND PERCENTAGE OF FOREMAN SAMPLE BY LOCATION AND DEPARTMENT	133
3.6	NUMBER AND PERCENTAGE OF FOREMAN SAMPLE ON THE VARIABLES OF TIME AS FOREMAN AND RACE	134
3.7	NUMBER AND PERCENTAGE OF FOREMAN SAMPLE ON THE VARIABLES OF TIME IN POSITION AND RACE	135
3.8	NUMBER AND PERCENTAGE OF RESPONSES ON TASK AND BEHAVIOR CATEGORY CRITERIA BY LOCATION	140
3.9	DESCRIPTION OF TERMS USED IN THE STUDY BY CODE AND CATEGORY	141
3.10	NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY DIRECT SUPERVISION (G1) ON TASK-ORIENTED CRITERIA - INITIAL APPRAISAL	142
3.11	NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY DIRECT SUPERVISION (G2) ON TASK-ORIENTED CRITERIA - FINAL APPRAISAL	143

LIST OF TABLES (CONTINUED)

TABLE		Page
3.12	NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY INDIRECT SUPERVISION (S1) ON TASK-ORIENTED CRITERIA INITIAL APPRAISAL	144
3.13	NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY INDIRECT SUPERVISION (S2) ON TASK-ORIENTED CRITERIA - FINAL APPRAISAL	145
3.14	NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY NON-RELATED SUPERVISION (I) ON TASK-ORIENTED CRITERIA	146
3.15	NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY DIRECT SUPERVISION (G1) ON BEHAVIOR ORIENTED CRITERIA - INITIAL APPRAISAL	147
3.16	NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY DIRECT SUPERVISION (G2) ON BEHAVIORAL ORIENTED CRITERIA - FINAL APPRAISAL	148
3.17	NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY INDIRECT SUPERVISION (S1) BEHAVIOR ORIENTED CRITERIA - INITIAL APPRAISAL	149
3.18	NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY INDIRECT SUPERVISION (S2) ON BEHAVIOR ORIENTED CRITERIA - FINAL APPRAISAL	150
3.19	NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY NON-RELATED SUPERVISION (I) ON BEHAVIOR ORIENTED CRITERIA	151
3.20	NUMBER AND PERCENTAGE OF RESPONSES ON "HARD-CORE" CRITERIA BY LOCATION . . .	154

LIST OF TABLES (CONTINUED)

TABLE		Page
3.21	SAMPLE, MEAN AND WEIGHTED MEAN ON HARD-CORE CRITERIA BY WEEK	155
3.22	JOB-RELATED CATEGORIES BY CRITERIA AND TABLE NUMBER	159
3.23	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE TASK CATEGORY OF PRODUCTIVITY . . .	160
3.24	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE TASK CATEGORY OF SAFETY	161
3.25	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE TASK CATEGORY OF MAINTENANCE . . .	162
3.26	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE TASK CATEGORY OF MOTIVATION	163
3.27	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE TASK CATEGORY OF MANAGEMENT ATTITUDE	164
3.28	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF LONG-RANGE PLANNING	165
3.29	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF SHORT-RANGE PLANNING	166
3.30	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF PROBLEM SOLVING	167

LIST OF TABLES (CONTINUED)

TABLE		Page
3.31	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF INTER-PERSONAL RELATIONS	168
3.32	CORRELATION BETWEEN COMPOSITE SCORES AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF GOAL SETTING	169
3.33	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF PERSONAL WORK STANDARDS	170
3.34	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE CATEGORY OF OVERALL	171
3.35	CORRELATION BETWEEN COMPOSITE SCORE AND TASK CATEGORY (T1) TEST AREAS. . . .	172
3.36	CORRELATION BETWEEN COMPOSITE SCORE AND BEHAVIOR CATEGORY (T2) TEST AREAS. . . .	173
3.37	CORRELATION BETWEEN COMPOSITE SCORE AND COMBINED BEHAVIORAL AND TASK CATEGORY (T1T2) TEST AREAS	174
3.38	CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON BEHAVIOR AND TASK CATEGORY CRITERIA .	175
3.39	RANK ORDER OF CORRELATIONS BETWEEN COMPOSITE SCORES AND RATER GROUPS AND CRITERIA/TEST AREA COMBINATIONS. . .	176
3.40	CORRELATION BETWEEN INITIAL AND FINAL SCORES ON TEST AREA CATEGORIES BY DIRECT SUPERVISION (G1, G2)	186
3.41	CORRELATION BETWEEN INITIAL AND FINAL SCORES ON TEST AREA CATEGORIES BY INDIRECT SUPERVISION (S1, S2)	187

LIST OF APPENDICES

APPENDIX		Page
A	PERFORMANCE-INCIDENT INSTRUMENT	204
B	HIGH-LOW RANKING FORM	207
C	"HARD-CORE" DATA - SUBJECTIVE RATING INSTRUMENT	213
D	JOB DUTIES AND JOB SKILLS SUBJECTIVE RATING INSTRUMENT	216
E	OVERALL PERFORMANCE SUBJECTIVE RATING INSTRUMENT	244

CHAPTER I
INTRODUCTION TO THE STUDY

PURPOSE

The purpose of this study was to determine whether, in a formal multitrait-multirater appraisal of managerial job performance in an industrial environment, subjective performance appraisal responses correlated significantly with the composite response of the sample population. Of particular interest was to determine which rater within the organizational structure and which specific job related criterion correlated the highest with the composite score.

The results of this study provide the basis for development of an instrument that would appraise managerial-level job performance with minimal time and cost considerations - regardless of the mission, composition or location of the organization - as well as provide guidelines to determine which position in the hierarchy of an organization should initiate the appraisal.

THE PROBLEM

In every organization, the performance of personnel is appraised, whether the organization is formal or informal, large

or small, productive or nonproductive. A consistent, accurate and fair system for the appraisal of job performance is necessary for successful manpower management. It is problematic that, while the notion of the individual performance appraisal (hereafter cited as "IPA") is necessary for successful manpower management, to date and for varied reasons, one specific system does not satisfy all regulatory, organizational or individual requirements.

Whether the IPA is for a decision-making function (e. g., promotion, transfer, retention) or a developmental function, the general assumption is that the more information available, the better the judgment. Within a large organization this assumption may be feasible; but smaller organizations may not have the necessary resources (e. g., fiscal, physical, manpower) for an all-encompassing system with respect to IPA of managerial personnel. Therefore, it must be determined whether a relatively simple, uncomplicated system of IPA can satisfy the regulatory, organizational and individual needs without adversely affecting consistency, accuracy and fairness.

The problem relative to this study is to determine whether it is necessary to have multiple raters, representing various backgrounds and/or work experience levels, evaluating performance criterion, or criteria.

In summary, assuming the optimum predictor is a composite of the entire system as measured by a selected sample, what criteria can adequately represent the entire composite and who can most effectively assign values to this criteria?

BACKGROUND INFORMATION

GENERAL

Because of the controversy and complexity that has existed for many years regarding the appraisal of managerial personnel, many studies have been initiated in an attempt to find the optimum appraisal system. A meaningful study to examine methods for improving the appraisal of management-level personnel was determined to be beyond the resources of one researcher. To conduct research to ensure reliable and valid results typically requires an availability of resources that are normally available only in a large or medium-size organization. A large multi-divisional manufacturing organization, hereafter cited as the "Firm," was conducting a project involving selection of "first-level" managers - commonly referred to in the industry as "foremen" - that was called the Foreman Selection System (hereafter cited as "FSS").

Organizational charts of the "Firm" follow. As applicable, within the text, divisions and plants within each division are designated by an alpha-numeric indicator (e. g., A3 represents Division "A," Plant "3." (Charts "A" and "B")

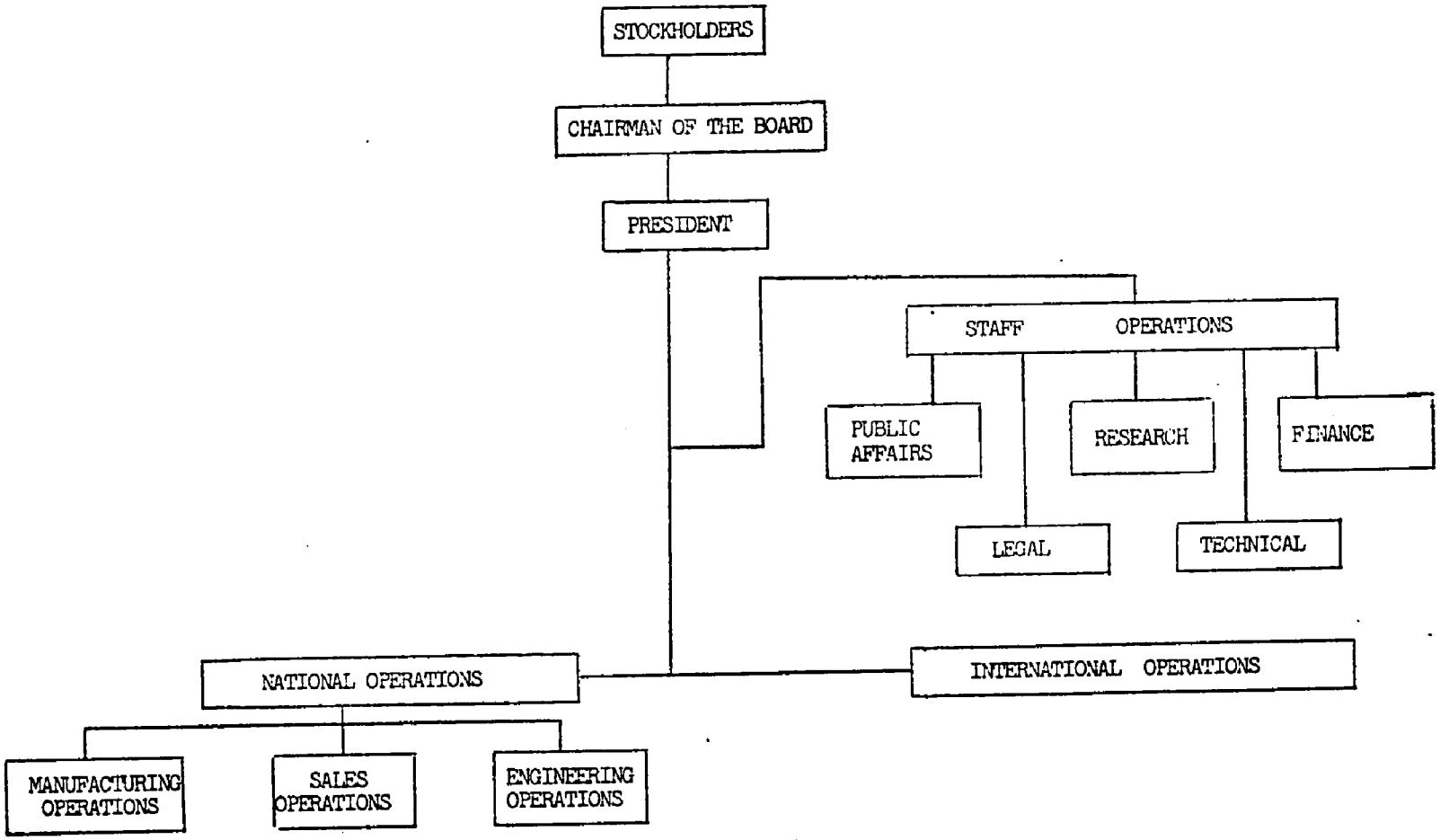
The data that was obtained by the "Firm" for the FSS project was temporarily terminated because of personnel constraints. The "Firm" approved the necessary fiscal amount to conduct the research, had the availability of physical facilities in geographically wide-spread locations with an ample and diverse sample, and had the expertise, interest and knowledge to conduct the project; but did not have the time or necessary human resources to continue the project. Through mutual assistance, the FSS project was reinstated and data were obtained for the "Firm" and this study.

FOREMAN SELECTION SYSTEM

Within the "Firm," the existing method of selecting foremen was in need of revision. This need for revision was two-fold. First, the "Firm" has a continuing need for high caliber foremen; from a business viewpoint, there was an internal indication that the existing selection system was not accomplishing its objective. Second, pressure from the Equal Employment Opportunity Commission (EEOC) required that employer selection procedures meet definite

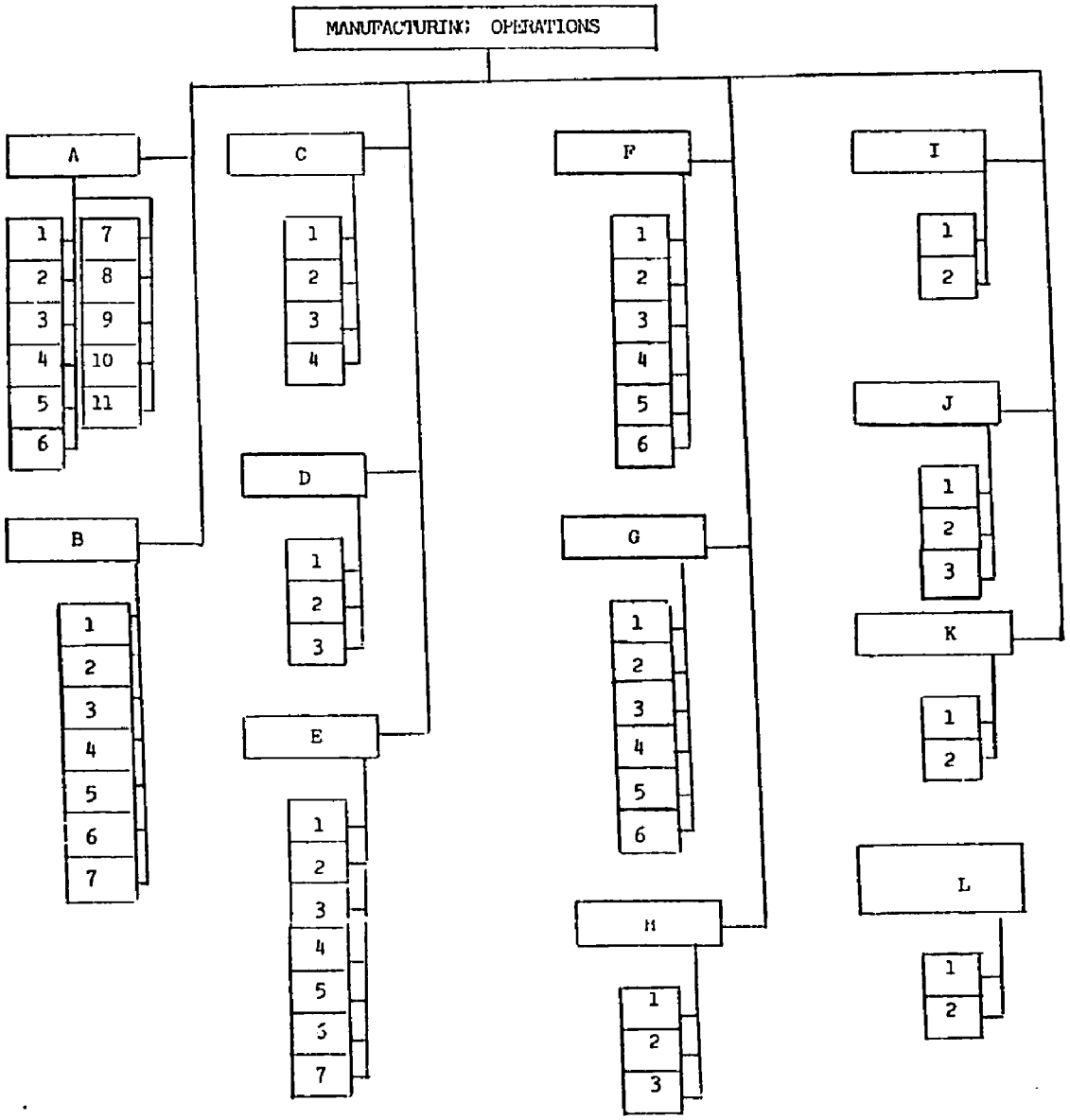
CHART "A"*

ORGANIZATION CHART



*See Chart "B"

CHART "B"
ORGANIZATION CHART



KEY:

DIVISION - Letter

PLANT - Number

FSS SAMPLE LOCATIONS

A3	E4
B5	F6
C1	G4
D3	I2

job relevancy criteria. The EEOC requirements indicate that organizations must be able to provide evidence that selection criteria are valid predictors of job performance.

The purpose of the EEOC Guidelines on Employee Selection Procedures, are ". . .based on the belief that properly validated and standardized employee selection procedures can significantly contribute to the implementation of non-discriminatory personnel policies required by Title VII" (of the Civil Rights Act of 1964).[1]

The FSS project was initiated by the research staff of the "Firm" as an attempt to revalidate existing selection procedures. Because an existing test battery would not revalidate with respect to job relevancy criteria of the EEOC, a standardized selection system was desired which could: (1) be utilized by all plants and divisions within the "Firm," (2) meet the needs of the "Firm," and (3) satisfy the EEOC requirements.

OBJECTIVES

The objectives of the FSS, as determined by the research staff of the "Firm," were to identify, develop, implement and validate an organizational-wide selection procedure for foremen

[1]William H. Brown III (Chairman), Federal Register (Washington: U. S. Government, Vol. 35, No. 139 (August 1, 1970), p. 12,333.

positions which would:

- Identify applicants least likely to be successful foremen;
- Minimize unfair and illegal discrimination against applicants through compliance with governmental regulatory standards;
- Be implemented and operated with minimal cost and time requirements and the greatest possible acceptance of the users of the system; and
- Allow sufficient monitoring and control to ensure that the system continues to operate as intended.[1]

The reasons for the organizational-wide application of the FSS was to ensure consistency in selection, to eliminate any haphazard selection procedures - particularly procedures that affected minority and women applicants - and to monitor and control selection procedures at the research staff level.

ORGANIZATION

To assist in the development of the project, representatives from each division were assigned to two project committees. A Steering Committee was composed of the Salaried Personnel Manager

---

[1]Taken from format developed by the Research Staff of the "Firm."

from each division. This committee served as a high-level decision-making group for the project. A project Working Committee was comprised of individuals as designated by the Steering Committee members and was responsible for the planning and implementation of each phase of the project.

TIMING

A representative timing plan, by phase, for the development of the FSS is presented on the following pages. (Note: Due to previously mentioned human resource constraints, the timing plan was not implemented by the "Firm" as proposed, Table 1.1).

JOB ANALYSIS

An analysis of the job of foreman was conducted within the "Firm" and involved all major foreman classifications within each division.

The major objectives of the job analysis, as outlined by the Research Staff of the "Firm," were to:

- Describe the job of foreman in terms of the demands made upon a job incumbent by focusing on the behaviors associated with successful and unsuccessful task performance;

TABLE 1.1

FOREMAN SELECTION SYSTEM
REPRESENTATIVE TIME PLAN[1]

	JAN.	FEB.	MAR.	APR.	MAY	JUNE	JULY	AUG.	SEPT.	OCT.	NOV.	DEC.
PHASE I	X	X	X	X	X							
Job Analysis and Development of Implementation plan												
Establish job analysis procedures	X	X										
Review existing foreman job descriptions	X	X										
Meet with division salaried personnel managers		X										
Establish Steering and Working Committees		X										
Define general foreman selection model		X	X									
Develop job analysis check list		X	X									
Conduct job analyses in divisions		X	X									
Consensus on accessible dimensions of behavior				X								
Develop implementation plans for divisions				X	X							
				X	X	X						
PHASE II												
Development of Selection System												
Develop exercises				X	X	X						
Develop assessor training program					X	X						
Involve division management					X							
Establish performance criteria					X	X						
Write selection system manual					X	X						
Establish monitoring system					X	X						

(TABLE 1.1 CONTINUED)

	JAN.	FEB.	MARCH	APR.	MAY	JUNE	JULY	AUG.	SEPT.	OCT.	NOV.	DEC.
PHASE III					X	X	X	X				
Implementation												
Involve plant management					X	X						
Pilot test						X						
Implement training						X	X					
Implement system							X	X				
							X	X	X	X	X	X
PHASE IV												
Evaluation and modification												
Collect criteria data							X	X	X			
In-plant audit of implementation of system							X	X	X			
Revise system where required									X	X	X	X
Validate system											X	X
Initiate monitoring system												X

[1]Taken from format developed by the Research Staff of the "Firm."

- Identify any differences in the behaviors associated with successful and unsuccessful task performance within foreman classifications across divisions and between foreman classifications within divisions (e. g., Production Foreman across divisions, Material Handling and Quality Control Foreman within one division);

- Identify the behavioral dimensions which define foreman performance that could be translated into an effective selection system methodology for selecting foremen across all organizational components.[1]

The job analysis consisted of an evaluation of information obtained from three sources. The primary source of information consisted of actual incidents of foreman performance to determine whether the foreman was successful or unsuccessful in the performance of duties. This method, which was based on the Flanagan reported critical-incident method of job analysis, required the observer to describe actual observed incidents of effective and/or ineffective performance. The second source of information was foreman task certification booklets which were developed to provide exposure of new foremen to the basic tasks they would be required to perform. The individual tasks focused primarily on administrative detail and record keeping and were reviewed by the direct supervisor of the foreman to ensure successful completion. The third source of information was direct observation of the

[1]Taken from format developed by the Research Staff of the "Firm."

work environment of the foreman. Information obtained from this source was obtained through tours of the work areas of the foreman on a selected basis.

Specifically, the primary source of information focused on the performance-incident method of job analysis and was selected over the critical-incident method as the most effective means for achieving the objectives of the job analysis. In this method, the direct supervisor recorded effective and ineffective performance-incidents which permitted a range of performance rather than the extremes of performance that is associated with the critical-incident method.

The use of direct supervision was based on the fact, in this industry, the direct supervisor is usually in a position to frequently observe the behavior and performance of the foreman and typically, has previously served in the capacity of a foreman. In addition, the direct supervisor provides a major role in defining the future and ultimate success of the foreman through performance reviews and recommendations for promotion, transfer or retention.

It was hypothesized that the basic elements of effective and ineffective foreman performance were universal to all foreman classifications within the "Firm." To ensure an adequate

test of this assumption, approximately one-third of the performance-incidents were written on non-production foremen.

A performance-incident collection instrument and incident writing training procedures were developed, pilot-tested, revised and implemented. A copy of the final performance-incident instrument and the training aids are included as Appendix A.

The first objective of the performance-incident analysis of the job of foreman was to identify behaviors associated with successful and unsuccessful job performance. These behaviors were identified and sorted into sixteen categories of job performance based primarily on the outcome of the incident or its context. The resultant job performance categories included:

1. Making Production
2. Housekeeping
3. Planning and Scheduling of Machines, Tools and Supplies
4. Using and Allocating Manpower Wisely
5. Controlling and Maintaining Quality
6. Motivating the Work Force
7. Cooperating and Dealing with Others in the Plant
8. Maintaining Safe and Healthy Working Conditions

9. Giving Job Instructions
10. Attending to Administrative Detail and Procedure
11. Knowing the Technical Aspects of the Job
12. Communicating Relevant Information to the Appropriate Persons
13. Disciplining the Work Force
14. Reducing Material Costs
15. Carrying Out Maintenance Responsibilities
16. Displaying a "Management Attitude"

Importance rankings were obtained for the job performance categories within each division. The importance rankings, by division and the total organization, are listed below. (Table 1.2) Brief descriptions of each job performance category are included under definitions and a clustering of the categories are listed in Table 1.3.

The second objective of the analysis was to identify differences in the behaviors associated with successful and unsuccessful task performance within foreman classifications across divisions and between foreman classifications within divisions. The results of the job analysis indicated that, in many foreman classifications, minimum technical job knowledge and skill requirements are necessary for a foreman to succeed. Therefore,

TABLE 1.2

JOB CATEGORY IMPORTANCE RANKINGS BY DIVISION[1]

	DIVISION									FIRM
	B	C	D	E	F	H	I	K	L	
Making Production	1	1	1	3	1	2	1	1	1	1
Housekeeping	7	6	9	7	5	7	3	6	10	6
Planning and Scheduling of Machines, Tools, Supplies	11	12	11	14	13	16	4	10	2	12
Using and Allocating Manpower Wisely	3	4	5	10	6	4	7	3	6	4
Controlling and Maintaining Quality	4	2	3	5	2	5	2	2	8	2
Motivating the Workforce	5	7	2	7	7	13	11	8	4	8
Cooperating and Dealing With Others	13	14	15	15	15	10	13	15	14	16
Maintaining Safe and Healthy Working Conditions	1	4	10	1	9	3	10	7	12	5
Giving Job Instructions	5	3	7	2	4	1	9	8	7	3
Attending to Administrative Detail	15	15	16	12	3	15	14	16	16	14
Knowing the Technical Aspects of the Job	16	13	13	13	16	13	15	13	15	15
Communicating Relevant Information	9	8	8	9	11	9	6	5	3	9
Disciplining the Workforce	14	11	6	4	14	8	5	11	13	11
Reducing Material Costs	10	10	11	12	11	11	8	4	9	7
Carrying Out Maintenance Responsibilities	11	16	14	16	8	12	16	14	5	13
Management Attitude	8	8	4	6	10	6	12	12	11	10

[1]Taken from format developed by the Research Staff of the "Firm."

TABLE 1.3

CLUSTER OF JOB ANALYSIS TASK CATEGORIES[1]

CLUSTER	CATEGORY
Productivity	Making Production
	Using and Allocating Manpower Wisely
	Controlling and Maintaining Quality
	Knowing the Technical Aspects of the Job
	Reducing Material Costs
	Giving Job Instructions
Safety and Housekeeping	Housekeeping
	Maintaining Safe and Healthy Working Conditions
Maintenance	Planning and scheduling of machines, tools and supplies
	Carrying out maintenance responsibilities
Motivating the Work Force	Motivating the Workforce
	Disciplining the Workforce
Management Attitude	Displaying a "Management Attitude"
	Attending to Administrative Detail
	Cooperating and Dealing with Others
	Communicating Relevant Information

[1]Taken from format developed by the Research Staff of the "Firm."

assuming specific technical skill requirements are not significant, inter-classification/intra-division or inter-division/intra-classification differences in the behaviors associated with successful or unsuccessful performance did not exist.

The final objective - to identify the behavioral dimensions which define foreman performance - were identified as follows:

FOREMAN SELECTION SYSTEM

Dimensions of Foreman Performance

I. Short-range Planning -

Up to one day: Planning resources to meet the day's schedule. Come in early to talk to prior shift and identify potential problems before start-up. Anticipation of likely problems. Problem-solving process. Early identification and action on problem indicators. Determining priorities. Recognizing routine problems. Evaluation of consequences of alternatives and preparation of contingency plans. Gathering and evaluating facts for decision-making, including adequacy of information search and the use of key indicators and critical information.

II. Long-range Planning -

From one day to one week: Structured situations for which he will have a few days to develop solutions or plans. Planning for schedule changes. Developing preventive maintenance schedules. Preparing vacation schedules and flex charts.

III. Problem-Solving -

Responsiveness to immediate crises, e. g., schedule changes, breakdowns, labor problems, absences, injuries. Decisiveness. Devising a fast fix. Time stress tolerance including decision-making under time pressure.

Problem diagnosis under time pressure, and crisis management. Speed of response (fast fix). Handling of immediate problems.

IV. Interpersonal Relations -

Consideration - Treats employees with respect (not to imply laxity). Willingness to listen to employee problems. Readily explains reasons for changes and actions. Concern for people. The use of systematic group strategies. Effective utilization of subordinates. Awareness and use of the informal organization.

Interpersonal Awareness - Sensitivity to interpersonal relationships, pressures and conflicts. Awareness of group dynamics. Awareness of the consequences of discipline. Conflict management and resolution. Sensitivity to individual differences including perceiving and reacting to the needs of others. Cooperation and negotiations with others, particularly peers.

V. Goal Setting -

Structuring Tasks - Defines relationships and roles. Sets expectations. Explains operating methods, procedures and work standards (including job instructions). Systematic development of employees. Acceptance of role of instructor and willingness to pass on instructional information. Use of discipline.

Goal Emphasis - Motivates by emphasizing goals (e. g., production, quality,

safety, etc.). Needless employees for greater effort. Tendency to stay with a problem, decision and conviction until the matter is settled.

VI. Personal Work Standards -

General Activity Level - Motivation and initiative. Effort beyond normal expectations.

Thoroughness - Attention to detail and follow through, e. g., willingness to communicate.

Internal Work Standards - Personal integrity. Management attitude. Commitment to plant and organization goals. Desire to do a good job for its own sake. Ability to obtain and use self-feedback.

Follows instructions.[1]

Results of the job analysis confirmed that the overlap of non-technical skill requirements across divisions and across foreman classifications was sufficient to justify common selection techniques provided that the assessment of technical skill requirements for specific foreman classifications is handled individually for each classification.

PREDICTOR DEVELOPMENT

Although a number of methods and techniques are utilized for predicting the job success of non-managerial workers, few

[1]Taken from format developed by the Research Staff of the "Firm."

management tools exist that predict managerial success. This scarcity of predictors is primarily due to the greater complexity of the abilities necessary to succeed at managerial levels. For example, while it might be possible to differentiate between applicants for a job on the basis of manual dexterity, no single skill is available that successfully assesses candidates for a managerial-level job.

The complexity of the abilities necessary to perform successfully as a foreman seem no less complex than those found in other managerial positions. A number of selection system predictor alternatives were evaluated by the "Firm" with their appropriateness to the objectives of the project. A summary of the alternative selection procedures evaluated against the objectives is presented on the following pages (Table 1.4).

The complexity of the job of foreman as defined by the job analysis, coupled with experimental applications of assessment centers within the "Firm" led the research staff to recommend to the Steering Committee the adaptation of an assessment center concept-based selection system in the selection of foremen within the "Firm."

TABLE 1.4

ALTERNATIVE SELECTION PROCEDURES[1]

PROCEDURE	SELECTION OF FOREMEN	COMPLIANCE WITH GOVERNMENT REGULATIONS	IMPLEMENTATION AND OPERATION	MONITORING AND CONTROL
Pencil and Paper Tests	Tests are not available to measure many foreman performance skills. Tests do yield valuable information and should be included in any foreman selection system.	Highly quantifiable, easiest procedure to conduct a validity study on but validity may be unacceptably low. Most tests show sex/race differences. Most likely procedure to receive close government scrutiny.	Most divisions had phased out the use of the Foreman Battery because of its low face validity. A new battery will be no more acceptable. Low cost and time requirements. Relationship between tests and job requirements not readily apparent.	Most quantifiable of the procedures, easiest to monitor and control.
Evaluation of on-the-job performance. (Temporary assignment to foreman position).	Would yield the maximum information about the applicant's ability to perform on the job. Accuracy would depend on the development of a valid performance appraisal system.	Highly likely to meet government guidelines. Must be validated. Validation study would be fairly complex because of nature of the "test" information. Procedure is the most likely to be valid. Unlikely to be scrutinized by the EEOC and OFCC. Should minimize sex/race differences.	Impractical without considerable prescreening of applicants. High cost and time requirements as applicant could only be evaluated after he had learned the job, i. e., one or more months. Would result in increased inefficiency. Development of appraisal system would be complex.	Difficult to monitor and control.

ALTERNATIVE SELECTION PROCEDURES (CONTINUED)

PROCEDURE	SELECTION OF FOREMEN	COMPLIANCE WITH GOVERNMENT REGULATIONS	IMPLEMENTATION AND OPERATION	MONITORING AND CONTROL
Selection by date of application (random selection)	A large number of unsuccessful foremen would be hired, leading to high foreman turnover, increased inefficiency, and potential morale and discipline problems with hourly.	Would meet government standards; little chance of illegality; least likely system to be challenged by the government; need not be validated.	Minimum cost and time requirements to implement and operate; would only be acceptable to those locations where the applicant population is nearly equal to the available job openings.	None required
Foreman Selection Committee - Group interview without training	Least likely to yield information on the performance potential of applicants.	High potential to discriminate unfairly; will not (present) meet government guidelines; must be validated but will be difficult to validate; highly subjective; likely to exaggerate sex/race differences and encourage government scrutiny.	Already used in most plants; fairly low time requirements; no cost. Most plants not satisfied with this method - not obviously related to the foreman's job. Moderate to low face validity.	Documentation and quantification of the selection process very difficult because of subjective nature.
Foreman Selection Committee - Interview with training	Will yield more information on performance potential but not enough to justify extra time and cost involved; increases the effectiveness of multiple evaluations of applicants.	Less likely to discriminate unfairly but still will not meet present guidelines. Must be validated but unlikely to validate; likely to exaggerate sex/race differences and encourage government scrutiny.	Will greatly increase time requirements. Training costs comparable to assessment center operation; low face validity, unlikely to be acceptable to most FSC members.	Documentation and quantification better than above, but still inadequate for effective monitoring.

ALTERNATIVE SELECTION PROCEDURES (CONTINUED)

PROCEDURE	SELECTION OF FOREMEN	COMPLIANCE WITH GOVERNMENT REGULATIONS	IMPLEMENTATION AND OPERATION	MONITORING AND CONTROL
Full-scale Assessment Center 2-3 days 6 Applicants at a time	Next to on-the-job evaluation, would yield the widest range of information on applicant's potential to perform the job. Exercises are simulations of foreman behavior. Makes maximum use of multiple evaluations. Assessment Center information can be used to develop foreman applicants.	Fairly easy to conduct a validation study and highly likely to be valid. Will meet EEOC, OFCC guidelines. Likely to minimize sex/race differences and less likely to be challenged by the government.	Can be expensive to implement and operate, especially if outside facilities are used. Requires high time commitment from plant personnel, and highly trained assessors. High face validity because exercises are simulations of job behavior. Centers cannot process less than six candidates at a time. Uses general foreman and superintendents as necessary.	Primarily uses subjective data but fairly easy to document and quantify the selection decision process. Direct monitoring of the center would be fairly easy as candidates are processed in groups of six.
Scaled-down Assessment Center less than one day	There is evidence from outside the "Firm" that a reduced assessment center is just as valid as a more elaborate center. See above.	Does not require a large block of assessor time. Low cost per candidate processed; high face validity because exercises are simulations of job behavior. Can handle one or more applicants at a time as required. Requires highly trained assessors. Uses general foremen and superintendents as assessors.	Does not require a large block of assessor time. Low cost per candidate processed; high face validity because exercises are simulations of job behavior. Can handle one or more applicants at a time as required. Requires highly trained assessors. Uses general foremen and superintendents as assessors.	Uses subjective data but more heavily emphasizes documentation and quantification than full-scale center. More difficult to audit than full-scale center.

[1]Taken from format developed by the Research Staff of the "Firm."

NOTE: Additional selection system issues not considered include the Labor Relations check and technical skill requirements of the position.

On the basis of previous organizational experience, a review of the objectives and needs of the organization, comparisons of the alternative selection methods and other staff research and recommendations, a scaled-down assessment center was selected as the model for the FSS.

SUMMARY

An analysis of the job of foreman was conducted within the "Firm" to determine which abilities differentiate successful foremen from unsuccessful foremen. The performance-incident technique of job analysis was utilized to obtain data and written and verbal incidents of foremen performance were collected from direct supervisors.

The performance-incidents were categorized into sixteen work outcome or job content areas and further categorized into elements of what the foreman did (process) and what was done (content). The elements were then sorted into logical groupings or dimensions of performance. The resultant dimensions became the basis for the development of the selection procedures.

It was decided that an assessment center concept-based selection procedure would best predict the foreman performance dimensions identified in the job analysis and, thus, became the

nucleus of the FSS.

NEED FOR THE STUDY

An evaluation technique for managerial-level personnel utilized by most organizations to periodically measure performance is the IPA. Typically, the basis for the IPA is either administrative decision-making or professional development of the individual. In addition, as an indirect result, IPA instruments have the potential to provide valuable data for research or improvement of existing systems. Recently, a new basis for the IPA has emerged - the requirement to comply with governmental mandates in accordance with the Civil Service Act of 1964.

Presently government agencies, such as the EEOC, are examining IPA programs regarding relevance to job requirements, cultural and racial biases, reliability and validity. In addition, the ultimate burden of proof in demonstrating the accuracy, consistency and fairness of IPA practices, policies, techniques and instrumentation has been placed with the organization by the regulatory agencies.

Therefore, there is a need for this study, on a generalized basis, to provide organizations - regardless of size, mission, or

location - a consistent, predictable and cost effective means to appraise performance of managerial personnel that:

(1) meets or exceeds external or internal regulatory standards,

(2) is in accord with organizational objectives and needs,

(3) ensures proper utilization of fiscal, physical and manpower resources,

(4) provides this appraisal with minimal time constraints to rater and ratee, and,

(5) provides the maximum possible acceptance to the organization and the individual.

HYPOTHESIS

This study was designed to test the hypothesis that: in the administration of the individual performance appraisal, there will be differences in how raters from different hierarchical levels within the organization appraise job-related criteria on subordinate members of the organization.

In testing this hypothesis, stated in the null hypothesis in Chapter III, it was hoped that the answers to the following questions could be obtained.

1. In the following test area-rater matrix, which test area, rater, and test area-rater combination have the highest correlation with the composite?

2. In the following test area-rater matrix, is there a difference in initial and final responses?

TEST AREA	RATERS G1 = G2	RATERS S1 = S2
T1	X	X
T2	X	X
T1T2	X	X

TEST AREA/CRITERIA/RATER MATRIX

CRITERIA/ TEST AREA CODE*	RATER CODE*									
	GF1	GF2	S1	S2	I	XGSI	XG	XS	XGXS	GSI
PROD	X	X	X	X	X	X	X	X	X	X
SAFETY	X	X	X	X	X	X	X	X	X	X
MAINT	X	X	X	X	X	X	X	X	X	X
MOTIV	X	X	X	X	X	X	X	X	X	X
MGTATT	X	X	X	X	X	X	X	X	X	X
T1	X	X	X	X	X	X	X	X	X	X
LRPLAN	X	X	X	X	X	X	X	X	X	X
SRPLAN	X	X	X	X	X	X	X	X	X	X
PSOLV	X	X	X	X	X	X	X	X	X	X
INTREL	X	X	X	X	X	X	X	X	X	X
GOALS	X	X	X	X	X	X	X	X	X	X
WRKSTD	X	X	X	X	X	X	X	X	X	X
OVERAL	X	X	X	X	X	X	X	X	X	X
T2	X	X	X	X	X	X	X	X	X	X
T1T2	X	X	X	X	X	X	X	X	X	X

* - See Table 3.9 for description of criteria, test area and raters.

METHODOLOGY

The source of data for this study was obtained through research conducted in conjunction with the FSS of the "Firm."

SAMPLE

The sample was comprised of a minimum of nine individuals per plant location who were in the position of foreman. The sample number and plant location for each site follows. The sample was divided between currently ranked above-average performing foremen and currently ranked below-average performing foremen. This division was obtained by having "second-level" supervision rank all of their foremen on the basis of overall job performance, using a multiple alternation ranking procedure. The instrument utilized for this ranking is displayed in Appendix B. The individual sample subsequently selected by research staff personnel was based upon the multiple alternation ranking, time in present position, length of service, current performance review rating, and other input provided by plant management (Table 1.5).

INSTRUMENTATION

The following resource instruments were distributed, explained and administered at each location with emphasis that information

TABLE 1.5

NUMBER AND PERCENTAGE OF SAMPLE POPULATION
BY LOCATION, SEX AND RACE

LOCATION CODE[#]	SEX/RACE			TOTAL
	MALE		OTHER*	
	WHITE	BLACK		
A3	6	4	-	10
B5	9	4	2	15
C1	9	-	-	9
D3	7	-	3	10
E4	7	2	1	10
F6	12	2	2	16
G4	10	-	-	10
I2	6	2	2	10
TOTAL	66	14	10	90
PERCENT	73.3	15.6	11.1	100.0

[#]See Organization Chart "B."

*Includes Female Foremen and Missing Data.

obtained was for "research purposes only" and strict confidence would be maintained. Instruments utilized in this study are presented as outlined below. Samples of instruments are displayed in the Appendices of this study.

SAMPLE INSTRUMENT	APPENDIX
"HARD-CORE" DATA - SUBJECTIVE RATING	C
JOB DUTIES AND JOB SKILLS - SUBJECTIVE RATING	D
OVERALL PERFORMANCE - SUBJECTIVE RATING	E

"HARD-CORE" DATA INSTRUMENT

The hard-core performance data instrument was developed through field interviews and pilot testing. The applicability of each criterion was determined by each plant on whether the data could be obtained on a given foreman and whether the foremen had any control or influence over the criterion.

The immediate supervision of each foreman selected for the sample "tracked" the foreman for a one-month period (two-month period at two locations - total N = 31) on selected quantifiable hard-core performance factors, such as scrap, absenteeism, and productivity, which were identified for each location. A matrix

of criteria by location is listed on the following page (Table 1.6).

On a daily basis, the rater recorded the selected data and, each week, the rater subjectively rated the foreman on the data that was recorded. A representative instrument of the combined daily/weekly record and weekly evaluation of hard-core data is displayed in Appendix C.

SUBJECTIVE DATA INSTRUMENT

The subjective rating instrument was developed utilizing criteria obtained during the job analysis phase of the FSS project and through consolidation and categorization by plant management. The instrument was comprised of rating scales on five categories of job duties and six categories of job skills. A listing of the final criteria is provided below; the final subjective rating instrument is displayed in Appendix D.

SUBJECTIVE CRITERIA

Job Duties -

Productivity
Safety and Housekeeping
Maintenance
Motivating the Workforce
Management Attitude

TABLE 1.6

PERFORMANCE DATA AVAILABILITY BY PERFORMANCE INFORMATION BY LOCATION

PERFORMANCE INFORMATION	LOCATION								TOTAL
	C1	A3	G4	I2	B5	D3	E4	F6	
Production Factors									
Scrap	X	X					X		3
Rejects		X		X				X	3
Defective Tag Operations					X				1
Off-standard Direct Labor	X	X	X			X	X	X	6
Direct Labor Variance				X					1
Net Productivity	X			X		X	X		4
Indirect Material Cost		X							1
SUB-TOTAL	3	4	1	3	1	2	3	2	19
Personnel Factors									
Absenteeism - Foreman	X	X	X	X	X	X	X	X	8
Absenteeism - Hourly	X	X	X	X	X	X	X	X	8
Pay Shortages		X	X		X		X	X	5
Ring Failures			X		X		X	X	4
Grievances				X					1
Disciplinary Action Reports				X	X				2
Safety Violations						X	X		2
Work Habits						X			1
SUB TOTAL	2	3	4	4	5	4	5	4	31
TOTAL	5	7	5	7	6	6	8	6	50

Job Skills -

Long-Range Planning
Short-Range Planning
Problem Solving
Interpersonal Relations
Goal Setting
Personal Work Standards

At the end of the one-month recording of hard-core data the same immediate supervisor subjectively rated the foreman. Concurrent with this rating, the "second-level" supervisor to the foreman completed an identical subjective rating form. Ratings were obtained immediately following a training session for both supervisors (direct and indirect supervision) regarding how to complete the rating scales and common rating errors.

In the same session, a member of the research staff (non-related supervision) conducted a performance-incident interview with both supervisors and, subsequently, also rated the foreman using an identical subjective rating form. Approximately one month later, the instrument was readministered with one additional sheet "overall performance." A facsimile of the additional rating sheet is displayed in Appendix E.

STATISTICAL ANALYSIS

Appropriate analyses were made utilizing the statistical and computer assistance available through The University of Michigan.

DEFINITION OF TERMS

For descriptive purposes, the title "foreman" is synonymous with and is implied as "first-level" management. Within the organizational hierarchy of the "Firm," the term "general foreman" is synonymous with "second-level" management and is the direct supervision or direct supervisor of the foreman. The "third-level" managerial position in this firm is titled "superintendent" and is the indirect supervisor or indirect supervision to the foreman. The term "non-related" supervisor/supervision is a managerial position within the organizational hierarchy (either line or staff) that is a "third-level" managerial position or higher.

A diagram of the line hierarchy of the "Firm" by managerial level is shown in Figure 1.

FIGURE 1

LEVELS OF MANAGEMENT
(LINE)

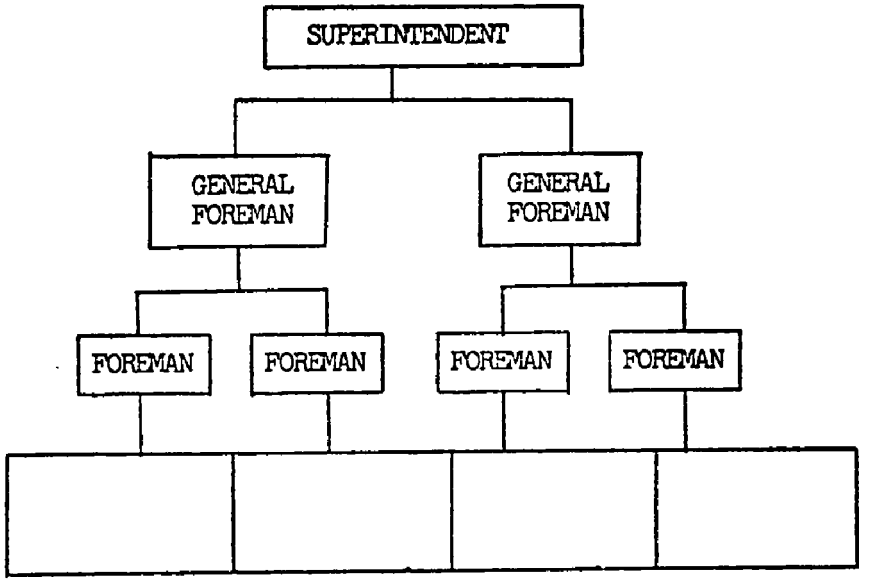
LEVELS OF MANAGEMENT

THIRD

SECOND

FIRST

HOURLY PERSONNEL



Within the organizational structure of the "Firm," one of the most important positions that determines the success of the organization is "first-level" management. The first-level" manager typically provides organizations that have a hierarchy of duties and responsibilities, the best opportunity to improve employee attitudes. This opportunity for improvement is available because the "first-level" manager is the direct representative and product of the organizational measurement and appraisal systems within an organization - and of the way in which they are administered and managed.

Following are other terms relative to this study:

Assessment - Any method used to evaluate characteristics of individuals.

Bias - Any constant error or systematic influence on measures irrelevent to the purpose of the measurement.

Correlation - The degree to which two or more sets of measurements vary together.

Criterion (plural, criteria) - Any behavior, performance level, or result of behavior to be predicted.

Composite Score - The sum of scores on different assessments to form a single score; a linear combination.

Predictor - Any measureable characteristic used to predict criterion performance.

Reliable - Consistent or dependable or repeatable; consistency of measurement.

Validity - The degree to which certain specified inferences from scores may be considered justified.

JOB PERFORMANCE CATEGORIES

- Making Production - Meeting assigned production quotas, meeting work standards for each operation, improvising if necessary to keep the lines running, preventing parts shortages.

- Housekeeping - Maintaining orderly work area, keeping grease, oil, etc., off the floor, providing waste containers.

- Planning and Scheduling of Machines, Tools, and Supplies - Includes repairing machines when necessary, making sure inventory of stock is sufficient to keep machines running, etc.

- Using and Allocating Manpower Wisely - Controlling direct labor costs, overtime equalization, assigning jobs to the "right" people.

- Controlling and Maintaining Quality - Not permitting defective parts to be used, supervising machine operations to ensure that bad parts are not produced, not trying to slip defective parts through by "selling" QC, handling parts gently, adhering to inspection process etc.

- Motivating the Work Force - Talking privately to a man to make him productive, complimenting the good worker, explaining reasons for actions, being friendly to employees, setting the example, etc.

- Cooperating and Dealing with Others in the Plant - Taking advice from engineering or other service departments, following procedures set by other departments, not taking advantage of new men in other departments, etc.

- Maintaining Safe and Healthy Working Conditions - Delivering the safety talk, following safety rules, making sure employees wear safety glasses.

- Giving Job Instructions - Showing a man exactly how a job operation should be accomplished, explaining details of a job to a new man, observing to be sure the new man knows how to do it, etc.

- Attending to Administrative Detail and Procedures - Promptly authorizing time sheet adjustments, noticing errors such as sweepers being carried direct rather than indirect, keeping up "blue cards," entering production count, following standard procedures, etc.

- Knowing the Technical Aspects of the Job - Getting trained on new machines in his area, understanding technical aspects of the work, etc.

- Communicating Relevant Information to the Appropriate Persons - Letting other shifts know about problems, informing the boss when important events occur, letting other departments know about problems.

- Disciplining the Work Force - Administering discipline when appropriate, knowing and enforcing the union contract and "Firm" rules.

- Reducing Material Costs - Preventing scrap, keeping damage at a minimum, reducing cost of indirect materials.

- Carrying Out Maintenance Responsibilities - Identifying and resolving maintenance problems, planning and scheduling machine maintenance.

- Displaying a "Management Attitude" - Being present for work consistently, cooperating with the general foreman and other management, refusing to pass the buck, setting an example, supporting the "Firm" and bosses when talking to others (especially employees).

JOB DUTIES

- Productivity - Consistently meeting assigned work goals; controlling for and maintaining acceptable quality levels; efficiently controlling labor costs by assigning the "right people" to jobs and providing employees with adequate job instructions and directions· controlling material costs by holding damage and waste to a minimum.

- Safety and Housekeeping - Maintaining an orderly work area and enforcing adequate clean up; promoting good housekeeping; ensuring that safety rules are followed.

- Maintenance - Ensuring that the right equipment and supplies are on hand to do the job; identifying and resolving maintenance problems when appropriate; scheduling necessary equipment repair; maintaining an adequate inventory of supplies and tools.

- Motivating The Workforce - Complimenting good workers; being courteous and explaining the reasons for actions concerning employees; setting a good example· enforcing the "Firm" rules and the union contract and disciplining employees fairly when appropriate; showing concern for employees.

- Management Attitude - Supporting the "Firm," particularly when talking to employees; refusing to pass the buck and setting a good example for employees; performing administrative responsibilities, i. e., labor counts; communicating relevant information to those who need to know; cooperating with other departments and areas.

- Long-Range Planning - Planning and organizing resources in the context of a relatively structured situation (e. g., vacation schedules, scheduled process changes, etc.) where one day to one week is available for developing solutions or plans.

- Short-Range Planning - Planning and scheduling resources to anticipate problems and meet job requirements where less than one day is available; anticipating likely problems, lining up the day's work; early identification and action on problem indicators.

- Problem Solving - Responding to problems demanding immediate attention; responding to crises; decisiveness and stress tolerance.

- Interpersonal Relations - Sensitivity to interpersonal relationships, pressures and conflicts; treating employees with respect and concern (not implying laxity); delegating; obtaining cooperation from employees; cooperating with others; persuasiveness; resolving conflict; effective use of discipline.

- Goal Setting - Emphasizing goal attainment to employees (e. g., production, quality, safety); applying pressure for productive output; define roles and relationships; setting expectations and explaining work procedures; giving job instructions; enforcing rules; disciplining employees.

- Personal Work Standards - Being thorough and attending to detail; following through; staying with a problem or decision until the matter is settled; expending effort beyond normal expectations; showing initiative; having personal integrity and high work standards; follows instructions.[1]

TABLE OF ABBREVIATIONS

IPA - Individual Performance Appraisal

FSS - Foreman Selection System

EEOC - Equal Employment Opportunity Commission

G1 - Direct Supervisor - Initial Rating

G2 - Direct Supervisor - Final Rating

S1 - Indirect Supervisor - Initial Rating

S2 - Indirect Supervisor - Final Rating

I - Non-Related Supervisor

T3 - Test Area - Hard-core Data

T1 - Test Area - Subjective Data - Task

T2 - Test Area - Subjective Data - Behavioral

N - Number of Individuals/Sample

---

[1]Taken from format developed by the Research Staff of the "Firm."

LIMITATIONS

Data, as well as subsequent findings and conclusions based on these data, pertain directly to the sample and locations related to the "Firm;" however, application of findings and techniques utilized could be transferred to any organization for purposes of performance appraisal, whether the organization is industrial, governmental, or educational.

The raters in this study were comprised only of supervisory personnel and only "first-level" managerial personnel (foremen) were selected for the sample and were evaluated. There were no peer, subordinate or self-evaluation ratings or rankings administered.

As administration of the instruments did not require highly sophistocated training, data were collected at widespread locations by persons other than the researcher. The raters were appropriately instructed with regard to procedures for administration of the instruments and they were informed of the confidentiality and "research purposes only" use of the data.

Because of the varied missions within each sample site location, the hard-core data varied by plant location. Also,

the multiple plant locations were subject to varying fiscal, physical, and manpower constraints, requirements and environment. Due to the varied site locations, not all individuals were rated or ranked by one individual.

For statistical purposes, the assumption that the composite score is the "optimal" score, and a linear combination is made.

The effort expended in obtaining the data used in this study was based on the general assumption that the more information available, the better the decision, provided the information is relative to the study. The use of three separate raters and rankers, over multiple time intervals, and the use of three distinct instruments is not practical or possible in most organizations due to time, manpower, and/or cost constraints.

SUMMARY

This study was, in general terms, a basis for determining which universe of raters would provide the most valid and reliable assignment of managerial performance appraisals.

A study of this nature is necessary to the conservation of manpower in any type of organization as well as ensure that

federal, state, local and organizational goals, standards and directives are implemented and maintained.

The defined hypotheses stated as questions, were discussed in detail and a thorough statistical analysis and interpretation was completed. Pertinent aspects of the study, as discussed in the literature, were reviewed. Design of the overall study, instrumentation, sampling, scoring procedures and sample selection were discussed in detail. An analysis of the results of the study and how the results relate to the hypotheses were developed.

Chapter I contains a discussion of the purpose of the study, defines the problem, presents hypotheses, outlines procedure and methodology utilized, lists applicable definitions, and identifies limitations of this study.

Related literature and discussion of typical methods of performance appraisal is contained in Chapter II.

Chapter III consists of data analyses and interpretations.

Chapter IV includes a summary of information, conclusions regarding the study, implications and recommendations for additional study.

CHAPTER II

REVIEW OF LITERATURE

INTRODUCTION

Because of the current regulatory, economic and fiscal pressures being imposed on many organizations, there has been an upward trend in an attempt to improve organizational effectiveness and efficiency. The trend is evident in business, government and education; applying not only to fiscal and physical commodities (e. g., inventory control, cost-benefit, facility use), but also to manpower management.

This chapter will provide a comprehensive and systematic review of one manpower management "tool" - the individual performance appraisal.

The nomenclature utilized to identify the process to appraise personnel varies from organization to organization. In this study, the following terms are assumed synonymous:[1]

[1] H. S. Roberts, Roberts' Dictionary of Industrial Relations (Washington: BNA Books, 1971), pp. 114, 320, 401.

Performance Appraisal - A systematic review of an individual employee's performance on his job to evaluate the effectiveness or adequacy of his work. . . .

Merit Rating - A system used to measure periodically the relative work performance of employees.

Efficiency Rating - The periodic appraisal of an employee's work performance, usually on the basis of factors which seek to measure the most important qualities required on his job. . . .

The following review of pertinent literature focuses on formalized management-level individual performance appraisal techniques in business, governmental and educational environments within the United States. Emphasis is directed toward organizations within the business environment, but specific differences in other organizational environments are discussed as appropriate. Discussion of organizations established for social reasons are not included in this study.

For convenience, the term "individual performance appraisal" will be cited as "IPA." In addition, the use of the terms "system," "program," "technique," "process," "approach," "procedure," and "method," in conjunction with the IPA are assumed to be

interchangeable as indicated:

Program = System

Approach = Technique

Process = Method = Procedure

Because of the multi-organizational environment, a cross-reference tabloid of terminology applicable to this review is provided in Table 2.1.

HISTORICAL OVERVIEW

Pay and benefits, job security, working conditions and a safe environment are among the most fundamental job characteristics utilized to attract, maintain and retain individuals in any organization. These characteristics are manifestations of the values stressed within the organization, the concern of the organization for its employees, the realities of the labor market and the manpower resource availability.

One effort by organizations to maximize the use of manpower resources has been the initiation and maintainance of IPA programs. Historically, IPA programs are a relatively recent acquisition on the management "tool shelf." The introduction of formal

TABLE 2.1

ORGANIZATION CROSS-REFERENCE
MANAGERIAL LEVEL

TERM	ORGANIZATIONAL ENVIRONMENT		LEVEL OF MANAGEMENT
	BUSINESS AND GOVERNMENT	EDUCATION	
MEMBER INDIVIDUAL	WORKER EMPLOYEE SUBORDINATE	TEACHER (K-12) INSTRUCTOR PROFESSOR	FIRST
LEADER	SUPERVISOR SUPERIOR EMPLOYER	PRINCIPAL (K-12) DEPARTMENT HEAD	SECOND
MIDDLE MANAGEMENT	SUPERINTENDENT MANAGER	SUPERINTENDENT (Local) DEAN	THIRD
TOP-MANAGEMENT	CHAIRMAN PRESIDENT DIRECTOR EXECUTIVE	SUPERINTENDENT (State) PRESIDENT	*
ENVIRONMENT	ON-THE-JOB WORK	IN-THE-CLASSROOM	-

*Depends on Organization.

performance appraisal techniques in the United States has been traced to the 1850's, where federal government offices were reported to have utilized personnel rating forms. In education, the use of rating forms in the appraisal of teachers began in the 1890's. Business began using formal performance appraisal programs in the early 1900's. [1] Since the early 1900's, it has been estimated that over three-fourths of the businesses in the United States have IPA programs.[2]

During the past five decades, many organizations have experienced dynamic growth and change. Concurrently, the appraisal of personnel performance has also experienced growth and change. Although the reasons for the appraisal of individuals have not drastically changed, many factors, external and internal to modern day organizations, have necessitated refinement and updating of existing IPA systems or the development and implementation of new, improved IPA systems.

---

[1] Thomas L. Whisler and Shirley F. Harper (Eds.), Performance Appraisal (New York: Holt, Rinehart and Winston, 1962), p. 423.

[2] Winston Oberg, "Make Performance Appraisal Relevant," Harvard Business Review (January-February, 1972), p. 61.

During this century, there have been significant changes in the structure of the labor market. First, the occupational structure of the labor force has shifted significantly from the agricultural environment. The largest recipient of this manpower shift has been in service-related industries and the government. Concurrent with this shift, a significant change in the mix of skills has occurred, with a resultant increase in white-collar workers and a decrease in blue-collar workers.

Second, the rapid development of automation has influenced the mix of skills as well as the labor market. Although automation has primarily influenced manufacturing operations, the advent and increased utilization of the computer has begun to significantly effect many non-manufacturing organizations and functions.

Third, the trend toward increasing organizational size continues, especially in service and government related organizations. The proverb "the bigger the better," has apparently provided the stimulus for this trend.

Fourth, the growth of unionization in the workforce has altered the viewpoints of individuals in the labor market. Although the unionization of the blue-collar worker appears to have reached a plateau, growth of unions is currently being

introduced or expanded in many white-collar labor markets.

Fifth, the increase of regulatory activity by federal, state and local legislators has significantly influenced the workforce and labor availability. Legislation influencing the age, sex and race of employees, have each altered the composition of the available workforce as well as the positions available for those seeking employment.

Sixth, not only have organizations increased in size, but the number of organizations within many industries has also increased. The effect of this change has permitted greater job availability for the individual. This increase in job availability is the result of automation, union control over work standards, expansion of staff activities, and federal, state and local legislative actions.

Seventh, the increased size and number of organizations has created more complex job structures. Because performance measurement is usually most difficult in large-scale activities and organizations, the relationship of individual effort to group output has become increasingly remote; therefore, new and improved methods for appraisal of performance have become necessary.

The changes in the labor market discussed have also significantly affected education and levels of academic achievement within the labor market. With the increased enrollment at all levels of education, especially in institutions of higher education, more resources (i. e., fiscal, physical, human) have become necessary. Concurrent with the increase of educational resources, taxpayers, legislators and doners have demanded and expect greater accountability of each resource. Because accountability is a basic consideration of efficient and effective management, demands for improved methods of appraising faculty and administrators have evolved.

Within organizations, changes have occurred which are both dependent and independent of the evolution in the labor market. First, primarily due to influence of military organizations, utilization and dependence on staff functions and staff activities has increased.

Second, in conjunction with the increased staffing trend, there has been an increase in individual specialization, especially in larger organizations.

Third, the "philosophy of management" has evolved from the exclusively traditional scientific approach to a blend that

includes the behavioral, or human relations, approach.

Fourth, intra-organizational change has been significantly influenced by the shifting of appraisal based primarily on a male, hourly workforce to appraisal involving men and women in both hourly and managerial positions.

PHILOSOPHY

Every organization has its own philosophy regarding operation. Likewise, the IPA that is successful for one organization may, for a variety of reasons, fail completely in another organization. This necessitates that each IPA program be individually designed to satisfy the unique characteristics of the organizational environment (e. g., mission, location, manpower).

The philosophical approaches for IPA programs have been classified as casual, traditional and behavioral.[1]

The casual IPA approach, by definition, is unsystematic, informal and often inconsistent in nature. This approach is

---

[1]Edwin B. Flippo, Principles of Personnel Management (New York: McGraw-Hill Book Company, 1966), p. 250.

mostly utilized in organizations that do not have, or do not desire to commit, the necessary resources to implement and maintain a formalized IPA program. Because of the varied methodology that may be utilized and the assumption that this report involves formal IPA programs, this approach will not be discussed.

The traditional IPA approach typically provides a highly systematic measurement of individual characteristics or contributions. The primary objective of this approach is accurate measurement of human performance by reducing, or eliminating, human bias or prejudice through utilization of a formal reporting instrument and system.

The basis for the traditional, or systematic, IPA approach is that it will provide information useful in making or enforcing decisions. The information obtained is usually available in advance of the time needed; thereby, reducing or eliminating the possibility of having to make "spot decisions." In many instances, the results obtained through the systematic approach, assist the decision-maker by providing data and information in a form that permits comparison. Comparison is possible because the individuals within the organization are appraised on the same criteria and in the same manner.

In addition, this approach tends to stimulate and aid in employee development. Individuals generally are interested in how their performance is perceived by others. The IPA may provide this type of information in a manner that is capable of being communicated to the individual. Typically, the variables used in the IPA process provide a pattern of desired behavior and, by comparing this pattern with actual performance, the individual is provided with a "snapshot" of strengths and weaknesses.

A procedural benefit from use of the traditional approach may be the improvement in the competency of the rater. The potential for increased rater competency results when the rater becomes better acquainted with the ratee as well as the position being evaluated.

The control aspect of the formal, traditional IPA program has been challenged on many occasions. This control issue is primarily in reference to the reward systems of the organization and the "power" the rater has in controlling the destiny or career of the individual. Many critics of this control aspect, such as McGregor, Drucker, and Whisler, maintain the rater should not be the sole individual ". . .for judging the personal worth of a fellow man."[1]

---

[1]Douglas M. McGregor, "An Uneasy Look at Performance Appraisal," Harvard Business Review, Vol. 35, No. 3 (May-June, 1957), in Whisler and Harper, p. 72.

McGregor and others have also indicated that second-guessing by the ratee of what the rater wants could, and often does, cause the ratee to suppress aspirations and to restrict creativity. Experience indicates that when no effective method of appeal is available, inequity becomes the result. An example of control, suppression and inequity related to one form of IPA was presented by Herman Wouk in his novel, The Caine Mutiny. In the book, Captain Queeg purposely suppressed the possibility of transfer to other duty stations through submission of mediocre individual performance ratings.

The limitations of the traditional IPA system have reduced its acceptance by many organizations. The "power role" of the rater in the traditional IPA system may produce dysfunctional consequences of conformity. The IPA format may not adequately measure the elements of performance designed to be evaluated and rating variables may often only be incidently related to actual results. In addition, the usual secrecy in which the rater completes the performance appraisal usually serves no useful purpose, especially to the ratee.

Although the traditional, systematic IPA approach has many pros and cons, the approach is more consistent and useful than the casual approach and is successfully utilized in many organizations.

The drawbacks of the traditional IPA approach have become the basis for a third approach, commonly known as the behavioral, or participative approach.

McGregor, a strong advocate of the participative approach, has indicated traditional IPA programs are designed to meet three needs. These needs, one for the benefit of the organization and the other two for the benefit of the individual, are as indicated:

> 1. They provide systematic judgments to back up salary increases, promotions, transfers, and sometimes demotions or terminations.
>
> 2. They are a means of telling a subordinate how he is doing, and suggesting needed changes in his behavior, attitudes, skills, or job knowledge; they let him know 'where he stands' with the boss.
>
> 3. They also are being increasingly used as a basis for the coaching and counseling of the individual by the superior.[1]

Although the three needs are often satisfied through unilateral, traditional techniques in many organizations, McGregor and others believe a participative or behavioral approach is more effective. The participative approach

---

[1] Ibid., p. 71.

advocated by the behavioralists represents a combination of current behavioral science research and theory with practical business objectives.

Because measurement of individual performance is an essential component in the measurement of organizational performance, the behavioralists believe each member of the organization should know for what he/she is accountable, how his/her position effects the overall performance of the organization, and how his/her contribution is measured.

General principles and comments relative to the IPA system that combines behavioral theory and business practices are as indicated:

> Improvement in job performance can be accomplished most effectively if the employee participates directly in establishing the achievement measures for his job. To do a better job, he should know what is expected of him, how he is doing on the job, and how and where he can get assistance when he needs it.
>
> Employees are inclined to accept suggestions for improving and/or maintaining their performance when they are offered feedback in a less concentrated form than is the case of an annual discussion. Studies of the learning process indicate that feedback is less effective the more time is allowed to elapse between performance and feedback (reinforcement).

> The supervisor's prime responsibility is as a coach and collaborator in individual development, not as a judge. Two major forces in an employee's development must be given full consideration to achieve maximum results:
>
> a) The employee must recognize the understanding and concerns of his immediate supervisor as that of an active partner in helping him.
>
> b) The supervisor must recognize and be concerned with the employee's personal aspirations, motivation, and growth needs.
>
> The primary purpose of conducting a formal performance appraisal is to further develop the productive skills of an employee. Forms and procedures are only tools.[1]

Accordingly, an effective IPA program should be one in which the evaluation of the performance of the subordinate is the product of supervisory practices, and not a process by itself.

The IPA system, according to the behavioralists, should meet two requirements:

> First, involvement of both the superior and subordinate in planning the work and establishing criteria is required. Specific needs must be identified

---

[1]Thomas N. Baylie, et al., "Appraisals of People Resources," in Dale Yoder and Herbert G. Heneman, Jr. (Eds.), Staffing Policies and Strategies (Washington: BNA Books, 1974), pp. 4-171-72.

> and a non-threatening environment provided in which the subordinate may be motivated to change.
>
> Second, a tangible record of performance related to the needs of the organization is required. This record is necessary for information in making personnel decisions, planning its manpower, and providing training and development services.[1]

The mechanics of the participative IPA program usually include two phases: (1) a continuing work-planning and progress review process between the supervisor and the subordinate and, (2) an annual performance review form.

The work-planning and progress review process is an informal ongoing appraisal that is dependent on the work situation, the job, and the individuals involved. The annual performance review form is primarily used to provide a formal record of performance, potential and development needs.

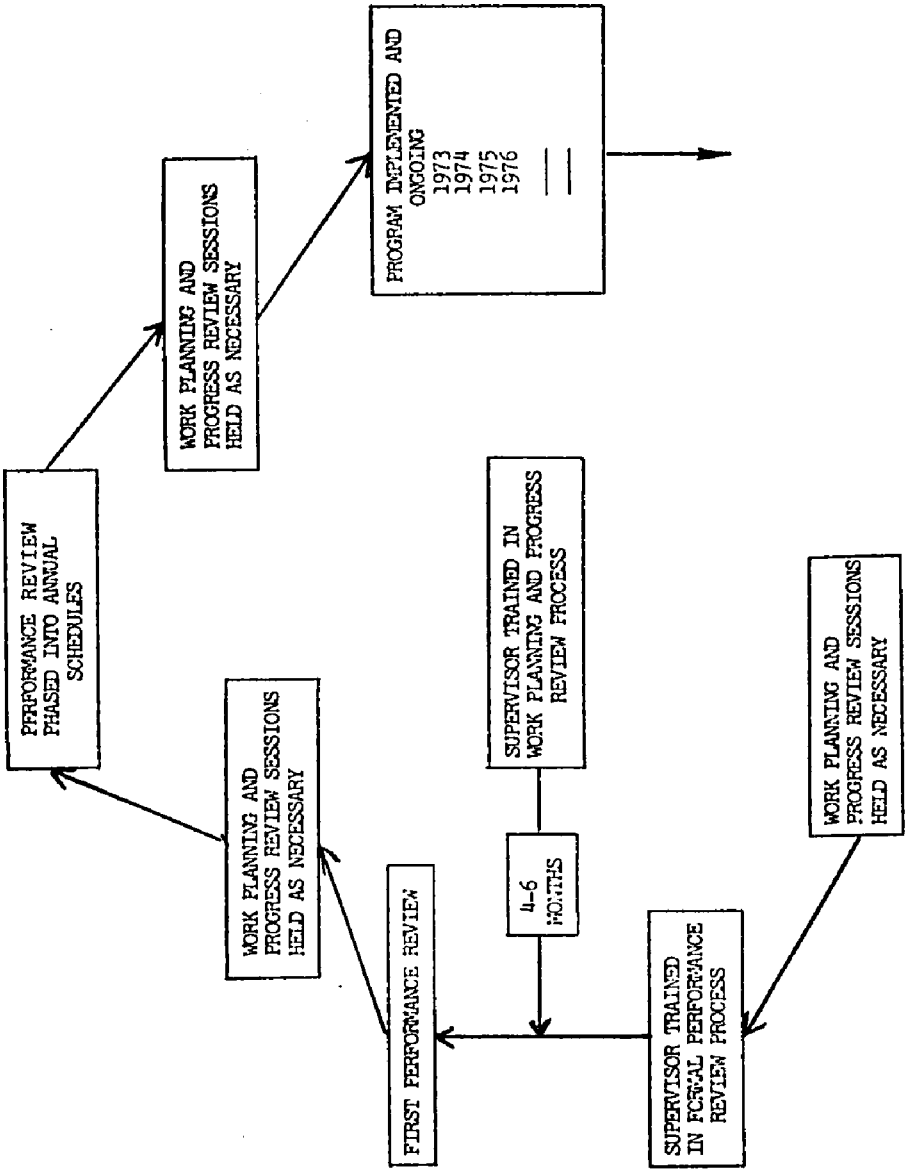
An example of a participative-oriented IPA program is shown in Figure 2.1.

In a direct comparison between the participative, behavioral IPA approach and the traditional, unilateral IPA approach, it

---

[1] Ibid., p. 4-170.

FIGURE 2.1
FLOW CHART OF PARTICIPATIVE-ORIENTED APPRAISAL PROGRAM



[1]Ibid., pp. 4-171. Figure 2.1 is a representation of the Figure presented.

has been noted the participative IPA program provides advantages as indicated; a more practical method for introducing new employees to their work and job environment, an IPA technique that relates more directly to day-to-day performance by the ratee, a basis for continuity of job functions regardless of personnel turnover and a better structure for guiding and developing employees over the long-term.

The results of a comparative study regarding use of trait appraisal (traditional) versus appraisal against specific objectives (participative) IPA programs are presented in Table 2.2.[1]

TABLE 2.2

HOW 141 COMPANIES APPRAISE MANAGERS*

TYPE OF APPRAISAL	PRESENT USE, %			FUTURE EMPHASIS, %		
	NONE	LITTLE	MUCH	LESS	SAME	MORE
TRAIT APPRAISAL	26	28	46	24	63	13
APPRAISAL AGAINST SPECIFIC OBJECTIVES	39	38	23	2	49	49

*NOTE: Columns do not add to 100 percent because some companies use more than one type of program.

[1]W. S. Wickstrom, "Developing Managerial Competence: Changing Concepts - Emerging Practices," National Industrial Conference Board Studies in Personnel Policy, No. 189, 1964, p. 126.

In another study of IPA programs comparing traditional and participative approaches, Blake and Mouton observed differences along three dimensions of superior-subordinate interpersonal relationships.[1] Their research indicated the participative IPA approach tended to produce "positive" results by promoting supervisor and subordinate interaction in seeking solutions to common problems. This increased collaboration resulted in higher levels of job satisfaction by both the supervisor and subordinate. The traditional IPA approach "power" characteristic was also suppressed due to the increased interpersonal relationships.

NEED

For an IPA program to effectively satisfy the needs of an organization, it must be determined what the program is expected to accomplish, what is the anticipated use of the program, and who will participate in the program.

A number of studies have been conducted to quantify the reasons organizations establish IPA programs. According to

---

[1]Robert R. Blake and Jane S. Mouton, "Power, People, and Performance Review," Advanced Management, Vol. 26, 1961, pp. 13-17.

Barrett, the reasons vary, but generally include one of the following:

- To give employees an idea of how they are doing,
- To identify promotable employees or those who should be demoted,
- For salary administration,
- To provide a basis for supervisor-employee communications,
- To help supervisors to know their workers better,
- To identify training needs,
- To identify employees for layoff or recall,
- To assist in placement,
- To validate the selection process and evaluate other personnel activities, such as training programs and psychological tests,
- To improve departmental employee effectiveness,
- To determine special talents,
- To determine the need for disciplinary action,
- To determine progress at the end of probationary periods,

- To furnish inputs to other personnel programs, such as manpower and succession planning, and,
- To supply information for use in grievance interviews.[1]

Although IPA programs should be based on need, to incorporate multiple needs into one program may provide counter-productive results. If multiple needs are trying to be achieved, confusion may result regarding the primary purpose of the IPA. In addition, the participants may become "hostile" toward the program and/or inconsistency of data may result, due to different interpretations of needs by the participants.

Various authors have indicated that IPA programs should be viewed and utilized as a "management tool" to satisfy one specific need. If information or data are required for other administrative or research purposes, supplemental IPA techniques should be developed and implemented to obtain desired information. However, if a multi-purpose IPA program is implemented due to necessity or resource limitations, constant control and monitoring is required to ensure fulfillment of all the desired needs. Multi-purpose IPA programs generally are more controversial and provocative to both the participants and management in the

---

[1]R. S. Barrett, Performance Rating (Chicago: Science Research Associates, 1966), p. 6.

determination of whether the results obtained are optimal.

OBJECTIVES

The objectives of IPA programs within an organization vary with the needs of the organization. One viewpoint concerning the objectives of IPA programs is presented below:

OBJECTIVES OF INDIVIDUAL PERFORMANCE APPRAISAL PROGRAMS

- Help or prod supervisors to observe their subordinates more closely and to do a better coaching job.
- Motivate employees by providing feedback on how they are doing.
- Provide back-up data for management decisions concerning merit increases, transfers, dismissals, and so on.
- Improve organization development by identifying people with promotion potential and pinpointing development needs.
- Establish a research and reference base for personnel decisions.[1]

The aforementioned objectives as written, apply primarily to business or government environments. Gage identified the following three reasons for IPA programs in education:

---

[1]Oberg, op. cit.

- The need for providing a broad base for administrative decision-making,
- A basis for professional improvement and development, and,
- The need for data for further research on teaching and learning.[1]

Miller subsequently added "feedback" and "guidance" as necessary requirements for a successful IPA program within an educational environment.[2]

Therefore, it is noteworthy that, with the exception of terminology differences, the objectives for establishing IPA programs in all environments (i. e., business, governmental, educational) are recognized to be common.

It is assured that, for any organization to be successful, all resources must be utilized and controlled efficiently and effectively. Having noted a commonality of organizational IPA objectives regardless of the environment, it is assumed there is a commonality regarding the primary resources within each organization. The primary resources assumed for this study

---

[1]N. L. Gage, "Ends and Means in Appraising College Teaching," in Richard I. Miller, Evaluating Faculty Performance (Washington: Josey-Bass, Inc., 1972), p. 11.

[2]Richard I. Miller, Evaluating Faculty Performance (Washington: Jossey-Bass, Inc., 1972), p. 11.

include fiscal resources, physical resources, and manpower resources. Typically, most organizations emphasize the administration and control of fiscal and physical resources, often at the expense of manpower resources. However, the management of manpower resources and, more specifically, the quality of management, is believed by many to be the key to the success of any organization. One viewpoint regarding the importance of the management of manpower resources was reported by Miller.

> If a company has nothing going for it except one thing - good management - it will make the grade. If it has everything except good management, it will flop.[1]

Assuming the aforementioned viewpoint is accurate, what can be done to optimize the management of manpower resources within the organization? The following sections of this review will attempt to provide an answer to this question.

PLANNING, IMPLEMENTATION AND UTILIZATION

In planning an IPA program, the categorization of the program must be defined in order to determine the use of the data. According to Barrett, IPA programs are included in one

---

[1] Ibid., p. 48.

of three major categories:

1. Administrative - This involves personnel actions such as salary administration, promotion, transfer or discharge.

2. Performance Improvement - This is used normally for delineating job responsibilities and goals with the supervisor and the subordinate jointly reviewing progress toward achieving these goals and planning together the employee's development.

3. Research - Performance measures are used to determine the value of the procedures utilized in such activities as selection and training.[1]

Early IPA program efforts tended to contain psychological overtones and usually consisted of an appraisal of traits that were considered important.

However, early IPA programs typically did not appraise performance in terms of results based on decisions influenced by the individual; but, were the result of preconceived characteristics, or criteria, that management personnel should possess. Examples of preconceived characteristics included such traits as leadership, initiative, dependability, and ambition. It was apparent that these criteria did not accurately measure the effectiveness of the individual on-the-job, but primarily what

[1]Barrett, op. cit.

was "thought" of the individual rather than what the individual "accomplished."

The major weakness in the early IPA programs tended to be the lack of adequately defined criteria that related to job responsibilities. The emphasis was on personality traits which overshadowed available objective criteria of on-the-job performance. This emphasis on trait criteria also tended to retard open communications between the rater and the ratee because of the difficulty of the rater to verbalize trait judgment of performance to the ratee. With the recognition and increased use of objective criteria, open communications have been noted to improve, primarily due to the identification of quantitative terms and data.

After the need and the categorization of the IPA program have been determined, the organization should decide which members in the organization will participate in the program. Typically, all personnel within an organization having an IPA system, are involved; however, it is important to determine if it is necessary to include every employee.

Regardless of the final selection of participants, identification of what variables to evaluate is also important. Variables selected for the IPA should be related specifically to

the need. Variables, such as loyalty, honesty, and integrity, are presumed by many to be personal quality prerequisites for any position; therefore, they should not be included. Further, if any personality characteristics are selected, the variables should be oriented toward job performance, not the individual as a person.

The variables selected should have enough flexibility to ensure accurate requirements of the different positions to be appraised, the levels of assignment, and operating conditions within the organization.

Numerous studies have been initiated to determine if specific variables are associated with success within an organization. One study, reported by Patton, was conducted within a large industrial organization. The study utilized demonstrably successful managerial personnel in an effort to determine if any common educational, economic, or social background variables were attributable to their successes.

The research included investigation of early family life, financial and social status, extent of education, including courses studied and grades received, and early work experience. The results of the survey indicated that during the formative years, these individuals ". . .tended to be as diverse, as America

itself."[1]

One of the significant contributions of this research was that it is apparently not possible to prejudge the capacity or potential of an individual with any certainty; therefore, to control manpower potential, effort should be focused on developing IPA techniques to judge current on-the-job performance. Although a review of past performance and background data may be helpful, present performance apparently is the most reliable determinant to future performance. The ability of management to accurately judge the performance of an individual is vital to the achievement and continuing success of the organization. In addition, the research results reinforced the notion that appraisal of performance should be a continuing process because individuals tend to achieve at different rates, at different times, and at different levels of competency.

Historically, the trend of initiating and maintaining formal IPA practices of managerial personnel has been on the increase, particularly in lower levels of management. This trend is shown in Table 2.3.

---

[1]Arch Patton, "How to Appraise Executive Performance," in Herbert J. Chruden and Arthur W. Sherman, Jr., Readings in Personnel Management (Chicago: South-Western Publishing Company, 1966), p. 188.

TABLE 2.3

PREVALENCE OF APPRAISAL PROGRAMS AT VARIOUS ORGANIZATIONAL LEVELS[1]

	NUMBER
Total Number of Companies Surveyed	567
Number Engaging in Formal Appraisal	357
Number Rating both Supervisors and Executives	243
Number Rating Executives Only	14
Number Which Have Discontinued a Program of Rating Executives	256
Number Rating Supervisors Only	100

The following trend chart, adapted from Scott and reported by Campbell, indicates the increased use of formal rating systems by organizations in the business environment.

In a more recent study, Wickstrom reported that the National Industrial Conference Board confirmed that 84% of 167 firms surveyed indicated they appraised management-level personnel to some extent.[2]

---

[1]W. R. Spriegel, "Company Practices in Appraisal of Managerial Performance," Personnel, 39 (1962), p. 79.

[2]Campbell, et al., op.cit. Refers to NICB study reported by W. S. Wickstrom, National Industrial Conference Board Studies in Personnel Policy, No. 189, 1964.

TABLE 2.4

SURVEY OF RESULTS: USE OF FORMAL RATING IN BUSINESS FIRMS[1]

YEAR	NUMBER OF COMPANIES IN SURVEY	PERCENTAGE OF RESPONDING COMPANIES ANSWERING "YES" ON APPRAISAL OF RANK AND FILE EMPLOYEES	PERCENTAGE OF RESPONDING COMPANIES ANSWERING "YES" ON APPRAISAL OF MANAGERS AND EXECUTIVES
1930	195	*	*
1940	231	95.3	45.6
1947	325	80.8	43.7
1953	628	77.5	47.5
1957	852	77.0	58.0

*Available evidence shows that only 41% of responding companies used rating scales.

Campbell also reported that approximately 80% of the firms sampled had formal appraisal of their managers.[2]

In a survey conducted in education by Gaff and associates and reported by Miller, the formal appraisal of educators indicated:

---

[1]John P. Campbell, et al., Managerial Behavior, Performance and Effectiveness (St. Louis: McGraw-Hill Book Company, 1970), p. 63. Refers to survey adapted from W. D. Scott, et al., Personnel Management (New York: McGraw-Hill Book Company, Inc., 1954).

[2]Campbell, et al., op. cit.

> Seventy-two per cent of the faculty said they favored a formal procedure to evaluate teaching. Eighty-two per cent of those in favor felt that students should be involved in the evaluation, 76 per cent felt that colleagues and 73 per cent felt that departmental chairmen should also be involved. . . .In [a] 1970 survey, 85 per cent of the respondents endorsed the idea that a formal program of teacher evaluation of faculty should be 'used by the college in making decisions about such matters as salary, promotion, and tenure.'[1]

Another important consideration in any IPA program is the determination of who should implement the IPA. The basic issues involve identifying and isolating the most reliable and valid criteria concerning the position being evaluated and selecting the rater who can provide the most objective judgment, given this information.

In most organizations, the individual generally selected as the rater is in a direct supervisory position, or is the immediate supervisor, to the individual to be rated. This rater selection is commonly utilized primarily because the immediate supervisor is generally accepted as the "best qualified judge" of a subordinate's performance within the leader-subordinate

---

[1]Miller, op. cit., p. 7.

hierarchy. However, the peer or colleague of the rater is considered the best evaluator of on-the-job performance of an individual.[1]

The acceptance of the immediate supervisor as the "best qualified judge" is because, traditionally, this supervisor is more familiar with his subordinate's work than any other person in the organizational hierarchy, has a better overall knowledge and understanding of organizational goals than other raters (e. g., peer, colleague), and maintains frequent contact with the subordinate in the work environment.

Some organizations use two or more raters in preparation of the IPA. In general, however, the multirater IPA method is not utilized by organizations based on the belief that raters other than direct supervision do not have sufficient direct contact with the individual in the position being rated, or the time and/or experience in using multiple raters is excessive for the resultant information causing a "negative" cost-benefit situation.

---

[1]Edward E. Lawler, III, "The Multitrait-Multirater Approach for Measuring Managerial Job Performance," Journal of Applied Psychology, 1967, 51, p. 370.

Whenever rater bias is suspected, or certain raters are using higher rating standards than other raters, or comparability of ratings is essential, a multirater technique may be the optimum methodology to utilize. Typical multirater IPA techniques include the field review and rating committee techniques.

The field review is generally conducted by a member of an administrative staff with small groups of raters from each supervisory unit. The staff member reviews each rating category or variable with the raters to: "(a) identify areas of inter-rater disagreement, (b) help the group arrive at a consensus, and (c) determine that each rater conceives the standards similarly."[1]

This group-judgment technique depersonalizes the final rating awarded while concurrently making the technique explicit and delibarate. The resultant ratings are generally more reliable and valid than ratings obtained from only one rater and, this technique permits staff organizations to readily recognize rating biases. However, the excessive time and expense drawbacks of the field review process make the utilization of this system attractive primarily as a final precaution or "check."

Another group-judgment technique in the preparation of the IPA is the rating committee. Although a consensus appraisal is

---

[1]Oberg, op. cit., p. 63.

the resultant, difficulty in obtaining informed raters is believed to be detrimental to validity. This method is also expensive due to multiple manpower commitments and, therefore, is usually used only for the IPA of higher echelons of management.

Because the judgment is a group effort, no one individual is responsible which could be perceived as either a "positive" or a "negative" asset. This method continues to be used in some organizations because of the belief that, with the addition of each rater, greater validity is obtained. The rationale for this belief is based on the thought that the more information available and contributed, the better the judgment.

Another source of multirater IPA information could be supplemental appraisals obtained from peers/colleagues of the ratee, subordinates of the ratee, or a self-appraisal by the ratee. The peer/colleague IPA methodology is considered by many to be the most valid means of obtaining information, primarily because of the close contact of the rater and ratee. This method of IPA is rarely utilized, however, because the peer/colleague is not considered to be as familiar with the overall goals of the organization as raters from supervisory positions. Also, anonymity, which is a prime requisite of this technique, is often difficult to maintain.

Currently, in business and government, subordinate ratings are utilized more for organizational interest than for decision-making or developmental reasons. In education, however, the student evaluation of faculty teaching has grown in popularity. In a dated survey, systematic student ratings of teacher effectiveness was used in all or most departments within the sample 12.4% of the time. Comparative sources of information are shown in Table 2.5.

An observation by Austin and Lee (1967) concerning the systematic student ratings is relevant to any organization:

> If the ultimate measure of the teacher's effectiveness is his (or her) impact on the student - a view which few educators would dispute - it is unfortunate that those sources of information most likely to yield information about this influence are least likely to be used.[1]

One other alternative in the selection of a rater, excluding the possibility of obtaining assistance external of the organization, is self-appraisal. This IPA technique is not normally utilized; but when used, it is normally requested to supplement other techniques of rating.

---

[1] Ibid., p. 24.

TABLE 2.5

FREQUENCY OF USE OF VARIOUS SOURCES OF INFORMATION IN THE EVALUATION OF TEACHING EFFECTIVENESS[1]

SOURCE OF INFORMATION	USED IN ALL OR MOST DEPARTMENTS %	NOT USED %
CHAIRMAN EVALUATION	85.1	3.4
DEAN EVALUATION	82.3	5.8
COLLEAGUE'S OPINIONS	48.9	8.7
SCHOLARLY RESEARCH AND PUBLICATIONS	43.8	21.6
INFORMAL STUDENT OPINIONS	41.2	9.6
GRADE DISTRIBUTIONS	28.0	37.6
COURSE SYLLABI AND EXAMINATIONS	26.4	28.0
COMMITTEE EVALUATION	25.1	52.4
STUDENT EXAMINATION PERFORMANCE	19.6	35.8
SELF-EVALUATION OR REPORT	16.3	57.2
CLASSROOM VISITS	14.0	39.5
SYSTEMATIC STUDENT RATINGS	12.4	47.6
ENROLLMENT IN ELECTIVE COURSES	11.0	49.9
LONG-TERM FOLLOW-UP OF STUDENTS	10.2	47.1
ALUMNI OPINIONS	9.9	46.8

[1]Miller, op. cit., p. 23.

For any type of system to be successful, support at all levels of the organization is mandatory. This also applies to any IPA program. The following list of general questions, directed toward the IPA, should be reviewed by each organization. The exclusion or improper attention to any question could significantly alter the success of the entire IPA program.

1. Are resources such as fiscal, physical and manpower available?

2. Will the inter- and intra-organizational environment support the program?

3. Do the participants understand the need for the IPA program?

4. Does the program relate to the specific needs of the individual and the organization?

5. Does the program have the time and effort commitment of participants as well as top management?

6. Will the information obtained be effectively utilized?

Assuming a supervisor/subordinate relationship, the necessity for any IPA program must be known and accepted by the ratee as well as the rater. Based on the understanding an IPA exists within an organization, the reaction to this program by the individual may vary as highlighted below:

1. The subordinate's desire to be evaluated varies in strength both within each individual and from individual to individual over time.

   - An individual is more ready to receive an evaluation from another when he is uncertain about his performance and suspects that it is poor.

   - An individual who believes his performance is good is less concerned about the opinions of others regarding his performance.

   - An individual who is success oriented in his performance is more interested in feedback than one who fears failure.

2. The anticipation of being evaluated is threatening. Hostile and untrusting attitudes are often directed toward performance appraisals and appraisers.

   - A subordinate views an evaluation by a superior as an occasion when he lacks control of his fate, and thus is apprehensive and anxious.

- Individuals with low self-esteem are more apt to be threatened than those with high self-esteem.

- Negative reactions to criticism are greater where those criticized received support from one another than when such support is not available.

3. The more an individual has the ability to influence the evaluator, the more likely he is to accept that evaluation as his own evaluation of himself.

   - The relationship between a supervisor and his subordinate determines whether the latter accepts the former's ratings.

   - An expert appraiser's opinions are more likely to be accepted than those of a nonexpert.

   - A subordinate is more likely to accept the ratings of a supervisor he likes and respects.

   - An individual is more likely to accept evaluations which are similar to those he has received from other appraisers.[1]

Historically, in too many instances, the careful preparation of an IPA has been nullified by improper, or faulty, management of the data. This improper, or faulty, management may result from any or a combination of the following: poor communications or faulty communication networks, users of the data having diverse goals and objectives from the organization, or a belief the system

---

[1]Baylie, op. cit., p. 4-179.

is a destructive (negative) rather than a constructive (positive) imposition.

McGregor noted several causes for resistance to IPA programs which directly involve the rater:

- A normal dislike of criticizing a subordinate (and perhaps having to argue about it).
- Lack of skill needed to handle the interviews.
- Dislike of a new procedure with its accompanying changes in ways of operating.
- Mistrust of the validity of the appraisal instrument.[1]

The utilization of the IPA program may significantly contribute to the success or failure of an organization. In addition, when the IPA program is combined with other performance indices, the IPA may significantly aid in developing an efficient and effective workforce as well as conserve human resources.

Several of the major uses of the IPA information are

---

[1]McGregor in Whisler and Harper, op. cit.

outlined below:

MAJOR USES OF IPA INFORMATION

1. Employee Development
   - Individual Training and Development
   - Manpower Planning

2. Administrative Decision-Making
   - Placement
   - Salary Administration

This notion is substantiated by survey results reported in Table 2.6 by Spriegel.[1]

TABLE 2.6

PURPOSES FOR WHICH APPRAISAL IS USED

PURPOSE	NUMBER OF FIRMS
COUNSELING	300
PROMOTION	298
TRAINING AND DEVELOPMENT	265
CONSIDERING RETENTION OR DISCHARGE	240
SALARY ADMINISTRATION OR MERIT INCREASE	237
BONUS PAYMENTS	54
PROFIT-SHARING PAYMENTS	14

[1]Spriegel, op. cit., p. 82.

According to this survey, the purposes for which the IPA is used stressed developmental, rather than decision-making, processes.

The IPA instrument may include information concerning the specific training or development planned for the individual. This information is provided to assist in improving the effectiveness of the individual on-the-job and to aid in career planning.

Typical developmental activities that may be recommended for managerial personnel, which are both work and leisure related, include participation in:

1. Business or Professional Organizations
2. Community or Civic Activities
3. Special Projects or Task Forces
4. Technical or Supervisory Courses
5. Management Conferences or Seminars
6. Professional Development Programs
7. College or Extension Courses

For manpower planning purposes, the IPA has the potential to assist in designation of the most capable individual, whenever and wherever needed. Input into the manpower planning process could include identification of strengths and weaknesses, identification of potential individuals to fill specific positions,

as well as identification of individuals that may need reassignment or retraining to be more effective.

Information from the IPA could provide valuable input in the selection, placement and retention processes within an organization. While successful performance of an individual in the present position does not assure an effective performance in a higher level position, criteria may be included in the IPA report that provides accurate predictive information; such as, results achieved, highlights of performance, overall evaluation of performance and potential. An advantage for utilization of the IPA for this type of data is that the available information is applicable not only to candidates for job placement, but also applicable in the search for candidates to fill a job vacancy.

The IPA program could also be developed and administered specifically for salary administration or incentive-bonus purposes.

In practice, IPA systems may be established for either performance improvement and employee development. A judicial role exists in administrative-oriented programs and a counseling, or coaching role, exists with developmental oriented appraisal. It is evident, therefore, a dicotomy would exist in attempting to

utilize one appraisal for both administrative and development purposes.

Another consideration in IPA programs includes the proper utilization of the data. The system must be manageable in terms of time required for processing, economical in terms of time expended and the results achieved, and cost justified for the effort expended.

The confidentiality of the IPA program information must be maintained. This includes physically separating the IPA program information from the personnel file or record of the individual.

Ideally, the IPA form or instrument should be retained only by the rater and/or coordinator of the program. This confidentiality is necessary to eliminate the possibility of compromising the integrity of either the counseling process or the administrative process.

Unfortunately, some organizations have utilized the IPA program as the exclusive factor for either developmental or decision-making processes. Sole reliance on the IPA may result in biased decision-making, or developmental processes, especially in areas involving the hiring, promotion, training and retention of individuals.

In a survey of thirty-three organizations, Campbell reported the following impressions regarding IPA procedures:

1. There appears to be a strong and growing tendency for firms to adopt a result oriented approach (although many firms also ask for global or trait ratings at the same time). In two instances where the nature of the technology makes managerial performance a highly interdependent process or where personnel change jobs and assignments rapidly, firms had been forced to discontinue management by objectives approaches because individual performance could not be identified.

2. A few firms have begun to experiment with peer ratings. Many of the management by objectives companies, of course, employ some measure of self-rating.

3. Firms who have developed manager replacement tables, although they may have a systematic and formal appraisal system, tend to use committees to determine lines of succession. These committees may or may not consider performance appraisals in making their decisions, and apparently they never use appraisals as the sole criterion.

4. At higher levels in the organization, appraisal is more likely to be performed by a committee, rather than by the superior alone.

5. A small minority of companies still do not have formal appraisal systems.

6. In large, diversified organizations, widely different systems may be used in various divisions.

7. Many companies have two or more appraisal systems operating concurrently. For example, a manager's budget performance may be appraised to determine whether he gets a bonus, while a global personality evaluation may be made to isolate his developmental needs.

8. A number of companies are beginning to feel that the traditional, relatively narrow ranges of wage and salary administration defeat whatever incentive value might be obtained from tying managerial pay to performance appraisal. On the other hand, those interviewees who stated that their firms did not have a systematic plan of wage and salary administration tended to believe that managers were relatively arbitrary in distributing increases--and that this also destroyed the incentive effects of pay.[1]

RATENG SYSTEMS

Ratings are formal, written appraisals assessing or recording judgments and observations related to the past, present, or future behavior and performance of an individual, usually in a work enrivonment. Typically, ratings are used either to appraise subordinates or to make comparisons among subordinates. In either situation, the appraisal is directed toward specific attributes of behavior (e. g., quantity or quality of performance) or a combination of attributes (e. g., overall performance or potential).

---

[1]Campbell, et al., op. cit., 65-66.

According to Wherry, a rating is ". . .a summing up of the totality of pertinent past observations of the appropriate behavior of the ratee. In addition, if the ratings are to reflect the true value of the ratee, the behavior observed must have been an adequate reflection of the ratee's behavior."[1]

Wherry also proposed a number of theorems regarding the rating process. These theorems are outlined below.

1. Rating scales or items that refer directly to tasks maximally controlled by the ratee rather than by the work situation will lead to more accurate rating.

2. Raters will vary in the accuracy of ratings given in direct proportion to the number and relevance of their previous contacts with the ratee.

3. Rating-scale items which refer to easily observed, frequently performed behavior categories will result in more accurate ratings than will those which refer to hard-to-observe behaviors.

4. The rater will make more accurate ratings when he has foreknowledge of the types of activity to be rated, since he will be better able to focus his attention on pertinent behavior.

5. The rater will be better able to focus his attention if he is furnished an easily accessible list of objective cues for the evaluation performance.

---

[1]R. J. Wherry, The Control of Bias in Ratings, VII, A Theory of Rating (PRB Report No. 922, Department of the Army, Personnel Research Branch, 29 February, 1952), in Baylie, p. 4-175.

6. Rating scales with physical features which facilitate recall of the actual observation of the ratee's performance will improve the accuracy of ratings.

   - Longer objective, descriptive statements will be more effective than single-value words or simple phrases in defining steps on an adjective scale.

   - The clearer and more unambiguous the scale, the more likely that the rater's attention will be centered upon the desired aspects of performance.

7. Objective recall will be improved when the rater keeps a continuing, written record of critical incidents of the performance he has specifically observed The function of such a record is twofold:

   - It gives concrete reinforcement at the time the performance is observed.

   - When reviewed just prior to completing the rating scale, it revives and strengthens pertinent aspects of the rater's recall.

8. When the rater believes the rating will have an immediate effect on the ratee, he is apt to be less accurate in his rating. Ratings obtained in advance, through a routine process, will be more accurate than those secured specifically at the time an administrative action (e. g., promotion) is contemplated.

9. Knowledge that the rating given will have to be justified may serve unconsciously to affect its accuracy.

   - If it is to be justified to the ratee, the rater is apt to recall more favorable perceptions.

- If it is to be justified to the rater's superior, the rater is apt to recall observations of interest to the superior regardless of their appropriateness.

- The above distortions can be offset by requiring that ratings be reviewed with both the ratee and the rater's superior.

10. Forgetting is largely a function of activities which occur between observation (learning) and recall. Ratings secured soon after the observation of performance will be more accurate than those secured after a considerable lapse of time.

    - A rating should be secured immediately when the work relationship of the rater or ratee changes through transfer or promotion.[1]

The following questions are pertinent in the observing and recording of rating systems.

1. What behaviors should be observed and recorded?

2. What are the psychometric characteristics of an adequate measure of job performance?

3. What problems or sources of error are often encountered in observing and recording job behavior and how may they be overcome?

---

[1] Ibid., p. 4-178.

4. What kinds of instruments--rating forms, checklists, etc.--may be used to record job beahvior observations and how well do they satisfy the requirements of an adequate measure?

5 How do the records or measurements of job behavior relate to other important organizational consequences, and what relevance do such relationships have for personnel selection and placement programs?[1]

The performance or behavior that should be observed and recorded should be relevant to getting the job accomplished properly and the ultimate objectives of the program. Behaviors should be identified through involvement and sampling by both raters and ratees.

The criteria of job performance utilized in the design of the instrument should be consistent and representative of the position. The entire range of behavior essential to the position should be included in the instrument, and, conversely, behavior not directly related to the performance should be eliminated to prevent possible contamination.

Potential errors that can exist in the instrument are numerous and varied. The fundamental source of error is the

---

[1]M. D. Dunnette, Personnel Selection and Placement (Belmont, Cal.: Wadsworth Publishing Co., 1966), pp. 85-102.

lack of participation by the rater in the design of the instrument, which often results in a lack of understanding or commitment to the task by the participants. Errors may also occur when an inadequate sampling of the job-behavior domain exists, when there are changes in the job environment, or as the result of the general variations in day-to-day performance.

Validity of the IPA is affected by two factors--one, whether the rater adheres to the principles and mechanics of the rating program, and two, whether the ability to rate accurately is impeded. The first factor may be influenced by several aspects, such as, whether the rater is unwilling or unable to adhere to the principles or whether the rater can identify with the ratee. The second factor involves whether ambiguity of the criteria exists, whether observation constraints are present, or whether personal bias is involved.

Through effort in planning, contamination and errors in IPA rating results can be minimized or eliminated. Recognition of the influences resulting from environmental constraints, the impact on performance resulting from actions external to the span of control of the individual, and an understanding between criteria evaluated and projected use of the results, all assist in minimizing contamination and errors.

APPRAISAL TECHNIQUES

In any organization, the various appraisal techniques and instruments produce unique combinations of strengths and weaknesses, depending primarily on the skills of the raters, the performance areas and individuals to be appraised, and the organizational environment.

In the selection of an IPA technique, an important consideration is whether the technique is compatible with organizational needs.

Several of the more commonly used IPA techniques are listed below:

1. Rating Scales
   - Graphic Rating Scale
   - Forced-Choice Rating Scale
   - Behavioral - Anchored
2. Essay
3. Critical-Incident
4. Ranking Methods
5. Assessment Centers

RATING SCALES

Rating scales comprise the most widely used format in the design of IPA instruments. Of the numerous rating techniques, three of the more widely used rating systems are reviewed - graphic rating scales, forced-choice rating scale, and behavioral-anchored rating scales.

GRAPHIC RATING SCALE TECHNIQUES

Linear or graphic rating scales probably are the most commonly used format in the IPA. Although appearing in many forms, graphic scales generally indicate relative strengths and weaknesses either by individual criteria or by overall performance. This technique typically requires the rater to record judgments by checking or circling a segment on a predetermined scale.

This technique usually is utilized when quality or quantity of performance is desired with the performance scale choice ranging from "outstanding" or maximum, to "unsatisfactory" or minimum. In comparison to other techniques, Oberg reported that in a classic comparison between the graphic scale and the forced-choice technique, the graphic scale was as valid as the best of

the forced-choice forms.[1] In addition, the graphic scale is generally less expensive to develop and is more acceptable to raters than the forced-choice technique. Although the graphic scale may not yield the depth of an essay appraisal, it is usually more consistent and reliable.

Graphic rating scales do have several noteworthy weaknesses. First, they are subject to error in observation and rating. Some of the more common errors in graphic rating scales are outlined below.

Reliability - results from a variation in the understanding of definitions.

Halo - results from raters judging individual aspects of performance in terms of the overall impression.

Systematic - results reflect variation in personal standards and expectations.

Harsh - results reflect lower ratings to selected individuals.

---

[1]Oberg, op. cit., Refers to Study by R. W. Highland and J. R. Berkshire, "Forced-Choice Performance Rating: A Methodological Study," Personnel Psychology (Autumn, 1953), p. 355.

Lenient - results reflect higher ratings to selected individuals.

Contamination- results when criteria are included that have little or no bearing on the performance.

Deficiency - results when critical criteria are omitted from the rating.

Central Tendency - results when judgments are near the mid-point of the scale.

Logical - results when rater evaluates one or more criteria based on the rating of another criteria.

Miscellaneous Biases - results due to influence caused by length of service, sex, race, religion, position in the organization, friendship, quick guesses, first impressions, setting, appearance, personal preferences, frame of reference (projection).

Second, they are typically after-the-fact appraisals of performance. Because the graphic scales are often after-the-fact, the individual has little, or no, opportunity to participate in setting performance standards or to change or improve performance through feedback.

Third, graphic scales, by design, are subject to various types of numeric scale operations. Realistic and appropriate scales should be used to ensure objectivity and to minimize inappropriate comparisons associated with mathematical and/or mechanical manipulation of the information and data received.

Examples of the graphic rating scale method are presented in Table 2.7, and Appendices C, D, and E.

FORCED-CHOICE RATING TECHNIQUES

The forced-choice rating technique was developed by the U. S. Army during World War II.[1] In this technique, raters describe, without evaluation, the behavior of an individual by selecting one or more statements from a group of statements, each of which is designed to reflect equal desirability and acceptability to the ratee.

[1] E. C. Sisson, "Forced Choice: The New Army Rating," Personnel Psychology, 1948, p. 365.

TABLE 2.7

GRAPHIC RATING SCALE

EXAMPLE

INTERPERSONAL RELATIONS SKILL - Treating Employee with Respect and Concern

INSTRUCTIONS:

Consider the Interpersonal Relations Skill of typical or average employee observed in the past. In relation to these employees, what level of Interpersonal Relations Skill does this employee have (circle one)?

A.

NA	9	8 7	6 5 4	3 2	1
	OUTSTANDING	EXCELLENT	SATISFACTORY	POOR	UNSATISFACTORY

B.

CANNOT RATE	MORE THAN MOST	A LITTLE MORE THAN MOST	ABOUT THE SAME AS MOST	A LITTLE LESS THAN MOST	LESS THAN MOST

C. CHECK MARK (✓) ONE

☐	☐	☐	☐	☐	☐
CANNOT RATE	WELL ABOVE AVERAGE	ABOVE AVERAGE	AVERAGE	BELOW AVERAGE	WELL BELOW AVERAGE

D. Indicate level of Interpersonal Relations Skill by indicating a score utilizing the following scale:

1-4 = Poor-Well Below Average

5-12 = Average

13-16 = Good-Well Above Average

_____SCORE

The responses are scored using empirically established weights and, by definition, individuals with high scores are better than those with lower scores.

Although in theory this is an excellent rating technique, several drawbacks have diminished the usefulness of this technique for IPA programs. These drawbacks include:

1. Validity is no greater than obtained from less costly graphic scales.

2. Forced decision-making may be rejected by the rater.

3. Instruments are typically complicated and costly to develop because commonality is limited.

4. Feedback value is questionable.

5. Scoring key secrecy is difficult.

An example of the forced-choice rating scale is outlined in Table 2.8.

TABLE 2.8

FORCED-CHOICE RATING SCALE

EXAMPLE

INTERPERSONAL RELATIONS SKILL

INSTRUCTIONS:

Rate the following statements to describe the manner this employee performs this skill on the job.

SCALE:

1 = Most Descriptive

2 = Descriptive

3 = Least Descriptive

RATING	STATEMENT
_____	Popular with employees
_____	Praises in public, reprimands in private
_____	Provides a positive example

BEHAVIORAL-ANCHORED
RATING TECHNIQUE

This technique was initially suggested by Smith and Kendall,[1] and incorporates the critical-incident approach developed by Flanagan.[2] The basis of this technique requires raters to participate directly in the development of the instrument. This permits the rater to consider the requirements for effective performance in a specific position and to define anchors for specific performance scales in job-behavior-related terms.

Advantages to the behavioral-anchored rating scale are numerous. Typically, the scales are developed by the personnel who use the scale; therefore, the examples and phraseology are understandable and valid to the rater. Contamination and errors associated with other techniques are usually not present as the scales and the techniques tend to reveal behaviors associated with success.

As a participative technique, a beneficial learning experience exists during development of the instrument. The scale is typically a long-term indicator of behavior as well as

---

[1]Patricia C. Smith and L. M. Kendall, "Retranslation of Expectations: An Approach to the Construction of Unambiguous Anchors for Rating Scales," Journal of Applied Psychology, 1963, 47, pp. 149-55.

[2]John C. Flanagan, "The Critical-Incident Technique," Psychological Bulletin, 1954, 51, pp. 327-55.

a technique to obtain feedback regarding job performance and expectations.

The primary drawbacks to this technique are the considerable time and manpower commitments required to develop and maintain the instrument. In addition, separate instruments must be developed for each specific job creating complexity and cost constraints.

An example of this rating technique is presented in Table 2.9.

ESSAY APPRAISAL TECHNIQUE

Essay appraisals are narrative form written judgments, with the rater describing his observations and opinions. Because of the informal nature of this technique, it is questionable whether this technique is as valid as more formal and complex IPA techniques. However, the primary disadvantages to the essay appraisal are the variability in length and content as well as the difficulty in combining or comparing.

An example of an essay appraisal technique question is presented in Table 2.10.

TABLE 2.9

BEHAVIORAL-ANCHORED RATING SCALE

EXAMPLE

INTERPERSONAL RELATIONS SKILL - Treating employees with respect and concern.

INSTRUCTIONS:

Indicate by a checkmark (✓) on the scale where this employee performs regarding Interpersonal Relation Skills.

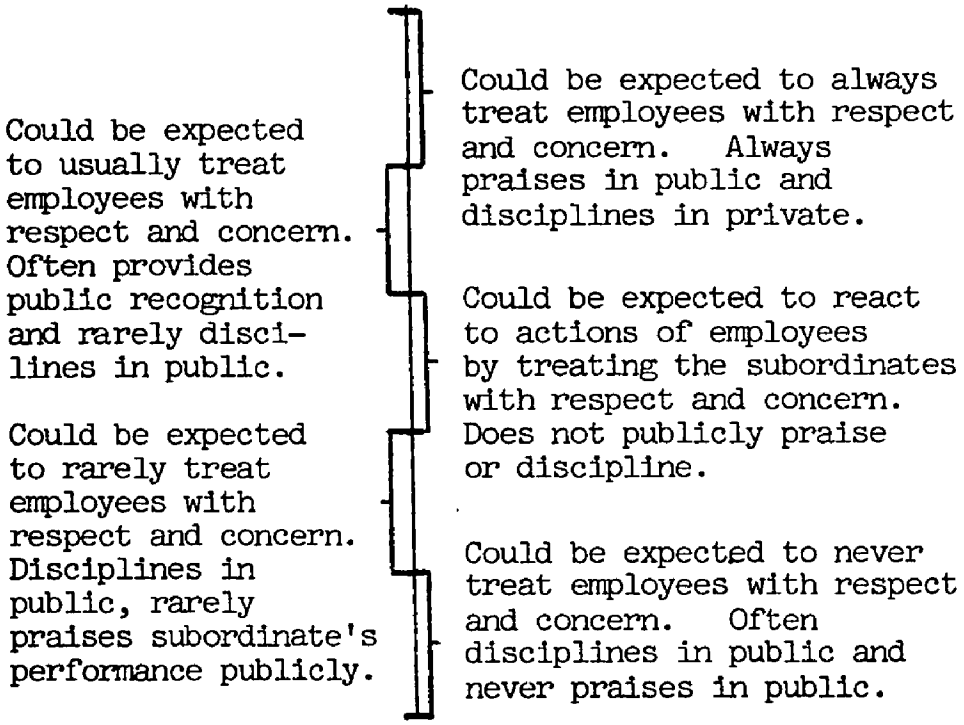


TABLE 2.10

ESSAY APPRAISAL

EXAMPLE

INTERPERSONAL RELATIONS SKILL

INSTRUCTIONS:

Describe in your own words, the Interpersonal Relation Skills of the employee. Include the strengths and weaknesses of the employee regarding treating others with respect and concern.

CRITICAL-INCIDENT TECHNIQUE

Flanagan was instrumental in introducing the development of a job-performance criteria technique initiated by U. S. Air Force research during World War II.[1]

This technique requires the rater to record, as observed, "incidents" of job behavior that are "critical" examples of either satisfactory or unsatisfactory performance. The basis of the technique is to provide, over time, an IPA as well as a media for feedback.

This technique is very useful as a personnel research instrument, particularly as a developmental technique. The technique forces raters to observe ratees more closely while concurrently permitting the rater to gain a better knowledge of his own performance standards and biases.

As an IPA technique, the critical-incident method has the following drawbacks.

1. The ratee can rarely influence the standards used to judge performance.

---

[1] Ibid., p. 329.

2. Feedback is after-the-fact.

3. The recording of the "incidents" is a tedious task.

A sample of an incident reporting instrument is shown in Appendix A.

RANKING METHODS

For comparisons of individuals within and between work groups, a ranking technique is advantageous. Two of the more common ranking techniques include alternation ranking and paired-comparison ranking. Although ranking techniques reportedly have limited reliability and validity, they are reliable, practical and easy to administer.

ALTERNATION RANKING TECHNIQUE

In alternation ranking, a listing of individuals to be ranked is provided to the rater. The rater first selects the "most valuable" individual with respect to characteristics being considered. The name of this individual is placed at the top of a new list. The next selection is the "least valuable" individual, whose name is entered at the bottom of the new list.

Alternation between "next highest" and "next lowest" continues until the initial list is exhausted.

The primary advantage of this technique is the ease in statistically comparing and combining similarly ranked groups. The limitations include assumption of equal "distance" between all ratees, ignorance of individual and situational variations, and forced classification of subordinates.

An example of the alternation ranking method is provided in Appendix B.

PAIRED-COMPARISON RANKING TECHNIQUE

The paired-comparison is a psychometric technique. Each individual being appraised is compared to every other individual, with one individual being judged better than the other. Ranking is provided by the summation of the scores established through counting the number of times the individual has been selected above the other individual.

Limitations are similar to the alternation method, except considerable time and effort is involved. For example, ranking

of ten individuals would require forty-five paired-comparisons.

$$\text{NUMBER OF COMPARISONS (C)} = \frac{N\ (N - 1)}{2}$$

Where N = Number of Individuals

An example of this technique is provided in Table 2.11.

GROUP ASSESSMENT TECHNIQUE

During World War II, assessment centers were developed and used primarily by the Armed Forces to test aptitude for special assignments. The concept has found wide acceptance, especially during the past decade, as more than 300 U. S. corporations (e. g., A.T.& T., General Electric, General Motors, Standard Oil of New Jersey, Sears, IBM) now have established assessment center programs.

The group assessment technique is commonly practiced in an assessment center. Assessment centers are designed to provide appraisals on a number of variables or criteria based upon a candidate's performance in a series of observed exercises. Each exercise is carefully designed to simulate conditions which

TABLE 2.11

PAIRED-COMPARISON RANKING

EXAMPLE

INTERPERSONAL RELATIONS SKILL

INSTRUCTIONS:

In comparison to one other, which individual ranks the highest with respect to this skill. (PLACE LETTER OF INDIVIDUAL RANKING HIGHEST IN COMPARISON BOX.)

INDIVIDUALS	COMPARISON COMBINATIONS A	B	C	D	E	HIGHEST RANKING TOTAL	OVERALL RANK
A		☐	☐	☐	☐	___	___
B			☐	☐	☐	___	___
C				☐	☐	___	___
D					☐	___	___
E						___	___

the individual might be expected to encounter on the job. The rationale for this technique is that the more nearly the exercise duplicates the real job situation, the more valid are the observations as predictors of a person's ability to do the job.

The exercises are observed and evaluated by trained assessors. These assessors closely observe the behavior of the candidate to see if he/she can effectively react to the problems that are presented. Because there are no "right or wrong" answers, the assessment is whether or not a candidate is logical and effective in his actions.

A trained assessor observes and assesses each participant on a number of performance factors, utilizing a carefully constructed evaluation form. The evaluation form and assessor training are both directed at specific indicators in the candidate's performance. The observations are made during and immediately after each exercise, and, in most instances, more than one assessor observes each candidate during each exercise. A number of assessible dimensions are deliberately repeated in several exercises to guarantee the highest validity and reliability possible.

An example of dimensions assessed are indicated in Table 2.12.

TABLE 2.12

ASSESSMENT CENTER DIMENSIONS ASSESSED[1]

CONFLICT	Ability to deal with the various types of conflicts faced in an organization in an effective fashion. 1. conflict between hourly workers; 2. conflict between peers; 3. conflict with hourly workers.
PRESSURE	Ability to deal with the various pressures faced in the organization. 1. pressure from higher management; 2. time pressure; 3. interpersonal pressure.
ORGANIZATION AND PLANNING	Handling the administrative details of the position. 1. scheduling work; 2. record keeping (time sheets); budget sheets, transfer sheets.
PERCEPTION	Ability to perceive the critical elements of a situation. 1. sensitivity to people; 2. sensitivity to situational constraints.

TABLE 2.12 (CONTINUED)

HUMAN RELATIONS	Ability to handle people in a fashion designed to elicit cooperation.
DECISION MAKING	Ability to make reasonable decisions. 1. decisions concerning people; 2. decisions concerning production.
ORAL COMMUNICATION	Ability to communicate effectively.

[1]Taken from format developed by the Research Staff of the "Firm."

Exercises that may be included in a managerial assessment center include the following:

1. An in-basket exercise designed to evaluate organizational and planning abilities, interpersonal skills, and thoroughness.

2. A scheduling problem involving assignment of individuals to several tasks.

3. A problem-solving exercise in such areas as production variances, stock handling or machine downtime.

The technique, which may be used for screening, identifying, selecting, or developing managerial talent, has increased in popularity because of the many advantages of the system. This technique has an effect of equalizing opportunity, improving morale, and providing a "pool" of qualified applicants. Additional advantages are achieved by providing all participants, including low status or low visibility areas of an organization, a chance for organizational exposure.

The assessment center technique also has a high degree of acceptance by the participants. This high acceptance is because the situational exercises, unique to the technique, contribute a more valid prediction of managerial performance than paper-and-pencil tests or trait ratings. In addition, the technique provides little differential-validity with respect to age, sex, race, present job/position, or job differences of the participants.

The assessment center technique is also advantageous with respect to the assessors. There are essentially no differences between trained managers in the role of an assessor and an industrial psychologist; therefore, the industrial psychologist may be utilized in the training and research associated with the technique. The training of the assessment center staff assists in the reliability of the technique as well as provides the assessor with skills that may be directly transfered to "regular" positions within the organization.

The most obvious limitations of the assessment center techniques are the typical considerations associated with special programs, such as, cost, time commitments for assessors and assessees, and a formal, professionally designed efficient program. In addition, effects on the participant may indicate the dampening

of individual initiative or incentive by producing managers who tend to conform to the values and views of the assessors.

The process of selection of individuals to attend the assessment center may result in reduced motivation for individuals not selected and may produce an "untouchable" syndrome for those individuals selected. Attendance at a center may also produce significant anxieties in participants, especially in the long-term.

OTHER TECHNIQUES

Other IPA techniques in use, but not discussed in this review, include:

1. Grading,
2. Check Lists,
3. Man-to-Man Comparisons,
4. Work Standards.

SUMMARY

In this review, various techniques, instruments and programs used to appraise personnel have been discussed. Basic guidelines

to consider when planning, developing, implementing and maintaining an IPA program have been concisely summarized by Baylie and reported in Yoder, and are presented below.

GUIDELINES FOR ESTABLISHMENT OF
AN IPA PROGRAM

- Be sure that an appraisal program is really wanted before one is started.
- Do not adopt another organization's appraisal or rating program without tailoring it to meet the needs of your individual organization.
- Be sure that top management and personnel are in agreement on what the appraisal program is expected to accomplish, how the results should be used, and how the results should not be used.
- Determine specifically who will be included in the program before designing it.
- Be sure that the appraisal program is compatible with the philosophy and objectives of the company.
- Be sure that management will actively support the program.
- Recognize the options or choices in designing the program. Consider the pros and cons of each one.
- Identify and develop the basic policy and see that the program is responsive to it.

- Involve users in developing the program and in planning its implementation.

- Be sure to test the program by running a pilot program before introducing the final program on a broad scale.

- Be aware, that the basic policy, the program, and its administration may require frequent revisions.

- Do not fail to communicate the objectives intent, and mechanics of the program to all participants prior to its implementation.

- Be sure that the supervisors who will serve as appraisers are adequately trained before they are expected to appraise and counsel their subordinates.

- As one element in maintaining an ongoing program, be sure that new supervisors, as they are appointed, receive training in the philosophy, mechanics, and objectives of the program and in how to perform appraisals.

- Be sure to monitor the program periodically to ensure that interviews and ratings are being accomplished, that the results of the appraisals are being used as intended, that the program is continuing to meet the needs of the organization, and that it is being revised as required so that it will be responsive to changing organizational goals.[1]

In conclusion, IPA techniques are being used at the management level, particularly at the lower levels of management.

---

[1]Baylie, op. cit., p. 4-200.

If the probability that good performance will be recognized and rewarded and, concurrently, poor performance will be corrected, an IPA system can contribute positively to both organizational and individual morale and performance. The use of IPA techniques for both management development and administrative purposes has been recognized to the extent many organizations are initiating two or more IPA systems to adequately handle important personnel functions.

CHAPTER III

ANALYSIS OF THE DATA

INTRODUCTION

This study was designed to test the hypothesis that, in the administration of the individual performance appraisal, there are differences in how raters from different hierarchical levels within an organization assess job-related criteria on subordinate members of the organization. Of particular interest was which rater, or combination of raters, within the organizational structure and which job related criterion, or criteria, correlate the highest with a composite score.

In addition, the study was designed to examine the reliability of the appraisal instrument through test-retest methodology. Specifically, how do the data obtained during the assessment of job-related criteria on subordinate members of the organization compare if an identical instrument is administered on

two separate occasions. The correlation between the initial and subsequent appraisal should be high if the test is to be considered reliable.

DEMOGRAPHIC INFORMATION

SAMPLE POPULATION

The sample utilized in this study was selected from a total available population of 988 individuals who were assigned to the position of foreman within the "Firm." Of the total foremen available, 937 were identified as male with 830 (84.0%) white, and 101 (10.2%) black. A total of 51 (5.2%) individuals were not identified by race/sex. (Table 3.1)

From the available foremen, a sample of 90 foremen was selected. This sample was 9.1% of the total foremen available in the eight locations and included 66 (73.3%) white males and 14 (15.6%) black males. (Table 3.2)

In a comparison of foremen evaluated and the foremen available for evaluation, based on the variables of sex and race, the division between white and black

TABLE 3.1

NUMBER AND PERCENTAGE OF FOREMEN AVAILABLE FOR SAMPLE ON THE VARIABLES OF RACE AND SEX

RACE	SEX				MISSING DATA		TOTAL	
	MALE		FEMALE					
	N	%	N	%	N	%	N	%
WHITE	830	84.0	4	0.4	-	-	834	84.4
BLACK	101	10.2	-	-	-	-	101	10.2
SPANISH SURNAMED AMERICAN	1	0.1	-	-	-	-	1	0.1
AMERICAN INDIAN	1	0.1	-	-	-	-	1	0.1
MISSING DATA	-	-	-	-	51	5.2	51	5.2
TOTAL	933	94.4	4	0.4	51	5.2	988	100.0

TABLE 3.2

NUMBER AND PERCENTAGE OF FOREMAN SAMPLE ON THE VARIABLES OF RACE AND SEX

RACE	SEX								PERCENT*
	MALE		FEMALE		MISSING		TOTAL		
	N	%	N	%	N	%	N	%	
WHITE	66	73.3	4	4.4	-	-	70	77.7	7.1
BLACK	14	15.6	-	-	-	-	14	15.6	1.4
MISSING DATA	-	-	-	-	6	6.7	6	6.7	0.6
TOTAL	80	88.9	4	4.4	6	6.7	90	100.0	9.1
%*	-	8.1	-	0.4	-	0.6	-	9.1	

* - Percentage of Total Available Foremen.

male foremen was lower for male, white foremen by 10.7% (84.0% - 73.3%) and higher for male, black foremen by 5.4% (15.6% - 10.2%). In view of the current emphasis on blacks and other minority groups in management level positions, the sample by race from actual seemed acceptable.

The availability and sample of female foremen was less than 1% and will not be considered significant in this study. Identified male foremen whose race was other than white or black was also considered insignificant for reporting because they had the potential to comprise a maximum of 5.4% of the male foreman sample (Table 3.1).

Although the foremen available at each location varied in size, the sample selected at each location ranged from a minimum of nine to a maximum of 16 individuals. The majority of the locations had a sample of 10, with two locations (B5 and F6) having samples of 15 and 16, respectively. The locations with the larger samples were unique in that evaluation of objective "hard-core" criteria occurred for eight weeks compared to four week programs at other locations.

The sample was also divided between currently ranked above-average performing foremen and currently ranked below-average performing foremen. This division was obtained by having "second level" or indirect supervision rank all of the available foremen on the basis of overall job performance, using a multiple alternation ranking procedure. A more detailed explanation of this ranking was discussed in previous chapters.

The high/low ranking of the sample included 43 male foremen ranked as above-average and 37 male foremen ranked as below average performers. The remaining 10 foremen (11.1%) were either female foremen or information regarding ranking was not available. (Table 3.3)

The majority of the sample (83.3%) worked the second shift (day shift). The midnight and afternoon shifts had 7.8% and 6.7% of the sample, respectively. (Table 3.4)

TABLE 3.3

NUMBER AND PERCENTAGE OF FOREMAN SAMPLE ON THE VARIABLES OF LOCATION, RANKING, SEX AND RACE

LOCATION CODE#	RANKING								
	MALE					TOTAL			
	HIGH		LOW			MALE			
	WHITE	BLACK	WHITE	BLACK	OTHER*	WHITE	BLACK	OTHER*	NUMBER
A3	3	2	3	2	-	6	4	-	10
B5	5	1	4	3	2	9	4	2	15
C1	6	-	3	-	-	9	-	-	9
D3	4	-	3	-	3	7	-	3	10
E4	5	-	2	2	1	7	2	1	10
F6	6	1	6	1	2	12	2	2	16
G4	6	-	4	-	-	10	-	-	10
I2	4	-	2	2	2	6	2	2	10
TOTAL	39	4	27	10	10	66	14	10	90
%	43.4	4.4	30.0	11.1	11.1	73.3	15.6	11.1	100.0

- See Organization Chart "B."

* - Includes Female Foremen and Missing Data.

TABLE 3.4

NUMBER AND PERCENTAGE OF FOREMAN SAMPLE BY LOCATION AND SHIFT

LOCATION CODE #	SHIFT CODE*			MISSING DATA	TOTAL
	A	B	C		
A3	-	10	-	-	10
B5	2	10	3	-	15
C1	1	8	-	-	9
D3	-	10	-	-	10
E4	4	5	1	-	10
F6	-	13	2	1	16
G4	-	10	-	-	10
I2	-	9	-	1	10
TOTAL	7	75	6	2	90
%	7.8	83.3	6.7	2.2	100.0

- See Organization Chart "B".

* - SHIFT CODE

A - MIDNIGHT

B - DAY

C - AFTERNOON

The sample was comprised of 61.1% (55) foremen working in production related departments. Other departments represented within the sample include Plant Engineering (12.2%), Quality Control (7.8%), Material Control (7.8%), Maintenance (5.6%), and Material Handling (4.4%). (Table 3.5)

Time-as-Foreman data were also collected. The largest single grouping, by years experience, was 39 individuals (43.3%) who had over five years experience as a foreman. The majority of these foremen were white (84.6%). Missing data on 28 individuals, coupled with the foremen with over five years experience accounted for 74.4% of the sample in this category. (Table 3.6) Time-in-Position data indicated 56.6% of the foreman sample had a time-in-position of five years or less, with 80.4% (41 of 51) of these foremen having three or less years experience in the position. (Table 3.7)

In summary, the sample selected was considered adequate in size, diverse in background and experience, and representative of the total foremen available.

TABLE 3.5

NUMBER AND PERCENTAGE OF FOREMAN SAMPLE BY LOCATION AND DEPARTMENT

LOCATION CODE#	DEPARTMENT CODE*						MISSING DATA	TOTAL
	P	MC	QC	MH	PE	M		
A3	6	-	-	2	2	-	-	10
B5	10	-	-	2	2	1	-	15
C1	5	2	-	-	2	-	-	9
D3	6	1	1	-	-	2	-	10
E4	6	-	2	-	-	2	-	10
F6	10	2	2	-	2	-	-	16
G4	6	2	-	-	2	-	-	10
I2	6	-	2	-	1	-	1	10
TOTAL	55	7	7	4	11	5	1	90
%	61.1	7.8	7.8	4.4	12.2	5.6	1.1	100.0

- See Organization Chart "B".

* - DEPARTMENT CODE

P - PRODUCTION/TOOL AND DIE
MC - MATERIAL CONTROL
QC - QUALITY CONTROL
MH - MATERIAL HANDLING
PE - PLANT ENGINEERING
M - MAINTENANCE

TABLE 3.6

NUMBER AND PERCENTAGE OF FOREMAN SAMPLE ON THE VARIABLES OF TIME AS FOREMAN AND RACE

TIME	RACE				MISSING DATA		TOTAL	
	WHITE		BLACK					
YEARS	N	%	N	%	N	%	N	%
0-1	2	2.2	1	1.1	-	-	3	3.3
1-2	5	5.6	-		-	-	5	5.6
2-3	5	5.6	2	2.2	-	-	7	7.8
3-4	1	1.1	-		-	-	1	1.1
4-5	6	6.7	1	1.1	-	-	7	7.8
OVER 5	33	36.6	6	6.7	-	-	39	43.3
MISSING DATA	18	20.0	4	4.4	6	6.7	28	31.1
TOTAL	70	77.8	14	15.5	6	6.7	90	100.0

TABLE 3.7

NUMBER AND PERCENTAGE OF FOREMAN SAMPLE ON THE VARIABLES OF TIME IN POSITION AND RACE

TIME	RACE				MISSING DATA		TOTAL	
	WHITE		BLACK					
YEARS	N	%	N	%	N	%	N	%
0-1	9	10.0	4	4.4	-	-	13	14.4
1-2	16	17.8	-	-	-	-	16	17.8
2-3	10	11.1	2	2.2	-	-	12	13.3
3-4	2	2.2	1	1.1	-	-	3	3.3
4-5	5	5.6	2	2.2	-	-	7	7.8
OVER 5	9	10.0	-	-	-	-	9	10.0
MISSING DATA	19	21.1	5	5.6	6	6.7	30	33.4
TOTAL	70	77.8	14	15.5	6	6.7	90	100.0

INSTRUMENTATION

The final instrument included a rating scale on five categories of job duties (task) and six categories of job skills (behavior). A categorical listing of the task and behavior criteria appraised included:

JOB DUTIES - TASK

Productivity
Safety and Housekeeping
Maintenance
Motivating the Workforce
Management Attitude

JOB SKILLS - BEHAVIOR

Long-Range Planning
Short-Range Planning
Problem Solving
Interpersonal Relations
Goal Setting
Personal Work Standards

The appraisal of each criteria within the categories was on a nine-point scale ranging from a maximum of

"Usually Better" (9) to a minimum of "Usually Worse" (1) for job duties and from a maximum of "More Than Most" (9) to a minimum of "Less Than Most" (1) for job skills.

METHODOLOGY

After completion of a four (or eight) week recording of selected "hard-core" data on each foreman, the immediate direct supervisor (G1) subjectively rated the foreman. Concurrent with this rating, the "second level," indirect supervisor of the foreman (S1) completed an identical rating instrument. Upon completion of the supervisor's rating, a non-related supervisor (I) conducted a performance-incident interview with each supervisor and on the basis of this interview, rated the foreman on a rating instrument identical to the instrument utilized by the other supervisors.

Approximately one month later, identical instruments, with one additional criteria - "Overall Performance," were administered to the direct (G2) and indirect (S2) supervisors.

In this study, a composite score was derived from the means of the aforementioned task and behavioral criteria as assigned by five individual rater groups (G1, G2, S1, S2, I). This composite score was used as a basis to determine what degree the criteria, or combinations thereof, varied together or were correlated. The coefficient of correlation (r) was calculated for comparative purposes to measure the magnitude of the relationships between criteria/test areas/raters.

In addition, the coefficient of determination, or variance (r^2) for each coefficient of correlation was identified to measure the significance of the relationships between criteria/test areas/raters. Variance is the coefficient of correlation squared. For example, squaring the coefficient of correlation that equals .90, provides $r^2 = .81$ which means that 81 per cent of the variance of B is shared in common with A or vice versa. If two correlated variables are identified as A and B, the percentage of variance in B associated with or determined by the variance in A is the coefficient of determination.[1]

---

[1]Fred N. Kerlinger, Foundations of Behavioral Research (Chicago, Ill.: Holt, Rinehart and Winston, Inc., 1973), p. 88.

RESPONSES

The percentage of task and behavioral criteria responses by location ranged from a maximum of 100% at two locations to a minimum of 52% at one location. The overall percentage of response was 79.8% (358 of 450 responses). (Table 3.8)

A description of terms used by code and category is presented in Table 3.9.

The tables that follow (Tables 3.10 through 3.19) provide the cell by cell number and percentage of responses assigned the foreman sample by multiple raters (G1, G2, S1, S2, I) for task and behavior category criteria.

For the task category, test area (T1) criteria (Tables 3.10 through 3.14) the number of responses by direct, indirect and non-related supervision ranged from 82 (G1: SAFETY and G1: MGTATT) to 51 (S2: MOTIV). The means of this test area criteria by rater ranged from a high of 6.44 (S2) to a low of 6.06 (I). The individual criteria mean ranged from a high of 6.71 (PROD) to a low of 5.49 (SAFETY).

TABLE 3.8

NUMBER AND PERCENTAGE OF RESPONSES ON
TASK AND BEHAVIOR CATEGORY
CRITERIA BY LOCATION

LOCATION CODE#	PLANNED	ACTUAL	PERCENTAGE RESPONDING
A3	50	26	52.0
B5	75	67	89.3
C1	45	36	80.0
D3	50	50	100.0
E4	50	34	68.0
F6	80	56	70.0
G4	50	39	78.0
I2	50	50	100.0
SUB-TOTAL	450	358	79.6
MISSING	-	92	-
TOTAL	450	450	-

- See Organization Chart "B."

TABLE 3.9

DESCRIPTION OF TERMS USED IN THE STUDY
BY CODE AND CATEGORY

#	CODE	DESCRIPTION
1	ABSENTF	ABSENCE-FOREMAN
2	ABSENTH	ABSENCE - HOURLY PERSONNEL
3	BEHAV	BEHAVIORAL CATEGORY TEST AREA
4	COMPOSIT	COMPOSITE - MEAN OF TASK AND BEHAVIORAL CATEGORY TEST AREAS
5	DEFTAG	DEFECTIVE TAG OPERATIONS
6	DISCACT	DISCIPLINARY ACTION REPORTS
7	DLABVAR	DIRECT LABOR VARIANCE
8	G	RATER MEAN - DIRECT SUPERVISION
9	G1	RATER-DIRECT SUPERVISION-INITIAL EVALUATION
10	G2	RATER-DIRECT SUPERVISION - FINAL EVALUATION
11	GOALS	GOAL SETTING SKILL
12	GRIEV	GRIEVANCES
13	GSI	RATER COMBINATION - MEAN OF G1 + S1 + I
14	HARD	HARD-CORE CATEGORY TEST AREA
15	I	RATER - NON-RELATED SUPERVISION
16	INTREL	INTERPERSONAL RELATIONS SKILL
17	LRPLAN	LONG-RANGE PLANNING SKILL
18	MAINT	MAINTENANCE
19	MATCOST	MATERIAL COST
20	MGTATT	MANAGEMENT ATTITUDE
21	MOTIV	MOTIVATING THE WORKFORCE
22	NETPROD	NET PRODUCTIVITY
23	OSDIRLA	OFF-STANDARD DIRECT LABOR
24	OVERAL	OVERALL PERFORMANCE
25	PAYSRT	PAY SHORTAGES
26	PROD	PRODUCTIVITY
27	PSOLV	PROBLEM SOLVING SKILL
28	REJECTS	REJECTS
29	RINGFAL	RING FAILURES
30	S	RATER MEAN - INDIRECT SUPERVISION
31	S1	RATER - INDIRECT SUPERVISION - INITIAL EVALUATION
32	S2	RATER - INDIRECT SUPERVISION - FINAL EVALUATION
33	SAFETY	SAFETY AND HOUSEKEEPING
34	SAFEVIO	SAFETY VIOLATIONS
35	SCRAP	SCRAP
36	SRPLAN	SHORT-RANGE PLANNING SKILL
37	T1	TEST AREA - TASK CRITERIA
38	T2	TEST AREA - BEHAVIORAL CRITERIA
39	T3	TEST AREA - HARD-CORE CRITERIA
40	TASK	TASK CATEGORY - TEST AREA
41	WORKHAB	WORK HABITS
42	WRKSTD	PERSONAL WORK STANDARDS
43	XG	RATER COMBINATION - MEAN OF G1 + G2
44	XGSI	RATER COMBINATION - MEAN OF XG + XS + I
45	XGXS	RATER COMBINATION - MEAN OF XG + XS
46	XG2S2	RATER COMBINATION - MEAN OF G2 + S2
47	XS	RATER COMBINATION - MEAN OF S1 + S2

TABLE 3.10

NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY DIRECT SUPERVISION (G1) ON TASK-ORIENTED CRITERIA - INITIAL APPRAISAL

CRITERIA CODE #	N	SCORE 9 N	9 %	8 N	8 %	7 N	7 %	6 N	6 %	5 N	5 %	4 N	4 %	3 N	3 %	2 N	2 %	1 N	1 %	MEAN	STD. DEV.
PROD.	81	12	14.8	12	14.8	29	35.8	10	12.3	11	13.6	4	4.9	2	2.5	1	1.2	0	0	6.74	1.59
SAFETY	82	9	11.0	5	6.1	18	22.0	15	18.3	9	11.0	20	24.4	4	4.9	1	1.2	1	1.2	5.82	1.87
MAINT	81	14	17.3	11	13.6	20	24.7	16	19.8	10	12.3	9	11.1	1	1.2	0	0	0	0	6.65	1.62
MOTIV	79	8	10.1	11	13.9	14	17.7	11	13.9	12	15.2	8	10.1	14	17.7	0	0	1	1.3	5.81	2.03
MGTATT	82	15	18.3	7	8.5	14	17.1	21	25.6	11	13.4	7	8.5	5	6.1	2	2.4	0	0	6.30	1.88

- See Table 3.9 for description of criteria.

TABLE 3.11

NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY DIRECT SUPERVISION (G2) ON TASK-ORIENTED CRITERIA - FINAL APPRAISAL

CRITERIA CODE #	N	SCORE 9 N	9 %	8 N	8 %	7 N	7 %	6 N	6 %	5 N	5 %	4 N	4 %	3 N	3 %	2 N	2 %	1 N	1 %	MEAN	STD. DEV.
PROD	57	8	14.0	11	19.3	14	24.6	8	14.0	9	15.8	6	10.5	0	0	0	0	1	1.8	6.60	1.73
SAFETY	58	6	10.3	7	12.1	9	15.5	11	19.0	14	24.1	10	17.2	1	1.7	0	0	0	0	6.07	1.64
MAINT	57	6	10.5	10	17.5	19	33.3	10	17.5	4	7.0	4	7.0	4	7.0	0	0	0	0	6.58	1.64
MOTIV	57	7	12.3	6	10.5	7	12.3	15	26.3	10	17.5	8	14.0	2	3.5	1	1.8	1	1.8	5.98	1.87
MGTATT	58	11	19.0	8	13.8	9	15.5	8	13.8	13	22.4	5	8.6	3	5.2	1	1.7	0	0	6.38	1.91

- See Table 3.9 for description of criteria.

TABLE 3.12

NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY INDIRECT SUPERVISION (S1) ON TASK-ORIENTED CRITERIA - INITIAL APPRAISAL

CRITERIA CODE #	N	SCORE 9 N	9 %	8 N	8 %	7 N	7 %	6 N	6 %	5 N	5 %	4 N	4 %	3 N	3 %	2 N	2 %	1 N	1 %	MEAN	STD. DEV.
PROD	74	14	18.9	10	13.5	25	33.8	7	9.5	9	12,2	2	2.7	5	6.8	1	1.4	1	1.4	6.68	1.89
SAFETY	74	8	10.8	8	10.8	7	9.5	13	17.6	12	16.2	17	23.0	7	9.5	2	2.7	0	0	5.62	1.94
MAINT	66	8	12.1	16	24.2	12	18.2	9	13.6	7	10.6	5	7.6	8	12.1	0	0	1	1.5	6.33	2.02
MOTIV	70	5	7.1	15	21.4	12	17.1	8	11.4	12	17.1	7	10.0	10	14.3	1	1.4	0	0	5.96	1.95
MOTATT	67	14	20.9	12	17.9	15	22.4	7	10.4	6	9.0	4	6.0	4	6.0	0	0	5	7.5	6.45	2.33

- See Table 3.9 for description of criteria.

TABLE 3.13

NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY INDIRECT SUPERVISION (S2) ON TASK-ORIENTED CRITERIA - FINAL APPRAISAL

CRITERIA CODE #	N	SCORE 9 N	9 %	8 N	8 %	7 N	7 %	6 N	6 %	5 N	5 %	4 N	4 %	3 N	3 %	2 N	2 %	1 N	1 %	MEAN	STD. DEV.
PROD	52	11	21.2	7	13.5	13	25.0	7	13.5	6	11.5	5	9.6	1	1.9	0	0	2	3.8	6.60	2.01
SAFETY	55	9	16.4	6	10.9	13	23.6	10	18.2	9	16.4	7	12.7	0	0	0	0	1	1.8	6.44	1.77
MAINT	53	11	20.8	3	5.7	11	20.8	13	24.5	5	9.4	5	9.4	4	7.5	0	0	1	1.9	6.34	1.97
MOTIV	51	6	11.8	9	17.6	14	27.5	9	17.6	6	11.8	2	3.9	2	3.9	1	2.0	2	3.9	6.40	1.97
MGTATT	53	7	13.2	16	30.2	6	11.3	8	15.1	7	13.2	2	3.8	4	7.5	1	1.9	2	3.8	6.42	2.15

- See Table 3.9 for description of criteria.

TABLE 3.14

NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY NON-RELATED SUPERVISION (I) ON TASK-ORIENTED CRITERIA

CRITERIA CODE #	N	SCORE 9 N	9 %	8 N	8 %	7 N	7 %	6 N	6 %	5 N	5 %	4 N	4 %	3 N	3 %	2 N	2 %	1 N	1 %	MEAN	STD. DEV.
PROD	80	8	10.0	24	30.0	22	27.5	7	8.8	8	10.0	7	8.8	2	2.5	1	1.3	1	1.3	6.71	1.76
SAFETY	81	4	4.9	8	9.9	13	16.0	16	19.8	14	17.3	14	17.3	9	11.1	2	2.5	1	1.2	5.49	1.83
MAINT	78	5	6.4	16	20.5	20	25.6	13	16.7	10	12.8	10	12.8	1	1.3	2	2.6	1	1.3	6.27	1.76
MOTIV	81	3	3.7	11	13.6	15	18.5	10	12.3	17	21.0	15	18.5	10	12.3	0	0	0	0	5.62	1.74
NGTATT	81	10	12.3	17	21.0	14	17.3	9	11.1	15	18.5	8	9.9	4	4.9	2	2.5	2	2.5	6.21	2.04

- See Table 3.9 for description of criteria.

TABLE 3.15

NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY DIRECT SUPERVISION (G1) ON BEHAVIOR ORIENTED CRITERIA - INITIAL APPRAISAL

CRITERIA CODE #	N	SCORE 9 N	9 %	8 N	8 %	7 N	7 %	6 N	6 %	5 N	5 %	4 N	4 %	3 N	3 %	2 N	2 %	1 N	1 %	MEAN	STD. DEV.
LRPLAN	80	8	10.0	18	22.5	15	18.8	16	20.0	10	12.5	9	11.3	4	5.0	0	0	0	0	6.44	1.69
SRPLAN	82	13	15.9	13	15.9	19	23.2	18	22.0	11	13.4	6	7.3	2	2.4	0	0	0	0	6.67	1.59
PSOLV	81	10	12.3	12	14.8	18	22.2	17	21.0	12	14.8	9	11.1	3	3.7	0	0	0	0	6.41	1.65
INTREL	82	7	8.5	20	24.4	10	12.2	12	14.6	16	19.5	10	12.2	7	8.5	0	0	0	0	6.17	1.82
GOALS	80	8	10.0	18	22.5	18	22.5	13	16.3	12	15.0	5	6.3	6	7.5	0	0	0	0	6.48	1.71
WRKSTD	82	13	15.9	14	17.1	18	22.0	13	15.9	10	12.2	11	13.4	3	3.7	0	0	0	0	6.54	1.75

- See Table 3.9 for description of criteria.

TABLE 3.16

NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY DIRECT SUPERVISION (G2) ON BEHAVIORAL ORIENTED CRITERIA FINAL APPRAISAL

CRITERIA CODE #	N	SCORE 9 N	9 %	8 N	8 %	7 N	7 %	6 N	6 %	5 N	5 %	4 N	4 %	3 N	3 %	2 N	2 %	1 N	1 %	MEAN	STD. DEV.
LRPLAN	57	6	10.5	10	17.5	14	24.6	13	22.8	4	7.0	8	14.0	1	1.8	1	1.8	0	0	6.44	1.68
SRPLAN	56	10	17.9	5	8.9	15	26.8	10	17.9	9	16.1	4	7.1	1	1.8	1	1.8	1	1.8	6.46	1.85
PSOLV	57	4	7.0	8	14.0	19	33.3	6	10.5	9	15.8	6	10.5	3	5.3	1	1.8	1	1.8	6.14	1.83
INTREL	56	7	12.5	8	14.3	10	17.9	13	23.2	10	17.9	5	8.9	2	3.6	1	1.8	0	0	6.30	1.73
GOALS	56	2	3.6	10	17.9	14	25.0	11	19.6	12	21.4	3	5.4	2	3.6	1	1.8	1	1.8	6.	1.
WRKSTD	57	9	15.8	7	12.3	13	22.8	13	22.8	8	14.0	4	7.0	1	1.8	1	1.8	1	1.8	6.46	1.81
OVERALL	57	3	5.3	11	19.3	16	28.1	9	15.8	8	14.0	7	12.3	1	1.8	1	1.8	1	1.8	6 .23	1.74

- See Table 3.9 for description of criteria.

TABLE 3.17

NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY INDIRECT SUPERVISION (S1) BEHAVIOR ORIENTED CRITERIA - INITIAL APPRAISAL

CRITERIA CODE #	N	SCORE 9 N	9 %	8 N	8 %	7 N	7 %	6 N	6 %	5 N	5 %	4 N	4 %	3 N	3 %	2 N	2 %	1 N	1 %	MEAN	STD. DEV.
LRPLAN	60	8	13.3	16	26.7	9	15.0	10	16.7	3	5.0	8	13.3	5	8.3	1	1.7	0	0	6.45	1.97
SRPLAN	69	11	15.9	16	23.2	17	24.6	10	14.5	9	13.0	3	4.3	3	4.3	0	0	0	0	6.84	1.62
PSOLV	68	9	13.2	15	22.1	13	19.1	8	11.8	10	14.7	3	4.4	9	13.2	1	1.5	0	0	6.34	1.99
INTREL	69	7	10.1	16	23.2	17	24.6	5	7.2	7	10.1	7	10.1	7	10.1	2	2.9	1	1.4	6.22	2.09
GOALS	64	4	6.3	9	14.1	14	21.9	15	23.4	13	20.3	3	4.7	5	7.8	1	1.6	0	0	6.09	1.65
WRKSTD	66	12	18.2	11	16.7	12	18.2	14	21.2	9	13.6	3	4.5	5	7.6	0	0	0	0	6.61	1.78

- See Table 3.9 for description of criteria.

TABLE 3.18

NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY INDIRECT SUPERVISION (S2) ON BEHAVIOR ORIENTED CRITERIA FINAL APPRAISAL

CRITERIA CODE #	N	SCORE 9 N	9 %	8 N	8 %	7 N	7 %	6 N	6 %	5 N	5 %	4 N	4 %	3 N	3 %	2 N	2 %	1 N	1 %	MEAN	STD. DEV.
LRPLAN	50	5	10.0	5	10.0	17	34.0	5	10.0	9	18.0	5	10.0	2	4.0	1	2.0	1	2.0	6.16	1.86
SRPLAN	53	5	9.4	14	26.4	11	20.8	9	17.0	9	17.0	3	5.7	0	0	0	0	1	3.8	6.55	1.79
PSOLV	52	5	9.6	11	21.2	11	21.2	9	17.3	9	17.3	3	5.8	2	3.8	0	0	2	3.8	6.33	1.90
INTREL	53	4	7.5	15	28.3	9	17.0	12	22.6	4	7.5	2	3.8	2	3.8	2	3.8	3	5.7	6.26	2.14
GOALS	54	6	11.1	2	3.7	11	20.4	16	29.6	9	16.7	5	9.3	2	3.7	1	1.9	2	3.7	5.89	1.87
WRKSTD	54	5	9.3	13	24.1	9	16.7	11	20.4	9	16.7	3	5.6	1	1.9	0	0	3	5.6	6.31	1.97
OVERALL	52	6	11.5	12	23.1	11	21.2	9	17.3	6	11.5	4	7.7	0	0	2	3.8	2	3.8	6.40	2.02

- See Table 3.9 for description of criteria.

TABLE 3.19

NUMBER AND PERCENTAGE OF RESPONSES ASSIGNED TO FOREMAN SAMPLE BY NON-RELATED SUPERVISION (I) ON BEHAVIOR ORIENTED CRITERIA

CRITERIA CODE #	N	SCORE 9 N	9 %	8 N	8 %	7 N	7 %	6 N	6 %	5 N	5 %	4 N	4 %	3 N	3 %	2 N	2 %	1 N	1 %	MEAN	STD. DEV.
LRPLAN	76	5	6.6	22	28.9	18	23.7	12	15.8	10	13.2	5	6.6	2	2.6	1	1.3	1	1.3	6.55	1.70
SRPLAN	79	5	6.3	25	31.6	23	29.1	11	13.9	11	13.9	1	1.3	2	2.5	1	1.3	0	0	6.82	1.45
PSOLV	80	6	7.5	23	28.8	18	22.5	9	11.3	9	11.3	9	11.3	6	7.5	0	0	0	0	6.46	1.76
INTREL	80	6	7.5	15	18.8	14	17.5	14	17.5	12	15.0	14	17.5	4	5.0	1	1.3	0	0	6.08	1.77
GOALS	77	3	3.9	19	24.7	19	24.7	12	15.6	10	13.0	8	10.4	5	6.5	1	1.3	0	0	6.27	1.70
WRKSTD	80	7	8.8	18	22.5	17	21.3	12	15.0	12	15.0	5	6.3	7	8.8	2	2.5	0	0	6.29	1.86

- See Table 3.9 for description of criteria.

In a comparison between initial and final responses by direct supervision (G1, G2), the relative task area ranking remained the same from appraisal to appraisal. (Tables 3.10 and 3.11)

In a comparison between initial and final responses by indirect supervision (S1, S2), the relative task area ranking of two criteria decreased (MGTATT, MAINT), one criterion increased (SAFETY), one criterion remained unchanged (PROD) and one criterion was utilized only once (OVERALL) as shown in Tables 3.12 and 3.13.

For the behavior category test area (T2) criteria, the number of responses by direct, indirect and non-related supervision ranged from 82 (G1: SRPLAN and G1: WRKSTD) to 52 (S2: OVERAL). The mean of this test area criteria by rater ranged from a high of 6.45 (G1) to a low of 6.27 (S2). The individual criteria mean ranged from a high of 6.84 (SRPLAN) to a low of 5.89 (GOALS).

In a comparison between initial and final responses by direct supervision (G1, G2), the relative test area ranking of three criteria increased (LRPLAN, INTREL, WRKSTD), one criterion decreased (GOALS), and one

criterion (PSOLV) remained unchanged (Table 3.15 and 3.16). In a comparison between initial and final responses by indirect supervision (S1, S2), the relative task area ranking of two criteria increased (PSOLV, INTREL), two criteria decreased (LRPLAN, WRKSTD), one criterion (SRPLAN) was unchanged and one criterion (OVERAL) was utilized only once (Tables 3.17 and 3.18).

HARD-CORE CRITERIA

The percentage of responses by location ranged from a maximum of 100% at one location (I2) to a minimum of 73.3% at one location (B5). The overall percentage of response was 88.0% (426 of 484 possible responses). (Table 3.20)

The response, sample mean, and weighted mean on the hard-core criteria appraisal follows (Table 3.21). In the tables, the number following the criteria code indicates the week in which the data was obtained. For example, NETPROD1 is comprised of the scores assigned to the foreman for the criteria Net Productivity during the first week of evaluation.

TABLE 3.20

NUMBER AND PERCENTAGE OF RESPONSES ON "HARD-CORE" CRITERIA BY LOCATION

LOCATION CODE*	PLANNED	ACTUAL	PERCENTAGE RESPONDING
A3	40	31	77.5
B5	120	88	73.3
C1	36	34	94.4
D3	40	40	100.0
E4	40	37	92.5
F6	128	124	96.9
G4	40	32	80.0
I2	40	40	100.0
SUB-TOTAL	484	426	88.0
MISSING	-	58	-
TOTAL	484	484	-

* - See Organization Chart "B."

TABLE 3.21

SAMPLE, MEAN AND WEIGHTED MEAN ON HARD-CORE CRITERIA BY WEEK

CRITERIA CODE*	N	SCORE 5	4	3	2	1	MISS.	MEAN	WEIGHTED MEAN**
NETPROD1	21	7	5	7	1	1	69	3.76	6.52
NETPROD2	26	9	7	9	0	1	64	3.88	3.77
NETPROD3	21	3	8	9	1	0	69	3.62	6.24
NETPROD4	22	4	5	12	1	0	68	3.55	6.09
NETPROD5	7	2	4	1	0	0	83	4.14	7.29
NETPROD6	2	0	1	1	0	0	88	3.50	6.00
NETPROD7	4	0	1	3	0	0	86	3.25	5.50
NETPROD8	2	0	1	1	0	0	88	3.50	6.00
MATCOST1	5	1	3	1	0	0	85	4.00	7.00
MATCOST2	5	2	2	1	0	0	85	4.20	7.40
MATCOST3	7	3	2	0	2	0	83	3.86	6.71
MATCOST4	2	2	0	0	0	0	88	5.00	9.00
SCRAP1	9	1	5	2	1	0	81	3.67	6.33
SCRAP2	9	1	5	2	0	1	81	3.56	6.11
SCRAP3	8	3	4	1	0	0	82	4.25	7.50
SCRAP4	7	3	2	1	1	0	83	4.00	7.00
SCRAP5	7	4	1	1	1	0	83	4.14	7.29
SCRAP6	2	0	2	0	0	0	88	4.00	7.00
SCRAP7	4	1	2	1	0	0	86	4.00	7.00
SCRAP8	2	1	0	1	0	0	88	4.00	7.00
REJECTS1	23	5	9	7	1	1	67	3.70	6.39
REJECTS2	17	2	6	7	2	0	73	3.47	5.94
REJECTS3	22	5	4	13	0	0	68	3.64	6.27
REJECTS4	17	4	5	6	1	1	73	3.59	6.18
REJECTS5	7	0	1	3	1	2	83	2.43	3.86
REJECTS6	10	0	4	3	2	1	80	3.00	5.00
REJECTS7	10	1	0	5	3	1	80	2.70	4.40
REJECTS8	8	1	2	3	2	0	82	3.25	5.50
DEFTAG1	4	1	1	1	0	1	86	3.25	5.50
DEFTAG2	5	3	1	0	0	1	85	4.00	7.00
DEFTAG3	3	3	0	0	0	0	87	5.00	9.00
DEFTAG4	3	3	0	0	0	0	87	5.00	9.00
WORKHAB1	7	0	3	2	1	1	83	3.00	5.50
WORKHAB2	7	0	3	1	2	1	83	2.86	4.71
WORKHAB3	5	0	2	3	0	0	85	3.40	5.80
WORKHAB4	5	0	2	2	1	0	85	3.20	5.40
SAFEVIO1	17	6	6	3	2	0	73	3.94	6.88
SAFEVIO2	22	10	2	9	1	0	60	3.95	5.55
SAFEVIO3	17	7	6	3	1	0	73	4.12	7.24
SAFEVIO4	18	8	4	6	0	0	72	4.11	7.22
SAFEVIO5	8	6	2	0	0	0	82	4.75	8.50
SAFEVIO6	5	3	0	2	0	0	85	4.20	7.40
SAFEVIO7	4	1	1	2	0	0	86	3.75	6.50
SAFEVIO8	4	3	1	0	0	0	86	4.75	8.50
OSDIRLA1	36	12	11	11	1	1	54	3.89	6.78
OSDIRLA2	33	13	6	11	2	1	57	3.85	6.70
OSDIRLA3	32	7	14	8	3	0	58	3.78	6.56
OSDIRLA4	29	7	8	13	1	0	61	3.72	6.45
OSDIRLA5	13	2	3	7	1	0	77	3.46	5.92
OSDIRLA6	14	3	6	4	1	0	76	3.79	6.57
OSDIRLA7	13	2	5	5	1	0	77	3.62	6.23
OSDIRLA8	9	4	2	2	1	0	81	4.00	7.00
DLABVAR1	6	0	3	2	1	0	84	3.33	5.67
DLABVAR2	6	0	4	2	0	0	84	3.67	6.33
DLABVAR3	5	1	1	3	0	0	85	3.60	6.20
DLABVAR4	5	0	3	2	0	0	85	3.60	6.20

(Table 3.21 COntinued)

CRITERIA CODE*	N	5	4	3	2	1	MISS.	MEAN	WEIGHTED MEAN**
ABSENTH1	83	36	22	22	2	1	7	4.08	7.17
ABSENTH2	77	30	18	27	1	1	13	3.97	6.95
ABSENTH3	72	32	15	23	2	0	18	4.07	7.14
ABSENTH4	63	17	20	23	1	2	27	3.78	6.56
ABSENTH5	21	8	5	6	1	1	69	3.86	6.71
ABSENTH6	20	5	8	5	1	1	70	3.75	6.50
ABSENTH7	20	7	6	5	1	1	70	3.85	6.70
ABSENTH8	17	7	5	5	0	0	73	4.12	7.23
ABSENTF1	87	-	-	-	9	78	3	1.10	1.21
ABSENTF2	86	-	-	-	13	73	4	1.15	1.30
ABSENTF3	82	-	-	-	11	71	8	1.13	1.27
ABSENTF4	73	-	-	-	12	61	17	1.16	1.33
ABSENTF5	26	-	-	-	5	21	64	1.19	1.38
ABSENTF6	27	-	-	-	4	23	63	1.15	1.30
ABSENTF7	22	-	-	-	2	20	68	1.09	1.18
ABSENTF8	20	-	-	-	3	17	70	1.15	1.30
GRIEV1	10	10	0	0	0	0	80	5.00	9.0
GRIEV2	10	10	0	0	0	0	80	5.00	9.0
GRIEV3	10	10	0	0	0	0	80	5.00	9.0
GRIEV4	10	10	0	0	0	0	80	5.00	9.0
DISCACT1	20	-	-	-	-	-	90	0	-
DISCACT2	20	-	-	-	-	-	90	0	-
DISCACT3	18	-	-	-	-	-	90	0	-
DISCACT4	17	-	-	-	-	-	90	0	-
RINGFAL1	46	19	13	9	3	2	44	3.96	6.91
RINGFAL2	40	20	13	4	3	0	50	4.25	7.50
RINGFAL3	40	26	10	4	0	0	50	4.55	8.10
RINGFAL4	35	17	10	8	0	0	55	4.26	7.51
RINGFAL5	21	12	5	3	1	0	69	4.33	7.67
RINGFAL6	20	8	8	4	0	0	70	4.20	7.40
RINGFAL7	20	10	4	6	0	0	70	4.20	7.40
RINGFAL8	17	11	3	2	1	0	73	4.41	7.82
PAYSRT1	53	37	7	6	0	3	37	4.42	7.83
PAYSRT2	50	32	12	5	0	1	40	4.48	7.96
PAYSRT3	47	33	10	4	0	0	43	4.62	8.23
PAYSRT4	37	28	6	2	1	0	53	4.65	8.30
PAYSRT5	21	17	3	1	0	0	69	4.76	8.52
PAYSRT6	20	14	5	1	0	0	70	4.65	8.30
PAYSRT7	20	15	3	2	0	0	70	4.65	8.30
PAYSRT8	17	14	3	0	0	0	73	4.82	8.65

* - See Table 3.9 for description of criteria.

** - WEIGHTED MEAN:

$$\frac{([5(n_5) + 4(n_5)] + [4(n_4) + 3(n_4)] + [3(n_3) + 2(n_3)] + [2(n_2) + 1(n_2)] + [1(n_1) + 0(n_1)]) \div N}{}$$

NOTE: n_n = Number of responses per "n" score.

N = Total number of responses.

After a review of the hard-core criteria data, it was determined that these data would not be utilized in the findings of this study. This decision was based on the diverse sample distribution and availability of the data. (See the performance data availability matrix in Chapter I.) In addition, the hard-core rating instruments typically required both subjective and objective responses and were administered by only one rater.

FINDINGS

The hypotheses were tested using correlation statistics. The MIDAS program at The University of Michigan was used for all analyses.

HYPOTHESES

For the purposes of statistical analysis, all hypotheses are stated in the null form in the following discussion.

Hypothesis 1 (Ho1)

There will be no differences in how raters from

different hierarchical levels within the organization appraise job-related criteria on subordinate members of the organization.

Results

Individuals in the job capacity of direct supervision, indirect supervision and non-related supervision were requested to appraise a selected foreman sample in two job-related categories - job duties (task) and job skills (behavior).

A listing of the job-related criteria and test areas by referent table follows (Table 3.22).

The coefficient of correlation and the coefficient of determination between the composite score and rater groups/combinations on criteria/test area categories were calculated and are presented in the tables that follow (Tables 3.23 through 3.37). A matrix of the coefficients of correlation by criteria/rater and a descending ranking of the coefficients of correlation and determination were developed. (Tables 3.38 and 3.39)

TABLE 3.22

JOB-RELATED CATEGORIES BY CRITERIA AND TABLE NUMBER

CRITERIA CODE*	TABLE NUMBER
JOB DUTIES	
PROD	3.23
SAFETY	3.24
MAINT	3.25
MOTIV	3.26
MGTATT	3.27
JOB SKILLS	
LRPLAN	3.28
SRPLAN	3.29
PSOLV	3.30
INTREL	3.31
GOALS	3.32
WRKSTD	3.33
OVERALL	
OVERAL	3.34
TEST AREAS	
T1	3.35
T2	3.36
T1T2	3.37

* - See Table 3.9 for description of criteria.

TABLE 3.23

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE TASK CATEGORY OF PRODUCTIVITY

RATER CODE*	MEAN	STD. DEV.	N	CORR	v^2
COMPOSIT	6.36	1.31	81	.78	.60
G1	6.74	1.59			
COMPOSIT	6.23	1.28	57	.74	.55
G2	6.50	1.73			
COMPOSIT	6.32	1.31	74	.81	.66
S1	6.68	1.89			
COMPOSIT	6.26	1.36	52	.80	.65
S2	6.60	2.01			
COMPOSIT	6.36	1.32	80	.83	.69
I	6.71	1.76			
COMPOSIT	6.36	1.29	85	.90	.80
XGSI	6.77	1.52			
COMPOSIT	6.37	1.29	83	.79	.63
XG	6.83	1.52			
COMPOSIT	6.37	1.32	78	.85	.72
XS	6.74	1.84			
COMPOSIT	6.36	1.29	85	.88	.77
XGXS	6.79	1.54			
COMPOSIT	6.36	1.29	85	.89	.79
GSI	6.74	1.54			

*See Table 3.9 for description of raters.

TABLE 3.24

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE TASK CATEGORY OF SAFETY

RATER CODE*	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.37	1.30	82	.53	.28
G1	5.82	1.87			
COMPOSIT	6.20	1.30	58	.55	.30
G2	6.07	1.64			
COMPOSIT	6.34	1.29	74	.62	.39
S1	5.62	1.94			
COMPOSIT	6.23	1.35	55	.63	.40
S2	6.44	1.77			
COMPOSIT	6.36	1.31	81	.60	.36
I	5.49	1.83			
COMPOSIT	6.36	1.29	85	.70	.49
XGSI	5.89	1.43			
COMPOSIT	6.37	1.29	83	.54	.29
XG	5.95	1.71			
COMPOSIT	6.37	1.32	78	.63	.40
XS	5.96	1.80			
COMPOSIT	6.36	1.29	85	.68	.46
XGXS	5.98	1.47			
COMPOSIT	6.36	1.29	85	.68	.46
GSI	5.67	1.55			

* - See Table 3.9 for description of raters.

TABLE 3.25

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE TASK CATEGORY OF MAINTENANCE

RATER-CODE*	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.40	1.28	81	.60	.36
G1	6.65	1.62			
COMPOSIT	6.23	1.28	57	.67	.45
G2	6.58	1.64			
COMPOSIT	6.33	1.32	66	.66	.44
S1	6.33	2.02			
COMPOSIT	6.19	1.35	53	.73	.54
S2	6.34	1.97			
COMPOSIT	6.36	1.33	78	.66	.43
I	6.27	1.76			
COMPOSIT	6.35	1.29	84	.78	.60
XGSI	6.52	1.40			
COMPOSIT	6.39	1.28	82	.66	.44
XG	6.69	1.45			
COMPOSIT	6.34	1.31	76	.73	.53
XS	6.43	1.84			
COMPOSIT	6.35	1.29	84	.77	.59
XGXS	6.57	1.43			
COMPOSIT	6.35	1.29	84	.70	.49
GSI	6.41	1.56			

* - See Table 3.9 for description of raters.

TABLE 3.26

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE TASK CATEGORY OF MOTIVATION

RATER CODE	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.36	1.32	79	.64	.41
G1	5.81	2.03			
COMPOSIT	6.20	1.31	57	.77	.59
G2	5.98	1.87			
COMPOSIT	6.29	1.33	70	.73	.53
S1	5.96	1.95			
COMPOSIT	6.25	1.34	51	.77	.59
S2	6.39	1.97			
COMPOSIT	6.36	1.31	81	.67	.45
I	5.62	1.74			
COMPOSIT	6.36	1.29	85	.84	.70
XGSI	5.96	1.60			
COMPOSIT	6.37	1.31	81	.75	.56
XG	5.94	1.88			
COMPOSIT	6.36	1.32	77	.79	.62
XS	6.25	1.84			
COMPOSIT	6.36	1.29	85	.84	.71
XGXS	6.11	1.68			
COMPOSIT	6.36	1.29	85	.76	.57
GSI	5.79	1.68			

* - See Table 3.9 for description of raters.

TABLE 3.27

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE TASK CATEGORY OF MANAGEMENT ATTITUDE

RATER CODE	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.37	1.30	82	.71	.50
G1	6.31	1.88			
COMPOSIT	6.20	1.30	58	.73	.54
G2	6.38	1.91			
COMPOSIT	6.37	1.22	67	.78	.61
S1	6.45	2.33			
COMPOSIT	6.23	1.32	53	.84	.71
S2	6.42	2.15			
COMPOSIT	6.36	1.31	81	.77	.60
I	6.21	2.04			
COMPOSIT	6.36	1.29	85	.87	.75
XGSI	6.35	1.77			
COMPOSIT	6.37	1.29	83	.78	.60
XG	6.42	1.78			
COMPOSIT	6.38	1.31	73	.84	.70
XS	6.47	2.14			
COMPOSIT	6.36	1.29	85	.86	.73
XGXS	6.42	1.83			
COMPOSIT	6.36	1.29	85	.84	.71
GSI	6.28	1.84			

* - See Table 3.9 for description of raters.

TABLE 3.28

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF LONG-RANGE PLANNING

RATER CODE*	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.37	1.31	80	.65	.42
G1	6.44	1.69			
COMPOSIT	6.19	1.31	57	.74	.54
G2	6.44	1.68			
COMPOSIT	6.45	1.24	60	.76	.58
S1	6.45	1.97			
COMPOSIT	6.21	1.31	50	.79	.62
S2	6.16	1.86			
COMPOSIT	6.39	1.34	76	.75	.56
I	6.55	1.70			
COMPOSIT	6.36	1.30	83	.84	.70
XGSI	6.47	1.51			
COMPOSIT	6.38	1.30	82	.72	.52
XG	6.52	1.59			
COMPOSIT	6.40	1.30	69	.83	.69
XS	6.41	1.77			
COMPOSIT	6.36	1.30	83	.83	.68
XGXS	6.44	1.58			
COMPOSIT	6.36	1.30	83	.79	.63
GSI	6.44	1.58			

* - See Table 3.9 for description of raters.

TABLE 3.29

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF SHORT-RANGE PLANNING

RATER CODE*	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.37	1.30	82	.71	.50
G1	6.67	1.59			
COMPOSIT	6.21	1.32	56	.70	.50
G2	6.46	1.85			
COMPOSIT	6.36	1.21	69	.61	.37
S1	6.84	1.62			
COMPOSIT	6.23	1.32	53	.83	.70
S2	6.55	1.79			
COMPOSIT	6.35	1.32	79	.69	.47
I	6.82	1.45			
COMPOSIT	6.35	1.29	84	.85	.72
XGSI	6.73	1.38			
COMPOSIT	6.37	1.29	83	.75	.56
XG	6.70	1.58			
COMPOSIT	6.37	1.30	75	.83	.68
XS	6.75	1.57			
COMPOSIT	6.35	1.29	84	.84	.70
XGXS	6.73	1.47			
COMPOSIT	6.35	1.29	84	.79	.62
GSI	6.75	1.34			

* - See Table 3.9 for description of raters.

TABLE 3.30

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF PROBLEM SOLVING

RATER CODE*	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.36	1.31	81	.66	.44
G1	6.41	1.65			
COMPOSIT	6.19	1.31	57	.80	.64
G2	6.14	1.83			
COMPOSIT	6.36	1.25	68	.67	.45
S1	6.34	1.99			
COMPOSIT	6.21	1.32	52	.80	.64
S2	6.33	1.90			
COMPOSIT	6.36	1.32	80	.76	.58
I	6.46	1.76			
COMPOSIT	6.36	1.29	85	.87	.75
XGSI	6.42	1.55			
COMPOSIT	6.37	1.29	83	.76	.58
XG	6.42	1.62			
COMPOSIT	6.35	1.31	75	.82	.66
XS	6.41	1.84			
COMPOSIT	6.36	1.29	85	.85	.72
XGXS	6.44	1.60			
COMPOSIT	6.36	1.29	85	.79	.63
GSI	6.41	1.57			

* - See Table 3.9 for description of raters.

TABLE 3.31

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF INTERPERSONAL RELATIONS

RATER CODE*	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.37	1.30	82	.65	.43
G1	6.17	1.83			
COMPOSIT	6.25	1.25	56	.69	.47
G2	6.30	1.73			
COMPOSIT	6.34	1.35	69	.76	.58
S1	6.22	2.09			
COMPOSIT	6.20	1.37	53	.74	.55
S2	6.26	2.14			
COMPOSIT	6.36	1.32	80	.78	.60
I	6.08	1.77			
COMPOSIT	6.36	1.29	85	.87	.76
XGSI	6.22	1.56			
COMPOSIT	6.37	1.29	83	.72	.51
XG	6.25	1.69			
COMPOSIT	6.38	1.32	77	.79	.63
XS	6.38	1.99			
COMPOSIT	6.36	1.29	85	.84	.71
XGXS	6.29	1.63			
COMPOSIT	6.36	1.29	85	.82	.66
GSI	6.13	1.65			

* - See Table 3.9 for description of raters.

TABLE 3.32

CORRELATION BETWEEN COMPOSITE SCORES AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF GOAL SETTING

RATER CODE*	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.37	1.31	80	.77	.59
G1	6.48	1.71			
COMPOSIT	6.23	1.29	56	.85	.72
G2	6.13	1.67			
COMPOSIT	6.41	1.24	64	.80	.63
S1	6.09	1.65			
COMPOSIT	6.20	1.35	54	.76	.57
S2	5.89	1.87			
COMPOSIT	6.37	1.30	77	.72	.51
I	6.27	1.70			
COMPOSIT	6.36	1.29	85	.93	.87
XGSI	6.25	1.44			
COMPOSIT	6.38	1.30	82	.84	.70
XG	6.41	1.62			
COMPOSIT	6.37	1.33	74	.82	.67
XS	6.12	1.75			
COMPOSIT	6.36	1.29	85	.91	.83
XGXS	6.28	1.52			
COMPOSIT	6.36	1.29	85	.86	.73
GSI	6.26	1.50			

* - See Table 3.9 for description of raters.

TABLE 3.33

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON THE BEHAVIORAL CATEGORY OF PERSONAL WORK STANDARDS

RATER CODE*	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.37	1.30	82	.77	.59
G1	6.54	1.75			
COMPOSIT	6.19	1.31	57	.76	.58
G2	6.46	1.81			
COMPOSIT	6.42	1.18	66	.69	.48
S1	6.61	1.78			
COMPOSIT	6.24	1.31	54	.85	.73
S2	6.31	1.97			
COMPOSIT	6.37	1.31	80	.76	.58
I	6.29	1.86			
COMPOSIT	6.36	1.29	85	.89	.80
XGSI	6.48	1.53			
COMPOSIT	6.37	1.29	83	.79	.63
XG	6.60	1.70			
COMPOSIT	6.42	1.28	72	.80	.65
XS	6.56	1.83			
COMPOSIT	6.36	1.29	85	.86	.73
XGXS	6.56	1.64			
COMPOSIT	6.36	1.29	85	.87	.75
GSI	6.44	1.54			

* - See Table 3.9 for description of raters.

TABLE 3.34

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G2, S2) ON THE CATEGORY OF OVERALL

RATER CODE*	MEAN	STD. DEV.	N	CORR.	r^2
COMPOSIT	6.19	1.31	57	.89	.79
G2	6.23	1.74			
COMPOSIT	6.17	1.35	52	.87	.76
S2	6.40	2.02			
COMPOSIT	6.28	1.32	67	.91	.82
XG2S2	6.41	1.81			

* - See Table 3.9 for description of raters.

TABLE 3.35

CORRELATION BETWEEN COMPOSITE SCORE AND TASK CATEGORY (T1) TEST AREAS

RATER CODE	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.37	1.30	82	.84	.71
G1	6.26	1.38			
COMPOSIT	6.20	1.30	58	.87	.76
G2	6.31	1.42			
COMPOSIT	6.32	1.30	75	.88	.78
S1	6.17	1.64			
COMPOSIT	6.23	1.35	55	.86	.74
S2	6.45	1.74			
COMPOSIT	6.36	1.31	81	.87	.76
I	6.06	1.48			
COMPOSIT	6.36	1.29	85	.97	.95
XGSI	6.29	1.30			
COMPOSIT	6.37	1.29	83	.89	.79
XG	6.36	1.33			
COMPOSIT	6.37	1.32	78	.90	.80
XS	6.37	1.60			
COMPOSIT	6.36	1.29	85	.96	.93
XGXS	6.36	1.32			
COMPOSIT	6.36	1.29	85	.95	.90
GSI	6.17	1.34			

* - See Table 3.9 for description of raters.

TABLE 3.36

CORRELATION BETWEEN COMPOSITE SCORE AND BEHAVIOR CATEGORY (T2) TEST AREAS

RATER CODE*	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.37	1.30	82	.83	.69
G1	6.44	1.42			
COMPOSIT	6.19	1.31	57	.86	.73
G2	6.30	1.57			
COMPOSIT	6.32	1.30	75	.83	.68
S1	6.40	1.58			
COMPOSIT	6.22	1.33	57	.87	.76
S2	6.28	1.81			
COMPOSIT	6.36	1.31	81	.86	.74
I	6.42	1.44			
COMPOSIT	6.36	1.29	85	.98	.95
XGSI	6.43	1.33			
COMPOSIT	6.37	1.29	83	.89	.79
XG	6.48	1.41			
COMPOSIT	6.37	1.32	78	.92	.84
XS	6.47	1.55			
COMPOSIT	6.36	1.29	85	.96	.92
XGXS	6.46	1.38			
COMPOSIT	6.36	1.29	85	.94	.89
GSI	6.42	1.30			

* - See Table 3.9 for description of raters.

TABLE 3.37

CORRELATION BETWEEN COMPOSITE SCORE AND COMBINED BEHAVIORAL AND TASK CATEGORY (T1T2) TEST AREAS

RATER CODE*	MEAN	STD. DEV.	N	CORR	r^2
COMPOSIT	6.37	1.30	82	.88	.77
G1	6.35	1.34			
COMPOSIT	6.20	1.30	58	.90	.80
G2	6.31	1.43			
COMPOSIT	6.32	1.30	75	.89	.80
S1	6.28	1.54			
COMPOSIT	6.22	1.33	57	.89	.79
S2	6.37	1.71			
COMPOSIT	6.36	1.31	81	.90	.82
I	6.24	1.40			
COMPOSIT	6.36	1.29	85	.97	.94
XG	6.36	1.27			
COMPOSIT	6.36	1.29	85	.95	.90
XS	6.41	1.30			
COMPOSIT	6.36	1.29	85	.97	.93
XGXS	6.39	1.28			
COMPOSIT	6.36	1.29	85	.98	.95
GSI	6.30	1.28			

* - See Table 3.9 for description of raters.

TABLE 3.38

CORRELATION BETWEEN COMPOSITE SCORE AND RATER GROUPS (G1, G2, S1, S2, I) ON BEHAVIOR AND TASK CATEGORY CRITERIA

CRITERIA CODE #	RATER CODE #									TOTAL*
	G1	G2	S1	S2	I	XG	XS	XGXS	GSI	
LRPLAN	.65	.74	.76	.79	.75	.72	.83	.83	.79	.84
SRPLAN	.71	.70	.61	.83	.69	.75	.83	.84	.79	.85
PSOLV	.66	.80	.67	.80	.76	.76	.81	.85	.79	.87
INTREL	.65	.69	.76	.74	.78	.72	.79	.84	.82	.87
GOALS	.77	.85	.80	.76	.72	.84	.82	.91	.86	.93
WRKSTD	.77	.76	.69	.85	.76	.79	.80	.86	.87	.89
T2	.83	.86	.83	.87	.86	.89	.92	.96	.94	.98
PROD	.78	.74	.81	.80	.83	.79	.85	.88	.89	.90
SAFETY	.53	.55	.62	.63	.60	.54	.63	.68	.68	.70
MAINT	.60	.67	.66	.73	.66	.66	.73	.77	.70	.78
MOTIV	.64	.77	.73	.77	.67	.75	.79	.84	.76	.84
MGTATT	.71	.73	.78	.84	.77	.78	.84	.86	.84	.87
T1	.84	.87	.88	.86	.87	.89	.90	.96	.95	.97
T1T2	.88	.90	.89	.89	.90	.97	.95	.97	.98	1.00
OVERAL	-	.89	-	.87	-	-	-	-	-	.91**

- See Table 3.9 for description of criteria and raters.

* - G1 + G2 + S1 + S2 + I (XGSI)

** - G2 + S2 (XG2S2)

TABLE 3.39

RANK ORDER OF CORRELATIONS BETWEEN COMPOSITE SCORES AND RATER GROUPS AND CRITERIA/TEST AREA COMBINATIONS

#	RATER CODE	CRITERIA/TEST AREA CODE	CORR.	r^2
1	XGSI	COMPOSIT	1.00	1.00
2	XGSI	BEHAV	.98	.95
3	GSI	T1T2	.98	.95
4	XGSI	TASK	.97	.95
5	XG	T1T2	.97	.94
6	XGXS	T1T2	.97	.93
7	XGXS	TASK	.96	.93
8	XGXS	BEHAV	.96	.92
9	XS	T1T2	.95	.90
10	GSI	TASK	.95	.90
11	GSI	BEHAV	.94	.89
12	XGSI	GOALS	.93	.87
13	XS	BEHAV	.92	.84
14	XGXS	GOALS	.91	.83
15	XG2S2	OVERAL	.91	.82
16	I	T1T2	.90	.82
17	XS	TASK	.90	.80
18	XGSI	PROD	.90	.80
19	G2	T1T2	.90	.80
20	S1	T1T2	.89	.80
21	XGSI	WRKSTD	.89	.80
22	XG	TASK	.89	.79
23	G2	OVERAL	.89	.79
24	GSI	PROD	.89	.79
25	S2	T1T2	.89	.79
26	XG	BEHAV	.89	.79
27	SI	TASK	.88	.78
28	XGXS	PROD	.88	.77
29	G1	T1T2	.88	.77
30	XGSI	INTREL	.87	.76
31	I	TASK	.87	.76
32	G2	TASK	.87	.76
33	S2	OVERAL	.87	.76
34	S2	BEHAV	.87	.76
35	GSI	WRKSTD	.87	.75
36	XGSI	MGTATT	.87	.75
37	XGSI	PSOLV	.87	.75
38	S2	TASK	.86	.74
39	I	BEHAV	.86	.74
40	G2	BEHAV	.86	.73
41	XGXS	WRKSTD	.86	.73
42	GSI	GOALS	.86	.73
43	XGXS	MGTATT	.86	.73
44	S2	WRKSTD	.85	.73
45	G2	GOALS	.85	.72
46	XS	PROD	.85	.72
47	XGXS	PSOLV	.85	.72
48	XGSI	SRPLAN	.85	.72
49	G1	TASK	.84	.71
50	GSI	MGTATT	.84	.71
51	S2	MGTATT	.84	.71
52	XGXS	MOTIV	.84	.71
53	XGXS	INTREL	.84	.71
54	XGSI	MOTIV	.84	.70
55	XGXS	SRPLAN	.84	.70
56	XS	MGTATT	.84	.70
57	XGSI	LRNPLAN	.84	.70
58	XG	GOALS	.84	.70

(Table 3.39 CONTINUED)

NUMBER	RATER CODE*	CRITERIA/TEST AREA CODE*	CORR.	r^2
59	S2	SRPLAN	.83	.70
60	G1	BEHAV	.83	.69
61	XS	LRPLAN	.83	.69
62	I	PROD	.83	.69
63	S1	BEHAV	.83	.68
64	XS	SRPLAN	.83	.68
65	XGXS	LRPLAN	.83	.68
66	XS	GOALS	.82	.67
67	GSI	INTREL	.82	.66
68	XS	PSOLV	.81	.66
69	S1	PROD	.81	.66
70	XS	WRKSTD	.80	.65
71	S2	PROD	.80	.65
72	S2	PSOLV	.80	.64
73	G2	PSOLV	.80	.64
74	S1	GOALS	.80	.63
75	XG	PROD	.79	.63
76	XS	INTREL	.79	.63
77	XG	WORKSTD	.79	.63
78	GSI	LRPLAN	.79	.63
79	GSI	PSOLV	.79	.62
80	GSI	SRPLAN	.79	.62
81	S2	LRPLAN	.79	.62
82	XS	MOTIV	.79	.62
83	S1	MGTATT	.78	.61
84	G1	PROD	.78	.60
85	XG	MGTATT	.78	.60
86	I	INTREL	.78	.60
87	XGSI	MAINT	.78	.60
88	I	MGTATT	.77	.60
89	XGXS	MAINT	.77	.59
90	G1	WRKSTD	.77	.59
91	G2	MOTIV	.77	.59
92	S2	MOTIV	.77	.59
93	G1	GOALS	.77	.59
94	G2	WRKSTD	.76	.58
95	S1	LRPLAN	.76	.58
96	I	WRKSTD	.76	.58
97	I	PSOLV	.76	.58
98	XG	PSOLV	.76	.58
99	S1	INTREL	.76	.58
100	S2	GOALS	.76	.57
101	GSI	MOTIV	.76	.57
102	XG	MOTIV	.75	.56
103	XG	SRPLAN	.75	.56
104	I	LRPLAN	.75	.56
105	G2	PROD	.74	.55
106	S2	INTREL	.74	.55
107	G2	LRPLAN	.74	.54
108	S2	MAINT	.73	.54
109	G2	MGTATT	.73	.54
110	XS	MAINT	.73	.53
111	S1	MOTIV	.73	.53
112	XG	LRPLAN	.72	.52
113	XG	INTREL	.72	.51
114	I	GOALS	.72	.51

(Table 3.39 CONTINUED)

NUMBER	RATER CODE*	CRITERIA/TEST AREA CODE*	CORR.	r^2
115	G1	MGTATT	.71	.50
116	G1	SRPLAN	.71	.50
117	G2	SRPLAN	.70	.50
118	GSI	MAINT	.70	.49
119	XGSI	SAFETY	.70	.49
120	S1	WRKSTD	.69	.48
121	I	SRPLAN	.69	.47
122	G2	INTREL	.69	.47
123	XGXS	SAFETY	.68	.46
124	GSI	SAFETY	.68	.46
125	S1	PSOLV	.67	.45
126	G2	MAINT	.67	.45
127	I	MOTIV	.67	.45
128	G1	PSOLV	.66	.44
129	S1	MAINT	.66	.44
130	XG	MAINT	.66	.44
131	I	MAINT	.66	.43
132	G1	INTREL	.65	.43
133	G1	LRPLAN	.65	.42
134	G1	MOTIV	.64	.41
135	S2	SAFETY	.63	.40
136	XS	SAFETY	.63	.40
137	S1	SAFETY	.62	.39
138	S1	SRPLAN	.61	.37
139	I	SAFETY	.60	.36
140	G1	MAINT	.60	.36
141	G2	SAFETY	.55	.30
142	XG	SAFETY	.54	.29
143	G1	SAFETY	.53	.28

* - See Table 3.9 for description of criteria/test areas/raters.

The highest total correlations by criteria category were .98 (XGSI: T2 and GSI: T1T2), followed by .97 (XGSI: T1, XG: T1T2, XGXS: T1T2). The highest correlation for a criterion was .93 (XGSI: GOALS), followed by .91 (XGXS: GOALS and XG2S2: OVERAL). The low highest correlation was .70 (XGSI: SAFETY). The highest number of correlations by criteria were all multirater combinations - XGSI (12), XGXS (1), XG2S2 (1) and GSI (1).

The highest individual correlation by criteria category was .90 (I: T1T2, followed by .90 (G2: T1T2) and .89 (S1: T1T2, G2: OVERAL, S2: T1T2). The highest individual rater correlation for a criterion was .89 (G2: OVERAL) followed by .87 (S2: OVERAL) and .85 (S2: WRKSTD, G2: GOALS). The low highest individual correlation was .63 (S2: SAFETY). The raters with the highest number of individual correlations by criteria were S2 (8), followed by G2 (3), I (3) and S1 (1).

The lowest total correlation by criteria category was .53 (G1: SAFETY), the next lowest criteria being .60 (G1: MAINT). The lowest correlation for a test area was .83 (S1: T2). The high lowest correlation was .88 (G1: T1T2) followed by .87 (S2: OVERAL). The raters with the lowest number of correlations were all individual raters - G1 (9), S1 (3), G2 (1), S2 (1), and I (1).

The following chart outlines the significant high/low correlation information by criterion, Chart "C."

By utilization of the highest positive mean of the coefficient of correlation as a basis, it was determined that the best criterion was OVERAL, the best test area was T1T2, and the best rater/rater combination was XG2S2. The rank order of all the criteria, test areas and raters is as indicated in Chart "D."

The exclusion of OVERAL data from the analysis, because the criterion was limited to two raters, would make the best criterion Productivity, the best test area would be T1T2, and the best rater/rater combination would be XGSI.

In this study, a coefficient of correlation of .75 or above was considered significant and to have met requirements of consistency. Utilizing this rationale, only two criteria (MAINT - .70 and SAFETY - .62) and one rater (G1 - .72) would be considered areas of concern and/or exclusion.

CHART "C"

HIGH/LOW CORRELATION SUMMARY

CRITERIA CODE*	CORRELATION								
	HIGHEST TOTAL			HIGHEST INDIVIDUAL			LOWEST TOTAL		
	CORR.	RATER	RANK	CORR.	RATER	RANK	CORR.	RATER	RANK
PROD	.90	XGSI	18	.83	I	62	.74	G2	105
SAFETY	.70	XGSI	119	.63	S2	135	.53	G1	143
MAINT	.78	XGSI	87	.73	S2	108	.60	G1	140
MOTIV	.84	XGXS	52	.77	G2	91	.64	G1	134
MGTATT	.87	XGSI	36	.51	S2	51	.71	G1	115
T1	.97	XGSI	4	.88	I	27	.84	G1	49
LRPLAN	.84	XGSI	57	.79	S2	81	.65	G1	133
SRPLAN	.85	XGSI	48	.83	S2	59	.61	S1	138
PSOLV	.87	XGSI	37	.80	S2	72	.66	G1	128
INTREL	.87	XGSI	30	.78	I	86	.65	G1	132
GOALS	.93	XGSI	12	.85	G2	45	.72	I	114
WRKSTD	.89	XGSI	21	.85	S2	44	.69	S1	120
T2	.98	XGSI	2	.87	S2	34	.83	S1	63
T1T2	.98	GSI	3	.90	I	16	.88	G1	29
OVERAL	.91	XG2S2	15	.89	G2	23	.87	S2	33

* - See Table 3.9 for description of criteria. Also see Table 3.39.

CHART "D"

RANK ORDER OF CRITERIA, RATERS, AND TEST AREAS

RANK ORDER	CRITERIA CODE*	CRITERIA CORR.	RATER CODE*	RATER CORR.	TEST AREA CODE*	TEST AREA CORR.
1	OVERAL	.89	XG2S2	.91@	T1T2	.93
2	PROD	.83	XGSI	.88	T1	.90
3	GOALS	.83	XGXS	.86	T2	.89
4	WORKSTD	.80	GSI	.83		
5	MGTATT	.80	XS	.82		
6	PSOLV	.78	XG	.81		
7	LRPLAN	.77	I	.80		
8	INTREL	.77	S2	.80@		
9	SRPLAN	.76	S2	.80		
10	MOTIV	.76	G2	.77@		
11	MAINT	.70	G2	.76		
12	SAFETY	.62	S1	.75		
13			G1	.72		

* - See Table 3.9 for description of criteria/test area/rater.

@ - OVERAL included.

To determine whether a difference in initial and final responses existed, the following matrix was utilized.

TEST AREA	RATERS G1=G2	RATERS S1=S2
T1	X	X
T2	X	X
T1T2	X	X

The coefficient of correlation and the coefficient of determination between the raters and test areas were calculated and are presented on the tables that follow (Tables 3.40 and 3.41).

The coefficient of determination for direct supervision was .62 for the task category test area and .55 for the behavior category test area for an overall variance of .07 between the task and behavior category test areas.

The coefficient of determination for indirect supervision was .66 for the task category test area and .40 for the behavior category test area for an overall variance of .26 between the task and behavior test areas.

supervision and indirect supervision in the assessment of job-related criteria on subordinate members of the organization if an identical instrument is administered on two separate occasions.

Results

In the study, direct supervisors and indirect supervisors were requested to test-retest the selected foreman sample utilizing an identical instrument at pre-determined time intervals.

The instruments utilized were identical except the final instrument had one additional criterion - a global criterion "Overall Performance." The global criterion was also rated on a nine-point scale, which ranged from a maximum of "Usually Better" (9) to a minimum of "Usually Worse" (1).

For statistical analysis, the Overall Performance rating was reported and evaluated separately - unless specifically noted.

To determine whether a difference in initial and final responses existed, the following matrix was utilized.

TEST AREA	RATERS GF1 = GF2	S1 = S2
T1	X	X
T2	X	X
TX	X	X

The coefficient of correlation and the coefficient of determination between the raters and test areas were calculated and are presented on the tables that follow (Tables 3.40 and 3.41).

The coefficient of determination for direct supervision was .62 for the task category test area and .55 for the behavior category test area for an overall variance of .07 between the task and behavior category test areas.

The coefficient of determination for indirect supervision was .66 for the task category test area and .40 for the behavior category test area for an overall variance of .26 between the task and behavior test areas.

TABLE 3.40

CORRELATION BETWEEN INITIAL AND FINAL SCORES ON TEST AREA CATEGORIES BY DIRECT SUPERVISION (G1, G2)

RATER: TEST AREA CODE*	MEAN	STD. DEV.	N	CORR.	r^2
G1: TASK (T1)	6.01	1.36	57	.79	.62
G2: TASK (T1)	6.31	1.43			
G1: BEHAV(T2)	6.18	1.38	56	.74	.55
G2: BEHAV(T2)	6.31	1.58			
G1: T1T2	6.09	1.31	57	.83	.68
G2: T1T2	6.31	1.45			

* - See Table 3.9 for description of test area and raters.

TABLE 3.41

CORRELATION BETWEEN INITIAL AND FINAL SCORES ON TEST AREA CATEGORIES BY INDIRECT SUPERVISION (S1, S2)

RATER: TEST AREA CODE*	MEAN	STD. DEV.	N	CORR.	r^2
S1: TASK (T1)	5.98	1.67	52	.81	.66
S2: TASK (T1)	6.35	1.73			
S1: BEHAV(T2)	6.15	1.57	54	.63	.40
S2: BEHAV(T2)	6.18	1.80			
S1: T1T2	6.05	1.53	54	.76	.58
S2: T1T2	6.27	1.70			

* - See Table 3.9 for description of test areas and raters.

The correlations ranged from a high of .83 (G1/G2: T1T2) to a low of .63 (S1/S2: T2).

Summary of Hypothesis #2

Significant difference was noted between the coefficient of correlation and coefficient of determination obtained from ratings by indirect supervision between the task category and behavior category test areas; therefore, the null hypothesis was rejected for this rater. However, no significant difference was apparent from data obtained between the task category and behavior category test areas for direct supervision and the null hypothesis could not be rejected.

SUMMARY

Hypothesis #1, stating no difference would exist in how raters from different hierarchical levels within the organization appraise job-related criteria on subordinate members of the organization, was rejected.

Hypothesis #2, stating no difference would exist in the test-retest responses of direct supervision and

indirect supervision in the appraisal of subordinate members of the organization, was rejected for indirect supervision, but could not be rejected for direct supervision.

CHAPTER IV

SUMMARY, FINDINGS, IMPLICATIONS AND RECOMMENDATIONS

SUMMARY

The purpose of this study was to examine a formal multitrait-multirater appraisal of managerial job performance in an industrial environment. Particular interest was focused on determining whether ratings assigned by personnel in supervisory positions are predictable and, more specifically, determining whether direct supervision, indirect supervision, non-related supervision, or a combination thereof, provides the most accurate and reliable performance appraisal.

The sample consisted of ninety foremen located in eight locations within an industrial organization. An instrument was designed to obtain data on job-related categories and was administered by individuals who were direct supervisors, indirect supervisors and non-related supervisors.

The instrument, which used a nine-point rating scale format, was considered sufficient to obtain the desired data. The task and behavioral test area data obtained were considered adequate for treatment. The hard-core test area data were not utilized, primarily because of the diverse sample distribution and availability of data.

Analysis consisted of number and percentage of responses by cell as well as coefficients of correlation and determination.

FINDINGS

Two major hypotheses were tested. The findings were as follows:

RATER COMPARISONS

1. An analysis of the data indicated that raters from different hierarchical levels within an organization appraise subordinate individuals differently. In addition, a difference existed when combinations of raters were compared.

2. Comparison of individual raters and the combinations of raters with the composite score indicated that the overall best correlation was a two rater combination comprised of direct supervision - final appraisal and indirect supervision - final appraisal.

3. The highest ranked single rater was the non-related supervisor, followed closely by the indirect supervisor - final appraisal and, in descending order, direct supervision - final appraisal, indirect supervision - initial appraisal and direct supervision - initial appraisal. A complete listing of the rank order of all raters/rater combinations is provided in Chapter III.

RATER RELIABILITY

Analysis of the data revealed indirect supervisor inconsistencies between initial and final appraisals; therefore, reliability for the rater group is questionable. The behavior category test data were not consistent with the task category or the combined behavior-task test areas. Indirect supervisors were not consistent from first to second ratings, nor were

they consistent between behavior and test scores. No significant differences were found in the scores of direct supervisors. They were consistent in their first and second ratings and behavior and task scores were highly correlated.

IMPLICATIONS

This study supported the notion that appraisal of managerial personnel is a complex matter and, in this particular instance, revealed that a review of current appraisal procedures within the "Firm" should be conducted. The findings of this study may have significant meaning concerning conservation of manpower, time and fiscal resources to the "Firm."

Although this study applies directly to one organization in a given environment during a specific time-frame, the findings have implications that apply not only to the "Firm" and similar organizations, but to organizations in other environments.

First, the findings of this study should be applicable

to most large organizations. The best rater/rater combination, discussed previously, should apply. Similarly, the Overall Performance was found to be the best test indicator and should be noted for reference in further exploration of the subject.

The next highest test criterion was Productivity, followed in descending order by Goal Setting, Personal Work Standards, Management Attitude, Problem-Solving, Long-Range Planning, Interpersonal Relations, Short-Range Planning, Motivating the Workforce, Maintenance, and Safety.

Assuming that human and fiscal resources are not limited, the selection of the optimum combination of criteria/test area and rater is possible. In this instance, the best criteria/test area and rater is the combination of the task and behavior category test areas with a five rater combination comprised of direct supervision - initial and final appraisal, indirect supervision - initial and final appraisal and non-related supervision. However, that was to be expected as the "combined" score was used as an optimum.

While an absolute measure may never be possible, the fact that differences between raters and test areas were found, indicates the need for information on which to base decisions, if resources are limited. The selection of the optimum combination of criteria/test area and rater was based on a correlation analysis with the composite variable in which the higher the combination, the more favorable the relationship. Several implications are drawn from the results obtained.

The multirater concept by virtue of having more than one rater and input, helps eliminate bias and prejudice in the appraisal. This implies that the more measures that are available, the more likelihood of an accurate appraisal. Although this may not exist in every instance, the results of this study indicate that multi-rater combinations had higher correlations with the composite than did single rater scores. It should be pointed out that the inclusion of each rater or test area in the composite against which all comparisons were made may have had slight effects on outcomes but, because numbers were comparable, the effects should be relatively insignificant.

If a high correlation exists between variables, it is assumed that the measurement instruments and procedures possess a reasonable degree of reliability. In this study, the data support this rationale; but a direct cause-effect relationship is not guaranteed.

The high correlation also indicates that the supervision for the "Firm" apparently based their appraisal ratings on similar observations of performance and judged the various categories according to similar standards.

The use of a rating system in the study has several interesting implications:

1. The resultant correlation analysis implied the system utilized is consistent with logic that with experience comes improvement. In three-fourths of the instances, the correlation increased where test-retest methodology was utilized.

2. The direct interest in the administration of the instrument was apparently favorably perceived by the raters. The implication is that personal attention and training of the raters may be significant variables. That is, in any organization, time expended to train raters prior to their completing performance appraisals may provide results that are more fair, accurate, reliable, and subsequently, more acceptable and desirable.

When either manpower, time limitations, or fiscal constraints are imposed which which prevent initiation of an appraisal program of the magnitude previously discussed, it becomes necessary to make sound judgments about including or eliminating raters or tests. No differentiation between initial and final appraisals is assumed.

The results of the study indicate that, if only one rater is used, the non-related supervisor would be the best rater, followed by the indirect supervisor and,

finally, the direct supervisor. The results indicate that the combined task and behavior areas are the best, followed by the task area and then the behavior area.

The optimum criterion, given only one rater, was found to be the direct or indirect supervisor on the criterion Overall Performance. However, for the non-related supervisor, Productivity was found to be the optimum criterion.

The top three criteria, given only one rater, were found to be, in order, as indicated. Direct Supervisor - Overall Performance, Goal Setting and Problem Solving - Indirect Supervisor - Overall Performance, Personal Work Standards, and Productivity. Non-related Supervisor - Productivity, Management Attitude, and Interpersonal Relations.

An overall analysis of the data revealed that direct and indirect supervision tended to stress production-related variables over personnel-related variables, while the non-related supervisor tended to

stress personnel-related variables. The implication is that the closer the proximity to the production-related areas, the more the emphasis. In addition, the difference in perspective concerning the job, the organization and the mission is evident.

There was a tendency for raters to be more liberal when rating the global criterion than specific criterion. This occurs because, typically, most existing appraisals are global. The reasoning for this is not clear. However, this fact must be considered in the appraisal of personnel when non-global criteria are used.

As a result of statistical analysis, several criteria could be eliminated from the appraisal. From the statistical analysis, two task category criteria are candidates for removal - Maintenance and Safety. One implication might be the terminology. To a foreman, these criteria are typically not considered influential to the outcome. Other implications might include reasoning that the rater may be apathetic toward the criterion (particularly Safety) and thus, rates the criteria accordingly. Or perhaps the criterion is seen as a hindrance to production.

Analysis of the data provided information that indicates that the rater who is organizationally closest to the ratee, is not best able to appraise the individual on the criteria utilized.

Implication concerning this finding may be the difference in perspective of the direct supervisor toward the individual based on more intimate knowledge and acquaintance with the ratee, or need for more extensive experience and knowledge within the organization and in completing performance appraisals, or a difference in perspective with respect to organizational goals and mission.

In determining which combination of raters and test criteria is best for any given organization, the first consideration must be resources. As a result of this study, it seems likely that multiple raters and multiple tests are best. But when resources are limited, it would be best to reduce raters (the more costly) first. Second echelon (indirect supervisors) seem to be the "best" raters and, therefore, should be included in the process over and above any other rater.

Task tests are "better" than behavior tests and should be included, if possible.

Organizations should vary rater and test groups by utilizing the information presented in Chapter III. Cost and outcome should always be considered. A change in correlation of .02 points may be insignificant or undesirable, if costs rise sharply as a result. In summary, the final decision as to performance assessment methodology should be left with the organization, so long as research and available resources are considered in proper perspective.

RECOMMENDATIONS

This study provided only a beginning. The following studies and activities are warranted:

1. An annual, or periodic, multitrait-multirater appraisal utilizing the same or similar variables.

2. A study concerning the effects on the results of this study with modifications as follows:

- Utilize the global criterion - Overall Performance - with all raters.

- Utilize test-retest methodology with non-related supervision.

- Consider demographic information of the raters.

3. A study researching the influence rater training has on appraisals.

4. A study that expands the sample to include peers and self-appraisal input.

5. A comparative cost-benefit study that explores the financial implications of multiple raters.

6. An extension of this study, eliminating selected criteria and/or utilizing different criteria.

7. A factor analysis of the criteria in this study.

APPENDIX A
PERFORMANCE-INCIDENT INSTRUMENT

FOREMAN SELECTION SYSTEM

We are asking you to assist this division in analyzing the job of foreman with the objective of improving our system for selecting the best quality foreman. To accomplish this objective, we would like you to discuss specific performance incidents associated with effective or ineffective performance on the following sheets.

TIPS ON COMPLETING PERFORMANCE INCIDENT RECORDS

1. You may want to sit and think about several incidents first and jot down a few key words on each to help you write them out later.

2. It may help you to think of one particular individual who you believe is highly effective or ineffective and think of a performance incident concerning him.

3. For the most part, incidents may actually represent the range of normal activities that are important aspects of the foreman's job.

4. Report only actual behavior that you have observed directly.

5. Report <u>all</u> relevant factors in the situation.

6. Be specific without excessive detail, be short and to the point.

7. Work individually.

8. Try to have as many effective incidents as you do ineffective.

9. Work as rapidly as you can, at your own pace.

10. We would like to average a total of 10 to 15 incidents from each person. Do more or less depending on the quality of the incidents.

11. <u>Do not</u> mention the name of the individual involved in the incident.

12. Be careful to report only one incident at a time.

13. If you have any questions about your incidents, ask the interviewer.

14. Please write legibly.

PERFORMANCE INCIDENT RECORD - FOREMAN
SELECTION SYSTEM PROJECT

Think of the last time you saw a foreman perform an activity required of his job that made you think of him as an effective or ineffective foreman.

In other words, think of a performance incident which added to or detracted from the successful completion of his job.

Was this an ___ effective or an ___ ineffective performance incident?

1. What were the circumstances leading up to the incident?

_____

_____

_____

2. Tell specifically what the foreman did that made him effective or ineffective.

_____

_____

_____

3. What were the results (outcome) of what the foreman did?

_____

_____

_____

Please complete the following items concerning the foreman in the incident above:

1. Production ___, non-production ___.

1a. If non-production, department name _____.

2 . Plant _____.

3. Division _____.

4. Time in position _____.

5. ____ Shift #1.

____ Shift #2.

____ Shift #3.

SOURCE: Research Staff of the "Firm."

APPENDIX B
HIGH-LOW RANKING FORM

HIGH-LOW RANKING FORM

CONFIDENTIAL

To be used for research purposes only.

RATER'S SIGNATURE_____ POSITION_____

DIVISION_____ PLANT_____ SHIFT_____

DATE_____

INSTRUCTIONS FOR COMPLETING THE HIGH-LOW RANKING FORM

1. List all General Foremen (first initial and last name only) that you supervise in the lefthand column of the opposite page.

2. In the second column, list all Foremen (first initial and last name only) who report to each of your General Foremen.

3. From the Foreman List, select the person whom you consider as ranking highest in overall performance. When determining overall performance, factors such as productivity, planning, controlling the workforce, interpersonal skills, etc., should be considered in evaluating each Foreman. Write his/her name on the top line (labeled "1-Highest") of the Rank Order column and cross the person's name off the Foreman list.

4. From the remaining Foreman names, select the person you feel ranks lowest, and write his/her name on the bottom of the Rank Order column where it says "1-Lowest." Now, cross his/her name off the list.

5. Continue by alternately selecting the next highest and the next lowest individuals remaining on the Foreman list until all names have been crossed out. For example if you have listed 12 Foremen, six of them should appear at the top of the Rank Order column, and six at the bottom. The middle space would be left blank.

6. If you are not sufficiently familiar with a Foreman's performance to rank him/her (new employee, etc.), or the individual's work is so unique that comparison is impossible, it is not necessary to record the individual in the Rank Order column.

   The fact that you rank an individual as lowest does not necessarily mean that you regard this person's performance as unsatisfactory. Your rating merely denotes how you would rank him/her in relation to others in the particular group.

7. In the "Sex/Race" column, please identify the sex and race of each of the Foremen by using the following codes:

	MALE	FEMALE
WHITE	1	3
NEGRO	2	4
ORIENTAL	5	6
SPANISH SURNAMED AMERICAN	7	8
AMERICAN INDIAN	0	9

8. For each Foreman, please indicate the approximate length of time the employee has worked under your supervision as a Foreman in the "Time On Present Position" column. The period of time should be recorded in the following manner:

   2 years, 10 months = 2-10
   10 years, 11 months =10-11
   6 months = 0-6

9. Please indicate the approximate length of time the employee has been classified as a Foreman in the last column. The time should be recorded in the same manner as requested in the "Time On Present Position" column.

SOURCE: Research Staff of the "Firm."

(Please Print)

HIGH-LOW RANKING FORM

Overall Performance

General Foreman List	Foreman List	Rank Order	Seniority/Price Code	Time on Price Pos.	Total Time as Foreman
		1 Highest			
		2 Next Highest			
		3 - Next Highest			
		4 - Next Highest			
		5 - Next Highest			
		6 - Next Highest			
		7 - Next Highest			
		8 - Next Highest			
		9 - Next Highest			
		10 - Next Highest			
		11 - Next Highest			
		12 - Next Highest			
		13 - Next Highest			
		14 - Next Highest			
		15 - Next Highest			
		16 - Next Highest			
		17 - Next Highest			
		18 - Next Highest			
		18 - Next Lowest			
		17 - Next Lowest			
		16 - Next Lowest			
		15 - Next Lowest			
		14 - Next Lowest			
		13 - Next Lowest			
		12 - Next Lowest			
		11 - Next Lowest			
		10 - Next Lowest			
		9 - Next Lowest			
		8 - Next Lowest			
		7 - Next Lowest			
		6 - Next Lowest			
		5 - Next Lowest			
		4 - Next Lowest			
		3 - Next Lowest			
		2 - Next Lowest			
		1 - Lowest			

APPENDIX C
"HARD-CORE" DATA - SUBJECTIVE RATING

FOREMAN _____ GENERAL FOREMAN _____

DIVISION #_____ SHIFT #_ _____ LINE #_____ WEEK ENDING_____ SEQUENCE #_____

	TOTAL LOAD	TOTAL THRU PAY POINT	STD.	TOTAL D.L. HOURS	TOTAL EARNED HOURS	VARIANCE + (-)	REJECTS	FOREMAN ABSENT YES/NO	HOURLY ABSENT ASSIGNED	GRIEVANCE	DISCIPLINARY ACTION REPORT
MONDAY											
TUESDAY											
WEDNESDAY											
THURSDAY											
FRIDAY											
SATURDAY											
SUNDAY											

Consider all of the factors which have influenced each of the above items this week. Evaluating only those factors which the foreman had control over, how would you rate his performance for the week on each of the following (circle the number which represents his performance):

Compared with other foremen performing this job in the past, his performance on this indicator was:

	ABOVE AVERAGE	SLIGHTLY ABOVE AVERAGE	AVERAGE	SLIGHTLY BELOW AVERAGE	BELOW AVERAGE
A PRODUCTIVITY	5	4	3	2	1

COMMENT: _____

Compared with other foremen performing this job in the past, his performance on this indicator was:

	ABOVE AVERAGE	SLIGHTLY ABOVE AVERAGE	AVERAGE	SLIGHTLY BELOW AVERAGE	BELOW AVERAGE
B DIRECT LABOR VARIANCE	5	4	3	2	1

COMMENT: _____

Compared with other foremen performing this job in the past, his performance on this indicator was:

	ABOVE AVERAGE	SLIGHTLY ABOVE AVERAGE	AVERAGE	SLIGHTLY BELOW AVERAGE	BELOW AVERAGE
C REJECT	5	4	3	2	1

COMMENT: _____

D FOREMAN ABSENT

Reason for Absence(S) _____

Compared with other foremen performing this job in the past, his performance on this indicator was:

	ABOVE AVERAGE	SLIGHTLY ABOVE AVERAGE	AVERAGE	SLIGHTLY BELOW AVERAGE	BELOW AVERAGE
E HOURLY ABSENTEEISM	5	4	3	2	1

COMMENT: _____

APPENDIX D

JOB DUTIES AND JOB SKILLS SUBJECTIVE RATING

CONFIDENTIAL: FOR RESEARCH PURPOSES ONLY

FOREMAN SELECTION SYSTEM
SUBJECTIVE RATING FORMS

Criteria Pilot Test

Please complete one booklet for each foreman being rated.

FOREMAN NAME _____
SHIFT_____ CLASSIFICATION/TITLE _____
RATER NAME _____
LEVEL: 1 2
TITLE _____
PLANT _____
CODE _____

This rating booklet is divided into two parts. Make sure that you understand the instructions for each part before proceeding. You will be given assistance in completing the ratings.

DO NOT WORK AHEAD OF THE SESSION LEADER

GENERAL INSTRUCTIONS

Purpose

These ratings will be used, along with additional "hard data" performance measures, to help Personnel Research determine which foreman performance indicators are valuable enough to be used in tracking the performance of trainee foremen hired through the new Foreman Selection System.
The ratings which you provide on the following pages will be used only for this purpose and will not be made available to any one else.

How to Complete the Scales

Described on each of the following pages is an area of job performance which is important to the job of foreman in this division. After reading the description of the performance area, you are to rate the foreman that you have been tracking for the past one or two months.

Complete the rating as follows:

- Read the definition of the performance area at the top of the page. Do not assume you know what the area is by just reading the title. The area may be defined differently by different people.

- Next read the general examples of the type of things that a foreman might do in that performance area. These items have been included to help you pin point specific things the foreman might have done in each performance area. These items are only representative of the factors that make up the area and you may wish to consider other factors under the performance area definition as well. Think of actual examples of similar things you have seen the foreman you are rating do.

- As you think of these actual performance examples, jot down one or two on the back side of the page. Be sure to give all of the important facts of the example, but do not include unnecessary detail.

- Finally, complete the rating by circling the number on the scale which best represents the foreman's typical performance in relation to foremen you have known in the past in similar jobs. Use the tips on the following page to help you avoid some of the common errors that are made in rating.

RATING TIPS

- Common Errors in Rating: Most raters are trying to do the best job possible. However, there are some common traps that raters tend to fall into. By being aware of them, you may be able to avoid them.

  1. Halo - Frequently raters judge individual aspects of a person's performance in terms of an overall impression they have of him. Very few people do everything equally well or poorly. Just because a man is good in one aspect of his performance, does not mean he is good in all aspects. Try not to let your general good or bad feelings about the foreman sway your rating on any one performance area. Be sure to rate the foreman's performance in each of the performance areas independently of his performance in other areas.

  2. Logical Error - Many raters tend to see one or more of the performance areas that follow as related and on that basis alone, assign the same rating to two or more performance areas. Each of the performance areas is sufficiently distinct from all the others that it must be considered independently. Again, consider the foreman's performance in each performance area independently.

  3. Leniency Error - Some raters have a tendency to rate a person by how well they know him. An "easy rater" will tend to give higher ratings to some one who is a friend or close fellow worker. A "hard rater," on the other hand, may try to compensate by rating people lower than they should. A third group of raters tend to be very hard raters or very easy raters regardless of how well they know the individual. This error can best be handled by being as objective as possible in considering specific examples of the foreman's performance. Compare actual things you have seen the foreman do with actual things you have seen other foremen do in similar situations.

  4. Central Tendency - Unlike leniency error, some raters hesitate to give extreme good or bad judgements about an individual and place their raters near the mid point of the scale. If an individual is about the same as other foremen you have known, rate him that way. However, if he is slightly above or below other foremen you have known, be sure to indicate the degree of that difference on the scale.

- Work on only one scale at a time. This will make it easier to avoid the errors mentioned above.

- Use actual, concrete examples of the foreman's typical performance to rate him, not what you feel about him generally. Objectivity is important.

- If you have a performance example that you are not sure fits the performance area, ask the trainer and he will help you judge if it should be included in rating the foreman's performance in that area.

- If you do not think you have enough information about a specific performance area to rate the foreman, leave that area unrated.

PART I

This part consists of ten pages representing five categories of foreman job duties:

1. Productivity
2. Safety and Housekeeping
3. Maintenance
4. Motivating the Workforce
5. Management Attitude

Read the description of each performance category thoroughly and list at least one or two examples of typical performance for that foreman before completing the rating.

CA

Code: _____

PRODUCTIVITY

Consistently meeting assigned work goals; controlling for and maintaining acceptable quality levels; efficiently controlling labor costs by assigning the "right people" to jobs and providing employes with adequate job instructions and directions; controlling material costs by holding damage and waste to a minimum.

In rating your foreman, consider the following:

- Does he typically meet his work goals? Does he occasionally exceed them or fall behind in meeting his work goals?
- When faced with the need to devise a fast fix to keep a job going, can he do so? Is he usually successful?
- Does he keep track of the skills of his employes and assign them to the jobs they are best suited for?
- If an equipment failure or material shortage occurs, do his men stand around or does he use them efficiently (e.g., reassign, send to early lunch, etc.)?
- Does he control for quality or does he tend to let poor workmanship "slip through"?
- Does he provide adequate job instructions for employes unfamiliar with their jobs? Does he follow up to make sure the job is being done correctly?
- Is he careful about wasting the materials (tools, supplies, etc.) assigned to him?

RATING

Consider typical or average foremen you have observed in the past performing a job similar to this foreman's. In relation to Productivity, is he (circle one):

NA	9	8	7	6	5	4	3	2	1
Can't Rate	Usually Better		A Little Better		About the Same			A Little Worse	Usually Worse

CA (con't)

What are some specific examples of things this foreman has done in terms of Productivity to justify the rating you gave him?

CB

Code:_____

SAFETY AND HOUSEKEEPING

Maintaining an orderly work area and enforcing adequate clean up; promoting good housekeeping; ensuring that safe work procedures are demonstrated to employes and that safety rules are followed.

In rating your foreman, consider the following:

- Does he regularly check to see that safety equipment is in place and operating or does he need to be reminded by his employes or the Safety Engineer?
- Does he enforce safe working conditions (e.g., wearing glasses, no horse play, etc.) or does he overlook unsafe acts as long as work goals are being met?
- Does he usually keep his work area clear or is he inclined to let things slide frequently? Does he assign employes to specific housekeeping assignments?
- Are all new employes in his area given a safety talk?
- Does he set a good example of safe working practices and good housekeeping?

RATING

Consider typical or average foremen you have observed in the past performing a job similar to this foreman's. In relation to Safety and Housekeeping, is he (circle one):

NA	9	8	7	6	5	4	3	2	1
Can't Rate	Usually Better		A Little Better		About the Same		A Little Worse		Usually Worse

CB (con't)

What are some specific examples of things this foreman has done in terms of Safety and Housekeeping to justify the rating you gave him?

CC

Code:_____

MAINTENANCE

Ensuring that the right equipment and supplies are on hand to do the job; identifying and resolving maintenance problems when appropriate; scheduling necessary equipment repair; maintaining an adequate inventory of supplies and tools.

In rating your foreman, consider the following:

- When he starts a job are the right tools and equipment lined up or does he have to delay the job to get them?
- Does he keep an adequate inventory of stock or parts that he needs?
- Does he notify other areas of maintenance or stock problems that may affect them?
- If he notices early signs that a piece of equipment is developing a problem, would he take action to get it fixed immediately or wait until the problem fully develops? Does he ignore equipment problems if it's not his responsibility?
- Does he follow up to make sure that appropriate repairs are made on equipment?
- Does he plan preventive maintenance to avoid untimely equipment failure?

RATING

Consider typical or average foremen you have observed in the past performing a job similar to this foreman's. In relation to Maintenance, is he (circle one):

NA	9	8	7	6	5	4	3	2	1
Can't Rate	Usually Better		A Little Better		About the Same		A Little Worse		Usually Worse

CC (con't)

What are some specific examples of things this foreman has done in terms of Maintenance to justify the rating you gave him?

CD

Code:_____

MOTIVATING THE WORKFORCE

Complimenting good workers; being courteous and explaining the reasons for actions concerning employes; setting a good example; enforcing the Company rules and the union contract and disciplining employes fairly when appropriate; showing concern for employes.

In rating your foreman, consider the following:

- Does he recognize his employes when they do a good job or does he assume that is what they are supposed to do and say nothing?
- When telling an employe to do something or when there is a change in procedure, does he explain why or just expect his orders to be followed without question?
- Does he treat each employe with respect and consideration or does he ignore their individual problems and concerns?
- Does he enforce Company rules uniformly and fairly or is he too lenient or unnecessarily strict?
- When an employe breaks a rule does he discipline him immediately and fairly, or does he try not to have to confront and discipline the employe if possible?
- When disciplining an employe, does he usually do it without causing an argument or labor relations problem?
- When disciplining an employe, does he try to get all of the facts or does he go off "half cocked"?
- Does he see if a disciplinary problem can be avoided in the future or does he just want to "punish" the employe?

RATING

Consider typical or average foremen you have observed in the past performing a job similar to this foreman's. In relation to Motivating the Workforce, is he (circle one):

NA	9	8.	7	6	5	4	3	2	1
Can't Rate	Usually Better		A Little Better		About the Same		A Little Worse		Usually Worse

CD (con't)

What are some specific examples of things this foreman has done in terms of Motivating the Workforce to justify the rating you gave him?

CE

Code:_____

MANAGEMENT ATTITUDE

Supporting the Company, particularly when talking to employes; refusing to pass the buck and setting a good example for employes; performing administrative responsibilities, i.e., labor counts; communicating relevant information to those who need to know; cooperating with other departments and areas.

In rating your foreman, consider the following:

- When employes do not agree with a Company policy or procedure, does he talk positively about the Company or does he try to side with the employes?
- If he takes an unpopular action, does he accept responsibility for it or try to blame it on the Company or someone else?
- Does he set a good working example for his employes?
- Does he occasionally put additional effort into his job by coming in early or accepting extra responsibility? Does he work by himself or require frequent supervision?
- Does he complete his administrative responsibilities (time cards, records, overtime equalization charts) on time and accurately or does he let it slide, make errors or have to be constantly reminded?
- Does he cooperate with other departments by following their advise or instructions or does he occasionally ignore them?
- Does he try to help other foremen or people in other areas or does he tend to be out for himself?
- When a problem or change occurs in his area that may affect another area or shift, does he notify the appropriate people to avoid delay and confusion?
- Does he let his supervisor know when he has a problem or when something important occurs that the supervisor should know about?

RATING

Consider typical or average foremen you have observed in the past performing a job similar to this foreman's. In relation to Management Attitude, is he (circle one):

NA	9	8	7	6	5	4	3	2	1
Can't Rate	Usually Better		A Little Better		About the Same		A Little Worse		Usually Worse

CE (con't)

What are some specific examples of things this foreman has done in terms of Management Attitude to justify the rating you gave him?

PART II

This part consists of twelve pages representing six skills a foreman should have to perform his job:

1. Long-Range Planning Skill
2. Short-Range Planning Skill
3. Problem Solving Skill
4. Interpersonal Relations Skill
5. Goal Setting Skill
6. Personal Work Standards

You are asked to rate the extent to which you believe the foreman being rated has these skills, not necessarily how well he used them. Consider the questions listed below the definition of the skill in arriving at the rating. Include at least one or two examples of typical performance that represent the foreman's skill in each area.

At the bottom of the second page of each skill is a second scale for you to indicate the relative amount of that skill a foreman must have to be successful.

DA

Code:_____

LONG-RANGE PLANNING

> Planning and organizing resources in the context of a relatively structured situation (e.g., vacation schedules, scheduled process changes, etc.) where one day to one week is available for developing solutions or plans.

In rating your foreman, consider the following:

- Is he able to plan in advance for anticipated changes in procedures or schedules?
- When he takes an action, is he able to anticipate what the consequences of that action are likely to be a day or two later on?
- Is he able to plan follow up for a decision or action he has made to make sure things are going well?
- When planning a future action, is he able to develop alternatives in case his original plan fails?
- In making a long-range decision or planning a long-range action, is he able to identify what information he needs to make that decision or take that action? Is he usually able to get the information he needs?
- Once he has made his mind up about a future action, is he able to reevaluate and modify his plans if he receives new or contradictory information?
- Is he able to identify situations which <u>may</u> become a problem in the future?

RATING

Consider the Long-Range Planning skill of typical or average foremen you have observed in the past. In relation to these foremen, what level of Long-Range Planning skill does this foreman have (circle one)?

NA	9	8	7	6	5	4	3	2	1
Can't Rate	More than Most		A Little More than Most		About the Same as Most		A Little Less than Most		Less than Most

DA (con't)

What are some examples of specific things this foreman has done that reflect his Long-Range Planning skills?

MINIMUM SKILL REQUIREMENT

What is the minimum amount of Long-Range Planning skill a foreman must have to perform this foreman's job successfully (circle one)?

10	9	8	7	6	5	4	3	2	1	0

A Great Deal Is Required ← → Very Little Is Required

No Skill Required

DB

Code:_____

SHORT-RANGE PLANNING

Planning and scheduling resources to anticipate problems and meet job requirements where less than one day is available; anticipating likely problems, lining up the day's work; early identification and action on problem indicators.

In rating your foreman, consider the following:

- Is he able to identify and effectively plan for the resources (men, materials, other departments, etc.) he will need to meet his daily work goals?
- Is he able to set priorities for the things he must do during the day? Does he use a systematic approach for setting priorities?
- Is he able to identify potential problems before they occur? When early signs of a problem occur, is he able to identify them and act on them before the problem becomes full blown?
- When he takes an action, is he able to predict the consequences of that action later on in the day?
- Is he able to identify most of the information he will need to make a decision? Is he usually able to get that information?
- Is he able to evaluate and make running change in his daily operational decisions as new or contradicting information is received?
- In planning a solution to a problem or planning a change of some type, is he able to plan an effective follow up or decision evaluation strategy?

RATING

Consider the Short-Range Planning skill of typical or average foremen you have observed in the past. In relation to these foremen, what level of Short-Range Planning skill does this foreman have (circle one)?

NA	9	8	7	6	5	4	3	2	1
Can't Rate	More than Most		A Little More than Most		About the Same As Most		A Little Less than Most		Less than Most

DB (con't)

What are some examples of specific things this foreman has done that reflect his Short-Range Planning skills?

MINIMUM SKILL REQUIREMENT

What is the minimum amount of Short-Range Planning skill a foreman must have to perform this foreman's job successfully (circle one)?

10	9	8	7	6	5	4	3	2	1	0

A Great Deal Is Required ← → Very Little Is Required — No Skill Required

DC

Code:_____

PROBLEM SOLVING

Responding to problems demanding immediate attention; responding to crises; decisiveness and stress tolerance.

In rating your foreman, consider the following:

- Is he able to diagnose problems? Is he able to disgnose problems under time pressure?
- When faced with a crisis (injury, serious breakdown, labor problems, etc.) is he able to handle it cooly and calmly?
- If two or more problems occur at about the same time, is he able to handle them effectively?
- Is he able to implement a "fast fix" for a problem when necessary?
- Is he willing to make a decision or give a judgment on an issue?
- Given a very short time period in which to make a decision, is he able to identify the right kind of information he needs to make that decision?

RATING

Consider the Problem Solving skill of typical or average foremen you have observed in the past. In relation to these foremen, what level of Problem Solving skill does this foreman have (circle one)?

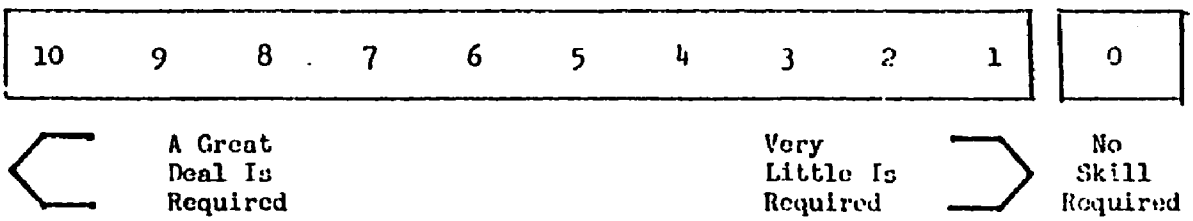
NA	9	8	7	6	5	4	3	2	1
Can't Rate	More than Most		A Little More than Most		About the Same As Most		A Little Less than Most		Less than Most

DC (con't)

What are some examples of specific things this foreman has done that reflect his Problem Solving skills?

MINIMUM SKILL REQUIREMENT

What is the minimum amount of Problem Solving skill a foreman must have to perform this foreman's job successfully (circle one)?



DD

Code:_____

INTERPERSONAL RELATIONS

Sensitivity to interpersonal relationships, pressures and conflicts; treating employes with respect and concern (not implying laxity); delegating; obtaining cooperation from employes; cooperating with others; persuasiveness; resolving conflict; effective use of discipline.

In rating your foreman, consider the following:

- Is he able to get along effectively with his employes? Is he able to handle conflict situations without creating additional problems?
- Is he sensitive to the needs and problems of individual employes? Is he able to treat employes with respect?
- Is he able to help employes develop and improve their performance?
- Is he able to delegate tasks to employes in such a way as to encourage them to accept the additional responsibility?
- Is he able to cooperate and deal with others, particularly fellow foremen?
- Is he able to persuade employes to cooperate in meeting work goals without creating resistance?

RATING

Consider the Interpersonal Relations skill of typical or average foremen you have observed in the past. In relation to these foremen, what level of Interpersonal Relations skill does this foreman have (circle one)?

NA	9	8	7	6	5	4	3	2	1
Can't Rate	More than Most		A Little More than Most		About the Same As Most		A Little Less than Most		Less than Most

DD (con't)

What are some examples of specific things this foreman has done that reflect his Interpersonal Relations skills?

MINIMUM SKILL REQUIREMENT

What is the minimum amount of Interpersonal Relations skill a foreman must have to perform this foreman's job successfully (circle one)?

10	9	8	7	6	5	4	3	2	1	0

A Great Deal Is Required ← → Very Little Is Required

No Skill Required

DE

Code:_____

GOAL SETTING

Emphasizing goal attainment to employes (e.g., production, quality, safety); applying pressure for productive output; define roles and relationships; setting expectations and explaining work procedures; giving job instructions; enforcing rules; disciplining employes.

In rating your foreman, consider the following:

- Is he able to motivate his employes by emphasizing the attainment of specific work goals such as safety, quality, etc.?
- Is he able to "needle" employes into greater work effort?
- Is he able to establish specific and clear-cut work goals and working relationships for his employes?
- Is he willing to give his employes specific work instructions when they require them?
- When a new employe is brought into his area or a procedure is changed, is he able to provide adequate job instructions and training?
- When it becomes necessary to enforce rules, is he able to do so without causing resentment?
- Is he able to use the "right amount" of discipline without creating additional personnel problems?

RATING

Consider the Goal Setting skill of typical or average foremen you have observed in the past. In relation to these foremen, what level of Goal Setting skill does this foreman have (circle one)?

NA	9	8	7	6	5	4	3	2	1
Can't Rate	More than Most		A Little More than Most		About the Same As Most		A Little Less than Most		Less than Most

DE

What are some examples of specific things this foreman has done that reflect his Goal Setting skills?

MINIMUM SKILL REQUIREMENT

What is the minimum amount of Goal Setting skill a foreman must have to perform this foreman's job successfully (circle one)?

10	9	8	7	6	5	4	3	2	1	0

A Great Deal Is Required ← Very Little Is Required → No Skill Required

DF

Code:_____

PERSONAL WORK STANDARDS

Being thorough and attending to detail; following through; staying with a problem or decision until the matter is settled; expending efforting beyond normal expectations; showing initiative; having personal integrity and high work standards; follows instructions.

In rating your foreman, consider the following:

- How well is he able to complete the things that are assigned to him?
- Does he pay fairly close attention to the details of the work he is doing?
- Does he think to follow through on decisions he has made or actions he has taken?
- Does he know when he needs more information to make a decision? Is he willing to seek additional information when required?
- Can he be expected to give above average effort in doing the things assigned to him?
- Does he set high work standards for himself?
- Is he committed to goals and the best interest of the plant and the organization?
- Is he able to take feedback on his job performance and use it to improve the quality of the work he does?

RATING

Consider the Personal Work Standards skill of typical or average foremen you have observed in the past. In relation to these foremen, what level of Personal Work Standards skill does this foreman have (circle one)?

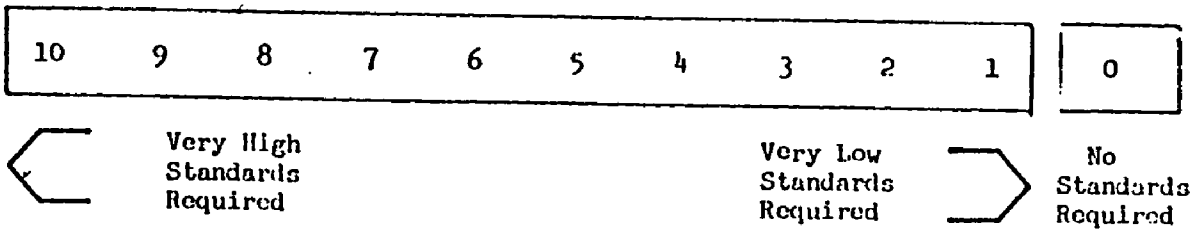
NA	9	8	7	6	5	4	3	2	1
Can't Rate	More than Most		A Little More than Most		About the Same As Most		A Little Less than Most		Less than Most

DF

What are some examples of specific things this foreman has done that reflect his Personal Work Standards?

MINIMUM SKILL REQUIREMENT

What is the level of Personal Work Standards a foreman must have to perform this foreman's job successfully (circle one)?



APPENDIX E

OVERALL PERFORMANCE SUBJECTIVE RATING

CONFIDENTIAL: FOR RESEARCH PURPOSES ONLY

FOREMAN NAME: _____

RATER NAME: _____

PLANT: _____

CODE (OFFICE USE ONLY): _____

OVERALL PERFORMANCE

The overall performance of the foreman in the major factors that make up his job.

In rating your foreman, consider the following:

- Making production
- Housekeeping
- Planning and Scheduling of Machines, Tools, and Supplies
- Using and Allocating Manpower Wisely
- Controlling and Maintaining Quality
- Motivating the Workforce
- Cooperating and Dealing with Others in the Plant
- Maintaining Safe and Healthy Working Conditions
- Giving Job Instructions
- Attending to Administrative Detail and Procedures
- Knowing the Technical Aspects of the Job
- Communicating Relevant Information to the Appropriate Person
- Disciplining the Workforce
- Reducing Material Costs
- Carrying Out Maintenance Responsibilities
- Displaying a "Management Attitude"

RATING

Consider typical or average foremen you have observed in the past performing a job similar to this foreman's. In relation to Overall Performance, is he (circle one);

9	8 7	6 5 4	3 2	1
USUALLY BETTER	A LITTLE BETTER	ABOUT THE SAME	A LITTLE WORSE	USUALLY WORSE

BIBLIOGRAPHY

BIBLIOGRAPHY

BOOKS

Anderson, Scarvia B. and Associates. Encyclopedia of Educational Evaluation. Washington: Jossey-Bass Publishers, 1975.

Barrett, Richard S. Performance Rating. Chicago: Science Research Associates, 1966.

Blum, Milton L. and James C. Naylor. Industrial Psychology: Its Theoretical and Social Foundations. New York: Harper and Row, Publishers, 1956.

Campbell, John P., et al. Managerial Behavior, Performance, and Effectiveness. New York: McGraw-Hill Book Company, 1970.

Chruden, Herbert J. and Arthur W. Sherman, Jr. Personnel Management. Chicago: South-Western Publishing Company, 1963.

Chruden, Herbert J. and Arthur W. Sherman, Jr. Readings in Personnel Management: Chicago: South-Western Publishing Company, 1966.

Costello, Timothy W. and Sheldon S. Zalkind. Psychology in Administration. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

Cummings, L. L. and D. P. Schwab. Performance in Organizations: Determinants and Appraisal. Glenview, Illinois: Scott Foresman and Company, 1973.

Donnelly, James H., Jr., et al. Fundamentals of Management. Austin, Texas: Business Publications, Inc., 1971.

Drucker, Peter F. The Practice of Management. New York: Harper and Row, Publishers, 1954.

Dunnette, Marvin D. Personnel Selection and Placement. Belmont, California: Wadsworth Publishing Co., 1966.

Eddy, William B., et al (Editors). Behavioral Science and the Manager's Role. National Institute for Applied Behavioral Science, 1969.

Flippo, Edwin B. Principles of Personnel Management. New York: McGraw-Hill Book Company, 1966.

Guilford, J. P. Personality. New York: McGraw-Hill Book Company, 1959.

Isaac, Stephan and William B. Michael. Handbook in Research and Evaluation. San Diego: Robert R. Knapp, Publisher, 1974.

Kazmier, Leonard J. Principles of Management. New York: McGraw-Hill Book Company, 1969.

Kerlinger, Fred N. Foundations of Behavioral Research. Chicago, Illinois: Holt, Rinehart and Winston, Inc., 1973.

Lake, Dale G., et al. Measuring Human Behavior. New York: Teachers College Press, 1973.

McGregor, Douglas M. The Human Side of Enterprise. New York: McGraw-Hill Book Company, 1960.

Miller, Richard I. Evaluating Faculty Performance. Washington: Jossey-Bass, Inc., Publishers, 1972.

Monsen, R. Joseph, Jr. and Borje O. Sexberg. The Business World: Introduction to Business Readings. New York: Houghton Mifflin Company, 1967.

Moore, Franklin G. (Editor). A Management Source Book. New York: Harper and Row, Publishers, 1964.

Newman, William H., et al. The Process of Management. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1967.

Roberts, H. S. Roberts' Dictionary of Industrial Relations. Washington: BNA Books, 1971.

Rowland, Virgil K. Evaluating and Improving Managerial Performance. New York: McGraw-Hill Book Company, 1970.

Schein, Edgar H. Organizational Psychology. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1965.

Scott, Walter D., et al. Personnel Management. New York: McGraw-Hill Book Company, 1954.

Scott, William G. and Terence R. Mitchell. Organization Theory: A Structural and Behavioral Analysis. Homewood, Illinois: Richard D. Irwin, Inc., 1972.

Thorndike, R. L. and E. Hagen. Measurement and Evaluation in Psychology and Education. New York: Jolin Wiley and Sons, 1955.

Whisler, Thomas L. and Shirley F. Harper (Editors). Performance Appraisal. New York: Holt, Rinehart and Winston, 1962.

Yoder, Dale, et al. Handbook of Personnel Management and Labor Relations. New York: McGraw-Hill Book Company, 1958.

Yoder, Dale and Herbert G. Heneman, Jr. (Editors). Staffing Policies and Strategies. Washington: The Bureau of National Affairs, Inc., 1974.

Yukl, Gary A. and Kenneth N. Wexley (Editors). Readings in Organizational and Industrial Psychology. New York: Oxford University Press, 1971.

Zander, Alvin. Performance Appraisals: Effects on Employees and Their Performance. Ann Arbor, Michigan: Foundation for Research on Human Behavior, 1963.

PERIODICALS

Austin, A. W. and C. B. T. Lee. "Current Practices in the Evaluation and Training of College Teachers." The Educational Record, 47, 1966.

Blake, Robert R. and Jane S. Mouton, "Power, People, and Performance Review." Advanced Management, 26, 1971.

Byham, W. C. "Assessment Centers for Spotting Future Managers." Harvard Business Review (July-August), 48, No. 4, 1970.

Byham, W. C. "The Assessment Center As An Aid in Management Development." Training and Development Journal, 25, No. 12, 1971.

Calhoon, A. R. and M. J. Epstein. "Performance Appraisal in Management by Objectives." Studies in Personnel Psychology, 4, No. 2, 1972.

Flanagan, John C. "The Critical Incident Technique." Psychological Bulletin, 51, 1954.

Flanagan, John C. and Robert K. Burns. "The Employee Performance Record: A New Appraisal and Development Tool." Harvard Business Review (May-June), 33, No. 3, 1955.

Highland, Richard W. and James R. Berkshire. "Forced-Choice Performance Rating: A Methodological Study." Personnel Psychology (Autumn), 6, No. 3, 1953.

Huck, J. R. "Assessment Centers: A Review of the External and Internal Validities." Personnel Psychology, 26, No. 2, 1973.

Kay, E. and E. F. Juse. "Try This Approach to Performance Appraisals." Supervisory Management. American Management Association (November), 1966.

Lawler, Edward E., III. "The Multitrait-Multirater Approach to Measuring Managerial Job Performance." Journal of Applied Psychology, 51, 1967.

Likert, Rensis. "Motivational Approach to Management Development." Harvard Business Review (July-August), 37, No. 4, 1959.

Mayfield, Harold. "In Defense of Performance Appraisal." Harvard Business Review (March-April), 38, No. 2, 1960.

McConkey, D. D. "Measuring Managers by Results." Personnel Journal, 41, 1962.

McGregor, Douglas M. "An Uneasy Look at Performance Appraisal. Harvard Business Review (September-October), 50, No. 5, 1972.

Meyer, Herbert H., et al. "Split Roles in Performance Appraisal." Harvard Business Review (January-February), 43, No. 1, 1965.

Oberg, Winston. "Make Performance Appraisal Relevant." Harvard Business Review (January-February), 50, No. 1, 1972.

Patton, Arch. "How to Appraise Executive Performance." Harvard Business Review (January-February), 38, No. 1, 1960.

Rieder, G. A. "Performance Review: A Mixed Bag." The Personnel Administrator (November), 1972.

Seashore, Stanley E., et al. "Relationships Among Criteria of Job Performance." Journal of Applied Psychology, 44, 1960.

Sisson, E. Donald. "Forced Choice: The New Army Rating." Personnel Psychology (Autumn), 1, No. 3, 1948.

Slevin, D. P. "The Assessment Center, Breakthrough in Management Appraisal and Development." Personnel Journal (April), 51, No. 4, 1972.

Smith, Patricia C. and L. M. Kendall. "Retranslation of Expectations: An Approach to the Construction of Unambiguous Anchors for Rating Scales." Journal of Applied Psychology, 47, 1963.

Spriegel, W. R. "Company Practices in Appraisal of Managerial Performance." Personnel, 39, 1962.

Taylor, Erwin K. and Robert J. Wherry. "A Study of Leniency in Two Rating Systems." Personnel Psychology, 4, 1951.

Thompson, P. H. and G. W. Dalton. "Performance Appraisal: Managers Beware." Harvard Business Review (January-February), 48, No. 1, 1970.

Turner, W. W. "Dimensions of Foreman Performance: A Factor Analysis of Criterion Measures." Journal of Applied Psychology, 44, 1960.

Whitla, Dean K. and John E. Tirrell. "The Validity of Ratings of Several Levels of Supervisors." Personnel Psychology (Winter), 6, No. 4, 1954.

Zeitlein, L. R. "Planning for a Successful Performance Review Program." Personnel Journal, 48, No. 12, 1969.

OTHER

Baylie, Thomas N. et al. "Appraisals of People Resources." In Dale Yoder and Herbert G. Heneman, Jr. (Editors). Staffing Policies and Strategies. Washington: BNA Books, 1974.

Bray, D. W. and D. L. Grant. "The Assessment Center in the Measurement of Potential for Business Management." Psychological Monographs, 80, No. 17, 1966.

Gage, N. L. "Ends and Means in Appraising College Teaching." In McKeachie, W. J. (Editor). The Appraisal of Teaching in Large Universities. Ann Arbor, Michigan. The University of Michigan, 1959.

Wickstrom, W. S. "Developing Managerial Competence: Changing Concepts - Emerging Practices." National Industrial Conference Board Studies in Personnel Policy, No. 189, 1964.

Wherry, Robert J. "The Control of Bias in Ratings, VII, A Theory of Rating." PRB Report No. 922. Department of the Army, Personnel Research Branch, 29 February, 1952.