

DEVELOPMENT OF AN
INTEGRATED PERFORMANCE APPRAISAL SYSTEM
FOR TRUCK DRIVERS IN THE WINE TRADE

PETER EDWIN ARTHUR WHITEHEAD, B.A. (HONS.) IN PSYCHOLOGY

*Thesis submitted to the Department of Psychology,
University of Cape Town,
in fulfilment of the requirement for the
degree of Master of Social Science
in Industrial Psychology*

SUPERVISOR : DR. A.M. KELLERMAN

CAPE TOWN
SOUTH AFRICA

MARCH 1986

The University of Cape Town has been given
the right to reproduce this thesis in whole
or in part. Copyright is held by the author.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

DEVELOPMENT OF AN
INTEGRATED PERFORMANCE APPRAISAL SYSTEM
FOR TRUCK DRIVERS IN THE WINE TRADE

ACKNOWLEDGEMENTS

The successful completion of this study was greatly dependent upon the carefully co-ordinated efforts of many individuals who brought with them diverse interests and expertise. I wish to acknowledge gratefully the significant contributions by these individuals and institutions.

- Dr. A.M. Kellerman, my supervisor, who so patiently guided me through the different phases of the study, and from whom I not only acquired a great deal of knowledge about the subject, but also a great deal of knowledge as a person.
- The Human Sciences Research Council for financial assistance.
- The wine manufacturing organization for assisting me in executing this study.
- Frank Bokhorst, the computer expert in the Psychology Department.
- Jo Torrance for the excellent typing of this manuscript.
- Carol Bower, Dirk le Roux and Francois Smit for special assistance.
- My parents, without whose education, faith and encouragement this study would not have been possible.
- Stephnie, for always assisting and encouraging me when things looked rather bleak.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	vii
LIST OF FIGURES	viii
ABSTRACT	ix
CHAPTER 1 INTRODUCTION	1
The Origin and Growth of Performance Appraisal	2
Performance Appraisal : Defining the Concept	4
The Role of Performance Appraisal in an Organization	5
Problem	7
Objectives	10
CHAPTER 2 GENERAL OVERVIEW OF PERFORMANCE APPRAISAL	13
The Performance Appraisal Process	13
The Job Description	15
Development of the Measuring Instrument	16
The Selection and Training of Raters	17
Collecting Job Performance Information	17
Measurement of Performance	18
The Appraisal Interview	19
Purposes of Performance Appraisal	20
Motivating Employees for Improved Performance	25
Determination of Training Needs	28
Promotion and Compensation	28
Counselling Employees	29
Manpower Inventory	29
Management Development	30
Communication	30
Research	30

Performance Evaluators	31
Appraisal by Subordinates	33
Peer Ratings	34
Combination of Raters	36
Client Appraisal	37
Self-Appraisal	38
Appraisal by Supervisors	39
Potential Performance Appraisal Problems and Proposals for Solving Them	41
Maintaining of Performance Appraisal Systems	41
Rater Problems with Performance Appraisal	43
Halo error	45
Recency of events	47
Central tendency	48
Leniency or harshness error	49
Personal bias	52
The one asset individual	53
Contrast	53
Projection	54
Effect of past record	54
Proposals for Solving Rater Problems	55
Criteria for Successful Performance Appraisal Systems	60
Job Relatedness	60
Definition of Success	61
Reliability and Validity	62
Standardization	63
Practicability	63
Formal Documentation	64
Acceptability to Trade Unions	65
Employee Participation	66
Summary	69
CHAPTER 3 DIFFERENT APPROACHES TO AND TECHNIQUES OF PERFORMANCE APPRAISAL	72
Approaches to Performance Appraisal	72
Trait-oriented Approach	73
Results-oriented Approach	74
Behaviour-based Approach	76
Different Performance Appraisal Techniques	78
Comparative Techniques	80
Ranking	80
Paired comparisons	81
Forced distribution	83
Evaluation of Comparative Techniques	86
Absolute Appraisal Techniques	87
Essay evaluation/narrative technique	88
Weighted checklist	89

Critical incidents	90
Graphic rating scale	93
Behaviourally anchored rating scale (BARS)	96
The behavioural observation scale (BOS)	101
Outcome Oriented Techniques	110
Direct index approach	110
Standards of performance	112
Management by objectives (MBO)	112
Motivation for Selecting the Behavioural Observation Scale as Performance Appraisal Technique	115
CHAPTER 4 METHODOLOGY	119
Definition of the Problem	119
Procedure	120
The Sample	122
Truck Drivers	124
Transport Foremen	125
Preliminary Study : Determining the Need for a New Performance Appraisal System	126
Construction of the Behavioural Observation Scale	130
Development of the Performance Appraisal Instrument	131
Job description	131
Description of Behavioural Dimensions	140
Responsibility	141
Safety	141
Supervision	141
Liaison with customers	143
Maintenance of truck	143
Perform administration	143
Reliability	144
Validity	145
Performance Appraisal at Three-monthly Intervals	145
Training of the Raters	146
The Actual Appraisal and Feedback Interviews	149
Statistical Analysis	154
The Hypotheses	156
Hypothesis 1	156
Hypothesis 2	158
Hypothesis 3	159
Hypothesis 4	160
Hypothesis 5	160
CHAPTER 5 RESULTS AND DISCUSSION	162
Preliminary Study : The Need for a Performance Appraisal System	163

Results of Performance Appraisal	170
Means and Standard Deviations	171
Reliability	171
Validity	179
Results of the Implementation of the BOS	180
Planned Comparisons	181
Hypothesis 1	182
Discussion	186
Hypothesis 2	194
Discussion	194
Hypothesis 3	196
Discussion	196
Hypothesis 4	200
Discussion	200
Hypothesis 5	202
Discussion	202
Comparison of five hypotheses	205
Unobtrusive Measures	207
Acceptability of the BOS	209
Comparison of experiences with the two systems	210
Experiences versus expectations of New Performance Appraisal System	214
CHAPTER 6 CONCLUSION	218
Recommendations to the Organization	221
REFERENCES	224
APPENDICES	
APPENDIX 1 - Previous Performance Appraisal System	232
APPENDIX 2 - Performance Appraisal Satisfaction Questionnaire	236
APPENDIX 3 - List of Items compiled from the Critical Incidents	240
APPENDIX 4 - Vehicle Check List - Daily	242
APPENDIX 5 - Scale of Importance	244
APPENDIX 6 - An Interaction Management Module	245
APPENDIX 7 - The Behavioural Observation Scale	246
APPENDIX 8 - Inter-Item Correlation Matrix	251

LIST OF TABLES

	PAGE
TABLE 1 Purposes of Performance Appraisal Programmes in Different Companies - First Study	22
TABLE 2 Purposes of Performance Appraisal Programmes in Different Companies - Second Study	23
TABLE 3 Purposes of Performance Appraisal Programmes in Different Companies - Third Study	24
TABLE 4 Illustration of a Weighted Checklist developed for Kitchen Managers	91
TABLE 5 Criteria for good Appraisal Systems included in Different Appraisal Formats	118
TABLE 6 Ratio of Interjudge Agreement for each Dimension	142
TABLE 7 Discrepancy between Experiences and Expectations of Truck Drivers regarding Performance Appraisal	165
TABLE 8 Means and Standard Deviations for Different Performance Appraisals	172
TABLE 9 Reliability of Measuring Instrument	173
TABLE 10 Reliability of Scale if each item is deleted	176
TABLE 11 Results of Hypothesis 1	183
TABLE 12 Percentage of Variance Explained	185
TABLE 13 Results of Hypothesis 2	195
TABLE 14 Results of Hypothesis 3	197
TABLE 15 Results of Hypothesis 4	201
TABLE 16 Results of Hypothesis 5	203
TABLE 17 Discrepancy between the Experiences of the Drivers with the First and Second Performance Appraisals	212
TABLE 18 Results obtained from Second Questionnaire determining the Experiences and Expectations of the Truck Drivers regarding the new Performance Appraisal System	215

LIST OF FIGURES

	PAGE
FIGURE 1 The Performance Appraisal Process	14
FIGURE 2 Potential Performance Appraisal Sources	32
FIGURE 3 Central Tendency Error	50
FIGURE 4 Distribution of Lenient and Harsh Ratings	51
FIGURE 5 Criteria for good Performance Appraisal Systems	70
FIGURE 6 Classification of different Performance Appraisal Techniques	79
FIGURE 7 The Paired Comparisons Appraisal Technique	82
FIGURE 8 Distribution of 40 Employees by Forced Distribution	85
FIGURE 9 The forced-choice Appraisal Technique	85
FIGURE 10 Different types of Graphic Rating Scales	95
FIGURE 11 One Dimension on a Behaviourally Anchored Rating Scale	99
FIGURE 12 Example of one BOS Performance Dimension for Evaluating Managers	102
FIGURE 13 The Difference between Critical Incidents, Behavioural Items and Behavioural Dimensions	104
FIGURE 14 The Transport Department	123
FIGURE 15 Performance Appraisal at Three-monthly Intervals	151
FIGURE 16 Hypotheses Graphically Illustrated	157

ABSTRACT

The purpose of this study was to develop a performance appraisal system for truck drivers in a wine manufacturing organization. One of the objectives of the system developed, was to improve the performance of the truck drivers. The sample involved in this study consisted of 80 truck drivers and six transport foremen.

A preliminary study was done to determine the satisfaction of the drivers with the previous appraisal system. A need for a new performance appraisal system was established and it was therefore decided to develop a new performance appraisal system for the truck drivers in the Transport Department of the organization.

A literature review indicated that the most suitable appraisal system for this specific situation was the behavioural observation scale (BOS). A job analysis of the truck driver's job was done, critical incidents were collected and behavioural dimensions were defined. This resulted in the final behavioural observation scale consisting of 37 behavioural items. The reliability of the appraisal instrument determined by Cronbach's coefficient alpha, was .98. An effort was made to achieve both content and face validity for the BOS.

To determine whether the performance of the truck drivers did increase as a result of the new appraisal system, an experimental and control group were defined. Their performance was appraised three times at three-monthly intervals with the BOS. The experimental group received feedback on its performance, which included setting goals to be achieved by the next appraisal. The drivers in the control group were unaware of the fact that their performance was being appraised. To determine whether the performance of the drivers in the experimental group had improved, planned comparisons were done. There was a substantial improvement in performance amongst the experimental group, whilst the control group's performance remained unchanged.

It was concluded that the intervention was successful. The use of a behavioural approach to performance linked with adequate feedback made a major contribution to the efficiency of these truck drivers as well as to their interaction with their supervisors, the transport foremen.

CHAPTER 1

Introduction

This study was undertaken at a wine manufacturing organization in the Western Cape, with the objective of developing a performance appraisal system for truck drivers in the organization. This thesis will deal mainly with various aspects underlying the development of a performance appraisal system.

In the process of developing such a system it was necessary to review the literature with regard to the performance appraisal process, and to get an overview of the different approaches to and the techniques of performance appraisal. A reasonable match between the needs of the organization, the requirements of the truck drivers involved and a reliable and valid appraisal system had to be achieved.

The first chapter defines performance appraisal and reviews the importance of performance appraisal systems to any organization. The problem of the organization involved in the study with regard to performance appraisal for truck drivers is sketched and the objectives of the present study are delineated.

The Origin and Growth of Performance Appraisal

The requirement that one person, usually a superior, evaluates another person, usually a subordinate, in terms of organizational criteria is as old as human society. As early as 221 AD, a Chinese philosopher in the Wei Dynasty, named Sin Yu, noted that individuals were not rated according to their merits, but according to the rater's likes and dislikes. In 1684 the Dublin Evening Post published a list of personal qualities which was used to evaluate each member of the Irish Legislature. In 1813 the United States army also used performance appraisal in some of its branches. The first industrial application of performance appraisal was made in the early 1800's by Robert Owen at his cotton mills in New Lanark, Scotland, whereas the first formal appraisal programme in the United States' Federal Civil Service was introduced in 1842 (Slivinski, 1975).

During World War II, the most important aspect regarding performance appraisal was the development of the forced-choice technique and a critical incidents approach to performance appraisal. Since the Second World War, the application, use and the development of different techniques and approaches to performance appraisal have mushroomed (Eichel & Bender, 1984; Slivinski, 1975).

In 1947 a bibliography listed over 600 books and pamphlets on performance appraisal and in 1953 a survey by the state of

California indicated that 95 per cent of the employees involved in this survey were interested in knowing how well they were doing in their jobs. In 1953 the Michigan Civil Service Commission conducted a study, the findings of which clearly indicated that those employees whose work performance was discussed with them frequently were more satisfied with this current practice than were those whose work performance was discussed with them infrequently or not at all (Slivinski, 1975). Slivinski further stated that this evidence seems to suggest that employees welcome performance evaluation as a major way of learning how they are doing on their jobs. Without a formal appraisal programme, supervisors would tend to speak to their employees only when they did something wrong.

By the early 1960's Wikstrom (1964) estimated that 84 per cent of all companies used some type of performance appraisal system. Some ten years later Slivinski (1975) estimated that 90 per cent of all companies and corporations used some type of system to appraise an individual's performance within an organizational setting.

It is obvious from the abovementioned that performance appraisal has in the past played a major role and is playing an increasingly important role in organizations. This highlights the necessity of a more in-depth exploration of the concept of performance appraisal.

Performance Appraisal : Defining the Concept

A vast number of definitions for this concept are available (Eichel & Bender, 1984; Henderson, 1984; Ivancevich & Glueck, 1983; Latham & Wexley, 1981). The definition given by Schuler will suffice: "Performance appraisal is a formal, structured system of measuring and evaluating an employee's job-related behaviours and outcomes, to discover how and why the employee is presently performing on the job and how the employee can perform more effectively in the future so that the employee, the organization and society all benefit." (1981, p. 221).

The main assumptions underlying this definition are the following:

- Behaviour must be observed by someone, usually a superior, in a formal, structured manner;
- An assessment of employee performance against predetermined job standards must take place;
- The information gathered in the appraisal, must be communicated to the employee.

Cascio (1982) supported the abovementioned when he stated that performance appraisal is composed of mainly two processes, namely observation and judgement.

Employees continually have their performance appraised on the job, whether formally or informally. Appraisals may be made from haphazard observation, memory, hearsay or intuition. A formal and rational system is of course more accurate, fair and useful to all concerned. Latham and Wexley (1981) defined a formal appraisal system as the systematic review of an employee's performance on the job which is used to evaluate the effectiveness of his or her work. Based on these definitions, it seems appropriate to review the role of performance appraisal in organizations.

The role of Performance Appraisal in an Organization

Nowadays, most organizations follow an active policy of continuously reviewing its past performance, its present progress and the future prospects of its human resources. In cases where this is not done, it may seem as if the organization is managing itself by intuition. The engagement of human talents in achieving the goals and objectives of the organization, the acceptance of the responsibilities implicit in the very function of management, and the necessity of a systematic manpower plan constitutes three reasons why sound management and modern organizations need a formal employee appraisal system (Slivinski, 1975). Therefore it is difficult to envisage how the abovementioned needs can be met effectively, without a specific sub-system of employee

performance appraisal which has the threefold purpose of development, providing information and motivation.

Such an appraisal system usually involves a process through which individual performance can be observed and evaluated so as to determine the quality and value thereof to the organization. This information is then transmitted to higher levels for information and decision-making affecting the employee's status and rewards. It is also used to develop a plan of action for an employee to increase his/her contribution to the organization and this heightens his/her sense of purpose and accomplishment. In other words, a performance appraisal system links together in one integrated network the goals of the organization, its manpower requirements and the goals of the individual.

Thus, a performance appraisal system makes its contribution to the organization's manpower planning, where the objectives of manpower planning are the knowledge of its present manpower resources, knowledge of the manpower environment, as well as knowledge of organizational objectives (Slivinski, 1975). The role played by a performance appraisal system in an organization can therefore be summarized as the maximum utilization of the existing human resources where the objectives of the organization and the individual are taken into consideration.

In view of the central role of performance appraisal discussed above, it therefore seemed appropriate to design a performance

appraisal system for the truck drivers in the present study. The discussion below will focus firstly on the problems of the truck drivers regarding performance appraisal as experienced by a wine manufacturing company and secondly, an effort will be made to clarify the objectives of the study.

Problem

In the company involved in this study, the delivery of bottled wine was done by means of trucks which visited the various outlets on a regular basis. In many instances the truck driver and his crew were the only direct contact between the company and its customers, apart from infrequent contacts from sales representatives. Orders for the company's products could be made per telephone or with the driver and the company was only represented directly when the truck driver arrived at the customer to deliver the goods ordered. Future or present customers' decisions to order products from the company may depend entirely on the behaviour and attitude demonstrated by the driver when loading or unloading cargo, or taking orders from them.

Added to that, the driver also spent a great deal of his working day on the open road, driving to and from customers. Therefore, he also represented the organization on the road. Thus, it was essential that the driver behaved in such a way as to promote the company's image at all times. A negative ex-

perience with the general public on the road may influence their choice of wine.

It is therefore vitally important to the organization to have truck drivers who are peak performers. The previous appraisal proved to be inadequate. The system consisted of a behaviourally anchored rating scale which included six trait- or personality-oriented items (to be discussed later). Raters had to rate employees on these items on a 9-point scale, the lower the score, the more negative the evaluation. This appraisal system, however, was used across jobs, resulting in no job-specific information being collected. An example of the existing performance appraisal form appears in Appendix 1. Although still in use, this system which was initiated eight years ago, was viewed as less important during the last two to three years because of the following deficiencies:

- The system was vague, i.e. criteria rated had no job specific relation and were therefore difficult to measure and too subjective.
- No record of employee behaviour throughout the review period was available and superiors could not give an accurate assessment of performance.
- The direct link of performance appraisal to remuneration determination led to a situation where superiors tended

not to give an honest assessment of an employee. They tended to rate higher in an effort to achieve better increases for their staff.

- The criteria included on the performance appraisal system were only applicable to the supervisory level of employees, which meant that they were too advanced for the group involved.
- The system was not proactive, i.e. it did not provide for the planning and setting of expected goals and performance standards, as well as for the improvement of unsuccessful job behaviour.
- The system did not allow supervisors to clearly distinguish between different workers.

To illustrate the image of the existing system, a personnel manager at the company said: "Managers no longer believe in the system. They see it simply as a bit of paperwork they must do annually by means of which they can try to get the best increases possible for their subordinates, irrespective of the real differences in their levels of performance."

To summarize the problem: Prior to this study, no appropriate appraisal system for truck drivers was in operation in the organization. As truck drivers provided the only means of transportation

of the organization's products to its clients, the job of truck driver was a critical aspect of the successful functioning of the organization as a whole. To achieve this goal, the truck drivers needed feedback on their performance. A performance appraisal which was specifically designed for truck drivers might satisfy this need.

Therefore, such a system must enable the truck drivers to get direct feedback on their behaviour. In addition, the system should also be understandable to them.

Following this broad outline of the problem as experienced by the company, some objectives for solving it may be identified.

Objectives

The general objective of this study was to develop a performance appraisal system which would meet the needs of both the truck drivers and the organization. An inherent aspect of this system had to be the improvement of on-the-job performance of these drivers. The system, however, had to be acceptable to them.

The steps taken to achieve this goal were firstly, to determine whether a real need for a new performance appraisal system existed amongst the truck drivers. After establishing this need, a system which would suit the organization, as well as the truck drivers, had to be found. Such a system had to involve regular ratings

to give feedback on a continuous basis. For the purposes of this study it was decided to appraise the truck drivers every three months. The results of the three appraisals were then compared to determine if improvement in the quality of performance and an increase in satisfaction of the drivers had taken place. Over and above the performance ratings, an effort was also made to get feedback on aspects such as accident rates, traffic violations and customer complaints. This also provided quantitative information on the functioning of the drivers.

In conclusion, the truck drivers' satisfaction with the new performance appraisal system was assessed to ensure that the new system would meet their needs.

Even though the author is aware of a trend to use a more general performance appraisal system which would satisfy the demands of the organization as a whole, it was clear in the case of the truck drivers that one should cater for their specific needs. Their role in the organization seemed vital enough to warrant a tailor-made system which would satisfy their particular needs.

To achieve the abovementioned objectives, the literature review focused on the performance appraisal process, the purposes for which performance appraisal is used, problems regarding performance appraisal and different performance appraisal techniques.

The study assessed the existing system and then focused on the development of a new performance appraisal system. The results of the new system are reported in chapter 5, and the conclusion drawn was that the intervention was significant.

CHAPTER 2

General Overview of Performance Appraisal

Some general aspects regarding performance appraisal were discussed in the first chapter. These included the origin of performance appraisal and the position thereof in an organization. In this chapter the actual performance appraisal process will be described in more detail. Initially, an overview of the process will be given to orientate the reader and afterwards the different aspects surrounding performance appraisal will be described. These aspects include the purposes of performance appraisal, who should appraise, potential performance appraisal problems and proposals to solve these problems, as well as criteria for good appraisal systems.

The Performance Appraisal Process

A study undertaken by Eichel and Bender in 1984, indicated that the performance appraisal process can be divided into several different steps. These steps include job descriptions, the development of the measuring instrument, selecting and training observers, collecting and evaluating the job performance information, performance feedback and the setting of new goals. These steps are illustrated in Figure 1. The discussion will follow the same order presented in the figure.

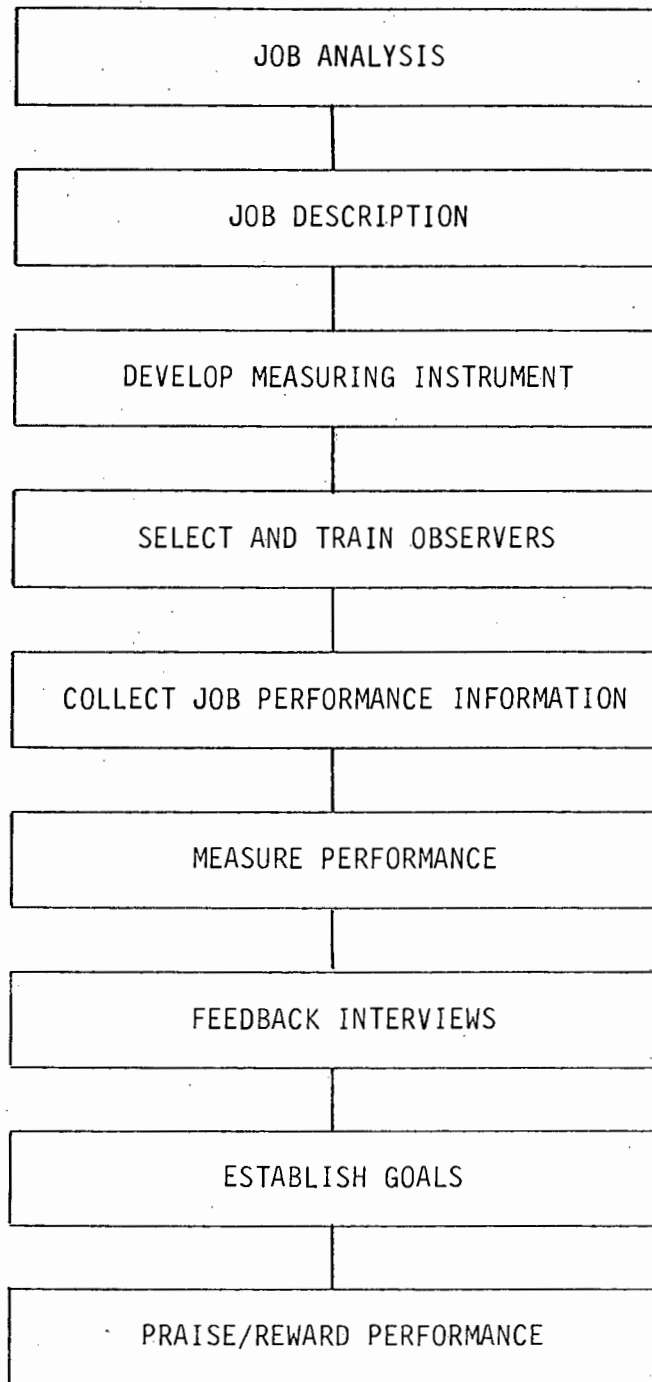


FIGURE 1 ,

The Performance Appraisal Process

Eichel and Bender (1984) included 588 organizations, which were all members of the American Management Association and the results obtained from this study indicated that the abovementioned steps were consistently present when these organizations implemented new performance appraisal systems. The different steps will now be discussed in more detail.

The Job Description

A job description usually results from a job analysis. This job description constitutes a record of existing and pertinent job facts and is a summary of the tasks, duties and responsibilities in a specific job, as determined by a job analysis. Therefore, before a performance-appraisal system could be developed, a job description would have to be compiled to indicate which aspects of a certain job should appear in the performance appraisal. In other words, the job description indicates what is done, why it is done, where it is done and how it is done. This information is very important for evaluating the performance of any employee, because it provides job-related criteria against which the employee can be evaluated (Alewine, 1982; Eichel & Bender, 1984).

A job description could serve a variety of purposes, and therefore it has extensive relationships with several other personnel and human resource management activities, such as selection,

compensation and training and development (Latham & Wexley, 1981). It is particularly important however, when it comes to performance evaluation, because to effectively evaluate employee performance the appraisal method that is used must reflect the duties of the job. As mentioned previously, a job description specifies these duties. According to Halloran (1981) the supervisor and the employee holding a specific position in question are the most suitable people to do the job description. In the case of this study, both parties were involved when writing the job description to ensure that agreement on the work to be done by the truck driver would be reached.

Development of the Measuring Instrument

There are a variety of performance appraisal systems with a diversity of methods for designing these systems. Some of these methods are briefly reviewed in a later section of this thesis. Thus, only a few general comments with regard to the development of an instrument will have to suffice.

According to Richardt (1976) the best appraisals are those which are tailormade for a specific job and where performance is stated in terms of measurable results to be achieved. This improves the objectivity of the appraisal system. This is what was intended by the appraisal system developed in this study. The appraisal was developed for a specific job, namely that of

the truck driver in the Transport Department of a wine manufacturing company. As this is a relatively undemanding job, attempts were also made to include items in the performance appraisal which were not too complex for the respondents to understand, thus achieving a reasonably tailor-made system.

The development of the measuring instrument as implemented in this study, is discussed in chapter 4.

The Selection and Training of Raters

Raters evaluate the actual performance of workers. The most popular method for appraising performance of employees is to use the direct supervisor of those being rated (Graves, 1982). There are, however, other options available which will be outlined in a later section.

Thorough training of raters is essential to ensure as high a degree of reliability as possible in the evaluation of performance. More attention is paid to this topic at a later stage.

Collecting Job Performance Information

Because appraisals are usually not done more than twice a year, information about an employee's work must be collected and filed until the next appraisal period. It is not always

easy for supervisors to remember certain incidents for long periods. As a possible solution, Cascio (1982) pointed out that goals and measurable standards of performance should be agreed upon at the beginning of each appraisal period. If not, it would be difficult for the supervisor to make an objective evaluation of the employee's performance when appraising it. These goals must then be achieved by the end of the appraisal period. Information about the performance of the employee is thus readily available. An added advantage of this system is that subordinates are encouraged to discuss their work with supervisors during this time, because they have mutually agreed on specific goals (Alewine, 1982; Eichel & Bender, 1984).

Measurement of Performance

During this stage, a rater must evaluate the performance of employees. This measurement of performance is done by means of the instrument developed earlier on in the performance appraisal process. The first measurement may pose some problems because achievement goals have not yet been clarified. However, as the process continues, one will expect some improvement in performance on the part of the worker, because clearly defined targets have been set (Latham & Wexley, 1981).

Of course the quality of the measuring instrument is of vital importance. A discussion of the various instruments and problems with regard to the measuring process follows later.

The Appraisal Interview

Mathis and Jackson (1982) stated that it is vitally important to communicate the results of a performance appraisal to an employee. A discussion of the results of the appraisal with employees helps them to achieve a clear understanding of how they are viewed by their supervisor and the organization (Alewine, 1982; Cascio, 1982; Eichel & Bender, 1984; Graves, 1982; Landy, 1985; Latham & Wexley, 1981; Wells, 1982).

Counselling and development should be emphasized rather than simply informing employees of how they were rated and why. By focusing on development, an opportunity is provided for consideration of the employee's performance and its improvement (Schuler, 1981). Wells confirmed this by saying that appraisers need to be trained to provide feedback "without generating animosity, to praise as well as to criticize, to confront employees constructively, to listen effectively and to be sensitive to the situations and conditions which the employee faces" (1982, p. 780).

In the study involving 588 companies mentioned previously, Eichel and Bender (1984) found that 98,8 per cent of the survey respondents held feedback interviews. Of these, 75,7 per cent held them once a year, 17,3 per cent held them twice a year and 8,1 per cent held them more than twice a year. The majority, 71,5 per cent, conducted structured feedback interviews, while another 24,7 per cent described their feedback as informal communication.

Burke, Weitzel and Weir (1978) found that participation of employees in the appraisal interview, as well as job related items such as problem solving and setting specific goals, correlated highly with the employees' satisfaction with the appraisal process.

In this study, the truck drivers played a major role in developing the appraisal system as well as in setting goals to be reached in the future. By doing this it was hoped that a higher degree of satisfaction with the appraisal system could be achieved.

After this very broad overview of the performance appraisal process, a more detailed review of some aspects relevant to the process, as mentioned earlier in this chapter, seems to be in order. Initially it seems appropriate to deal with the purposes of performance appraisal before going on to other more applied areas.

Purposes of Performance Appraisal

The decision to design a performance appraisal system stems from a particular need in the organization. In other words, it is necessary to clarify why performance is to be assessed (the purpose of appraisal) before deciding upon the method to be used for assessing that performance.

Various authors studied the purposes of performance appraisals in different organizations. Firstly, French (1970), did a survey for the National Industrial Conference involving 166 companies. All of these companies used a systematic type of performance appraisal. The results of this survey are reported in Table 1. It was found that appraisals were used mostly for determining which employee should be promoted, and for wage and salary determination. Lay-off of personnel was one of the least used purposes.

Secondly, Schuster and Kindall (in Klatt, Murdick & Schuster, 1978) conducted a similar survey involving 403 companies using performance appraisal systems. The results obtained by them are illustrated in Table 2. It became evident that performance appraisals were used mostly for counselling, and for planning of training and development of employees. It was least used for retention or discharge of employees, or improving company planning.

In a third study, Eichel and Bender (1984) found that in 588 companies, performance appraisals were used mostly for compensating and counselling employees, with limited use for retention or discharge. The results obtained in this study are given in Table 3.

TABLE 1

Purposes of Performance Appraisal Programmes in Different Companies
: First Study

<u>PURPOSES</u>	<u>NUMBER OF COMPANIES</u>	<u>%</u>
Promotion	122	73
Wage or Salary determination	144	69
Training and Development	102	61
To help supervisors know their employees	101	61
To let workers know their progress	102	61
Transfer	98	59
Follow-up interviews	57	34
Discharge	77	46
Personnel research	48	29
Lay-off	44	27
Total companies with performance appraisal programmes	166	100

(Adapted from: French, 1970, p. 292).

TABLE 2.

Purposes of Performance Appraisal Programmes in Different Companies

: Second Study

<u>PURPOSES</u>	<u>NUMBER OF COMPANIES</u>	<u>%</u>
Counselling the ratee	278	88,0
Planning training or development for the employee	270	85,4
Motivating the employee to achieve higher levels of performance	269	85,1
Promoting the employee	266	84,2
Merit increases	238	75,3
Retention or discharge of the employee	184	58,2
Improving company planning	178	56,3
Other purposes of performance appraisal	28	8,9
Total companies with performance appraisal programmes	316	

(Adapted from Klatt, Murdick & Schuster, 1978)

TABLE 3

Purposes of Performance Appraisal Programmes in different Companies
: Third Study

PURPOSE	PERCENTAGE OF THOSE RESPONDING
Compensation	80
Counselling	65
Training and Development	64
Promotion	45
Manpower Planning	43
Retention / Discharge	30

(Adapted from: Eichel & Bender, 1984, p. 12).

In conclusion, it appeared that the three main purposes for which performance appraisals were used are counselling employees, determining their development and training needs, as well as determining their compensation. It was evident that performance appraisals were rarely used for retention or discharge of employees.

The abovementioned purposes of performance appraisal, as well as other purposes for which appraisal systems can be used, will now be discussed.

Motivating Employees for Improved Performance

According to Glueck (1979) the single most important purpose is to improve employee performance. Thus, an analysis of employee performance should identify weaknesses in performance. On the basis of this information, some plans for improving the performance can be made. After implementation of these plans, another review is done which starts the cycle all over again. This guides the employee in continuously setting goals for improvement, which may lead to the enhancement of his or her career (Biesheuvel, 1984).

If an employee is to be motivated through performance appraisal, five basic steps should be involved (Latham & Wexley, 1981). In the first place, a supervisor must determine what is expected from the employee doing a particular job.

Secondly, the direct supervisor must be sure that the people responsible for completing the performance appraisal are able to recognize effective and ineffective performance when they see it.

In the third instance, the supervisor must involve the employee in the setting of specific goals for the employee, and fourthly, the supervisor must take steps to ensure that the consequences of goal attainment are positive. If not, the goals will not be accepted. In terms of the learning theory, positive behaviours which are rewarded, increase the probability that these behaviours will be repeated.

Finally, employees must also be involved in solving problems which concern both the employee and the supervisor.

The key issues in this process are the setting of goals and the possibility of achieving such goals. In a laboratory experiment, Locke (1968) assigned individuals different types of goals on a variety of simple tasks. Individuals who were assigned difficult goals performed better than individuals who were assigned moderate or easy goals. Individuals who were simply trying to do their best were outperformed by individuals who had specific, challenging goals.

Latham and Wexley (1981) stated three reasons why goal setting affects performance. In the first place the setting of goals has a directive effect on what people think and do. Goals focus activity in one specific direction. They also regulate the expenditure of energy, since people typically put forth effort in proportion to the difficulty of the goal, provided that the goal is accepted. Difficult goals also lead to more persistence than easy goals. These three dimensions, namely direction, effort and persistence, are the central aspects of the motivation and appraisal process.

McClelland (1961) found that goal setting may account for improved performance, because it may create a desire in employees to perform in terms of a standard of excellence. He defined this process as the need for achievement. McClelland further stated that since high achievers selected a goal they tend to be totally preoccupied with the task until it is successfully completed. In other words, goals should be set for employees which ought to stimulate their need for achievement, thus resulting in improved performance. These goals should be moderate to be more easily accepted by the achiever.

According to Luthans (1977) high achievers have a desire for immediate feedback on performance and they find accomplishing a task intrinsically satisfying.

As the abovementioned aspect, namely goalsetting, was an integral part of the appraisal system to be developed, it was assumed that performance of employees being appraised would improve when applying the newly developed system.

Determination of Training Needs

If the appropriate facts are included in the performance appraisal, raters could be able to identify the strengths and weaknesses of their subordinates. In doing this, they may be able to assist their training department in establishing objectives for group training programmes and develop these programmes according to the needs of specific employees (Schuler, 1981). Added to that, performance appraisal can also inform employees about their progress and indicate to them what skills they need to develop to become eligible for promotion, pay raises or both (Mathis & Jackson, 1982).

Promotion and Compensation

Performance appraisals are also used to determine whether employees should be promoted or have their salaries increased. To do this, superiors need to establish whether a particular employee's performance is above, below or at the expected level (Strauss & Sayles, 1980).

This approach to promotion and compensation supports the idea that raises should be given for merit rather than for seniority, where merit indicates that an employee receives a raise based on performance (Graves, 1982).

Counselling Employees

The fact that employees are counselled on different aspects regarding their jobs, could lead to an employee experiencing a higher degree of job satisfaction (Latham & Wexley, 1981).

During a counselling session, the appraiser should create an open and supportive atmosphere at the beginning of the appraisal interview. It should be made clear to employees that the emphasis of the discussion is on development and counselling. Thus, the purpose of appraisal is to help employees and the employee should also be allowed to air his/her views during the appraisal interview.

Manpower Inventory

Information emanating from the performance appraisal process also provides a useful input to the preparation of inventories of manpower skills and for manpower planning in general. Thus, persons eligible for promotion can be identified and succession

planning is greatly facilitated. Furthermore, by matching planned business strategies against the skills and abilities of different employees, management is in a better position to provide for orderly growth or change (Schuler, 1981).

Management Development

Performance appraisal provides a framework for future employee development by identifying and preparing individuals for increased responsibilities (Kirkpatrick, 1984; Slivinski, 1975). Employee development programmes can be designed to meet the needs of employees and the organization.

Communication

Performance appraisal provides a format for dialogue between superior and subordinate and could thus improve the mutual understanding of personal goals and concerns (Wells, 1982; Zippo & Miller, 1984).

Research

Slivinski (1975) is of the opinion that information obtained from a performance appraisal can often be used as a criterion to assess the validity of personnel selection and training procedures.

From the above discussion of the purposes of performance appraisal, it is evident that a variety of benefits are to be

derived from installing a good performance appraisal system. In the case of the truck drivers at the wine manufacturing company, the main objectives were to motivate employees, to improve communication and to find a more reliable and valid system for promotion and compensation. Of course, the aim of improving the public image of the organization stated at the outset of the study, remained on the agenda. To achieve the objectives of a performance appraisal system, it is however also important to pay attention to the evaluators who will be involved in goal setting and the achievement of those goals.

Performance Evaluators

Evaluation of employees' performance can be done by subordinates, peers, a combination of raters, clients or the direct supervisor of the employee. In Figure 2 possible sources of appraisal are indicated.

A distinction is made between external and internal sources for employee evaluation. A brief discussion of these different sources follows.

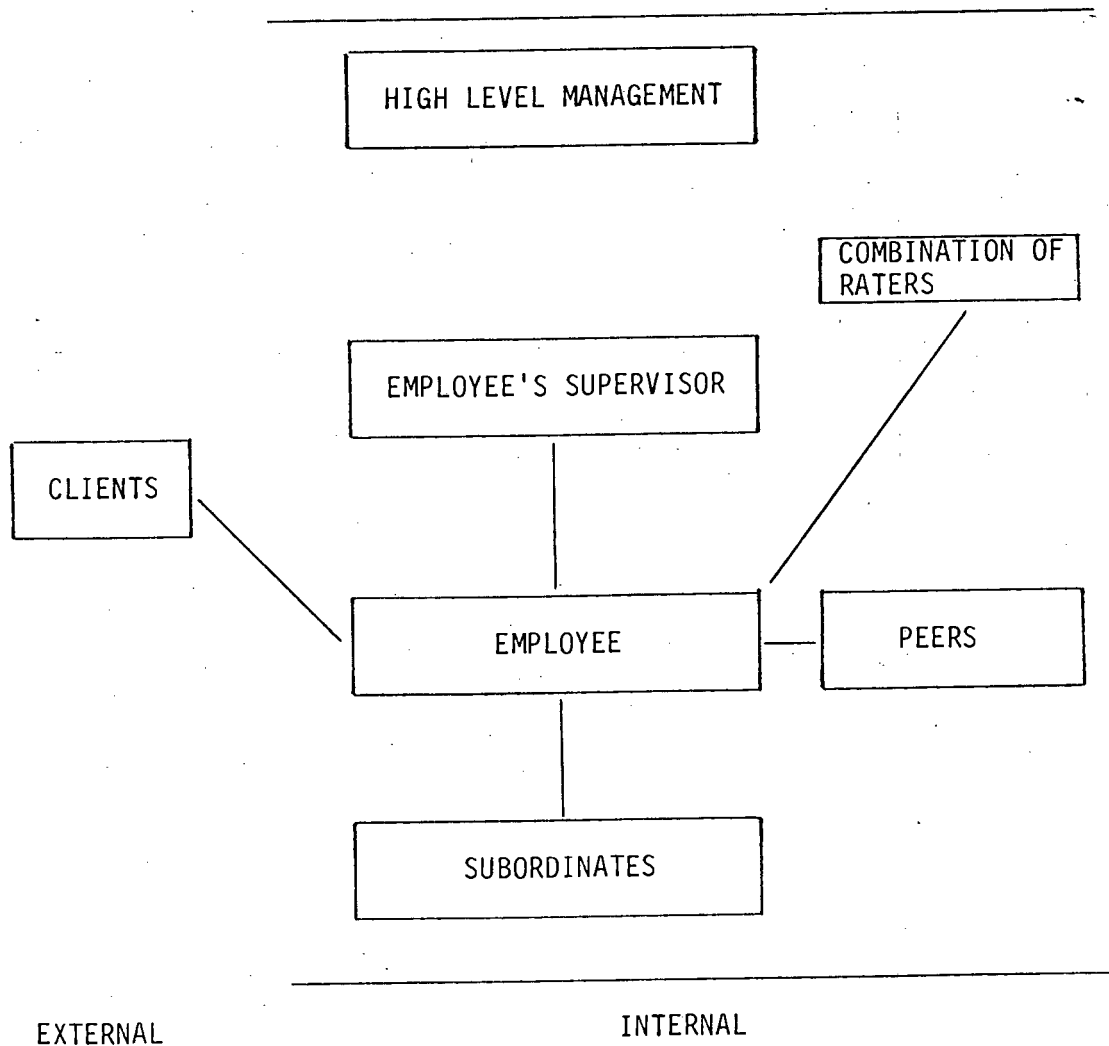


FIGURE 2: Potential Performance Appraisal Sources

Appraisal by Subordinates

This type of appraisal offers a different perspective on a superior's performance. The subordinate has access to "inside information" about a superior which may not generally be known by top management. The subordinate knows exactly the extent to which he/she plans and organizes, the type of leadership style he/she is most comfortable with and how well he/she communicates (Cascio, 1982). According to Beer and Ruh (1976), this approach is used more regularly by universities (students evaluating their lecturers) and large corporations, where a manager may have many subordinates.

There are, however, several significant disadvantages when using subordinates to appraise the performance of their superiors. Firstly, the subordinates may experience a request to formally appraise their superiors as threatening, because they could be reprimanded by their supervisors for an honest, but unfavourable appraisal. Secondly, subordinate appraisals tend to undermine a superior's legitimate, positional power, as well as his/her reward-and-punishment power (Cummings and Schwab, 1973). In the last instance, the possibility always exists that a subordinate's appraisal of his/her superior will focus primarily on the extent to which that superior fulfills the subordinate's needs, instead of emphasizing the organizational accomplishments of the superior.

The abovementioned method of appraisal was clearly not suitable for the present study. Truck drivers did not always supervise

the same employees, and the employees being supervised by them were not well enough qualified to give an objective evaluation of the drivers. Some of these employees did not have much schooling and they were appointed in the organization as general labourers, on job grade 1, the lowest grading of any employee within the organization (the grading system used by the organization is explained at a later stage). Thus, considering the abovementioned facts, the validity and reliability of an appraisal by subordinates, would surely be in question.

Peer Ratings

According to Cummings and Schwab (1973), two considerations would seem to facilitate the effective use of peer appraisals, namely: a high level of interpersonal trust and sharing among peers, coupled with a noncompetitive reward system and situations where information about an employee's performance methods or outcomes are uniquely available to his peers.

Research on peer appraisals has shown them to consistently meet acceptable standards of reliability (Gordon & Medland, 1965) and interobserver reliability (De Jung & Kaplan, 1962). In addition to their reliability, peer appraisals are also valid predictors of job performance, even more valid than supervisory appraisals (Korman, 1968). Cummings and Schwab (1973) stated that research on peer appraisals showed them to be predictive of success or correlated with both objective and other subjective evaluations of success in naval flight training and performance,

military officer performance, scholastic performance, field sales performance as well as middle-management performance.

Peer ratings were found to be more harsh than self-ratings (Kavanagh, MacKinney & Wollins, 1971). They also differed from supervisory ratings (Campbell, Dunnette, Lawler & Weick, 1970) in that peers and supervisors may perceive different aspects of an employee's performance, largely because employees behave differently when the boss is present.

A major drawback of peer ratings is the fact that in order for them to be valid, group members must have close contact with one another (Cummings & Schwab, 1973; Latham & Wexley, 1981). According to Mathis and Jackson (1982), most of the research on peer ratings was done on military personnel at the management or pre-management level where peer groups are closely knit. However, this does not hold true for peer groups in an industrial setting. In some organizations, it may be difficult to find peers who have first-hand knowledge of each other's behaviour. This in fact was the case in this study. A truck driver operated alone with a few subordinates, and his peers could only observe his behaviour when he loaded or unloaded cargo at the depot, if they happened to be performing the same function at the same time.

Another drawback of peer ratings may be the competitive nature of many organizational reward systems. A win-lose attitude amongst peers could inhibit honesty in appraising perceived rivals

(Cummings & Schwab, 1973). Peer rating may also cause psychological conflicts in the rater. Conflicts may arise between evaluating one's peer highly and increasing one's own chances of a large salary increase, or between evaluating him/her poorly and maintaining his/her friendship. According to Cummings and Schwab (1973) it is for such reasons that peer appraisals have been found invalid, or even disruptive in some organizations. Their arguments were applicable to the circumstances under which the truck drivers in this study operated. Thus peer ratings were ruled out.

Combination of Raters

In situations where employees work under supervision of more than one supervisor and where they are known by more than one, a committee of supervisors may be appointed to evaluate specific employees' performance. Thus more information about a particular employee may become available (Mathis & Jackson, 1982).

Schneier (1977) described five potential advantages when a combination of raters is used for performance appraisal. Firstly, it generates a larger data base on which personnel decisions can be made. Secondly, it helps to identify extremely biased or different ratings. In the third instance, it allows appraisal from multiple perspectives and in the fourth instance, it permits assessment of the reliability of ratings as well as the sensitivity

of the rating format. Lastly, it allows participation by others besides superiors in the appraisal process, which can foster commitment to the system.

Schneier (1977), however, also identified certain weaknesses when a combination of raters is used for performance appraisal. It may complicate the task of giving feedback on appraisal ratings, because interpreting ratings made by others can be difficult and inaccurate. Furthermore, to combine or average appraisal ratings can mask the differences that come from various perspectives. DeCotiis and Petit (1978) further stated that individual raters may feel less accountable for the appraisal in a multiple rater system, and therefore the accuracy of their ratings may decrease.

In the case of the truck drivers involved in this study, this method could not be applied because supervisors were only familiar with those drivers reporting directly to them.

Client Appraisal

Clients, forming part of the external environment of the organization, may also be used to assess performance. As a total assessment of an employee cannot be achieved in this way (a client does not know how an employee is behaving away from the client), this method is rarely applied.

Self-Appraisal

Self-appraisals seem to be justified when there are strong reasons to believe that the performer himself is in the best position to observe and assess his own methods of work and results. This may be applicable, for example, when a performer is working under conditions of extreme physical isolation, or is the unique possessor of a rare skill (Cummings & Schwab, 1973).

Several advantages have been found to be associated with self-appraisal, including more involvement and commitment of the performer to goals set for him/her (Schuler, 1981), more satisfying and constructive appraisal interviews, improved job performance and less defensiveness by performers regarding the appraisal interviews as well as the overall appraisal process (Cummings & Schwab, 1973). Latham and Wexley (1981) further stated that a supervisor may learn how an employee perceives the job responsibilities, performance on the job, as well as problems encountered in carrying out job responsibilities. Self-appraisals also help to clarify differences of opinion between the employee and the manager regarding job requirements and job performance (Basset & Meyer, 1968).

The abovementioned advantages have to be countered by the evidence indicating several problems associated with the use of self-appraisals. Firstly, several studies have found low agreement between self- and supervisory appraisals. However, according to Latham and Wexley (1981), the research is not consistent

as to the direction of the disagreement. Studies by Beatty, Schneier and Beatty (1977) and Thornton (1968) found self-ratings to exceed those given by managers. Kirchner (1965) also found that self-ratings were more lenient than supervisory ratings. In contrast, Heneman (1974) and Teel (1980) reported that self-ratings were harsher than ratings made by superiors.

Thornton (1980) reviewed the literature on self-appraisal and concluded that self-ratings generally tend to show more leniency, less discriminant validity, less reliability, less halo and less agreement with other sources, than other's ratings (eg. supervisor, peers and subordinates).

Appraisal by Supervisors

It is typical of most performance appraisal systems to have the evaluation of employees done by their immediate supervisors (Cascio, 1982). The supervisors are probably more familiar with the performance of each individual than anybody else, because they have had the best opportunity to observe the individual (Dessler, 1983; Wells, 1982). The supervisor also knows the job better than any other evaluator, having usually done the same job before being promoted to supervisor. He/she is probably also responsible for reward and punishment decisions such as pay, promotion or discipline, and must therefore be able to tie effective or ineffective performance to the personnel actions taken.

According to Guion (1965) ratings must come from people who are qualified to rate, the primary qualification being firsthand knowledge. A rater therefore can be qualified only on the basis of his/her own observation, and not if his/her evaluation is based on hearsay or prejudice. Typically, an immediate supervisor is best qualified for this function.

Despite these abovementioned facts, Barrett (1966) concluded that an employee's appraisal depends heavily on how each supervisor **thinks** the work **should be** performed rather than how well **it is** **actually** performed by an employee. Latham and Wexley (1981) also stated that supervisory appraisals are frequently loaded with subjectivity and bias and that they are neither as reliable nor valid as peer ratings.

The abovementioned discussion focused mainly on the individuals more directly involved in the performance appraisal process, recognizing that appraisal is to some extent a personal event. The role of appraiser is typically filled by the direct supervisor, often imperfectly (Landy & Farr, 1983). Alternative sources of raters were examined, including the employee's subordinates, peers, a combination of raters or the employee him/herself. In this study, none of these alternatives were likely to generate more accurate ratings than the direct supervisors with their knowledge of the truck driver's job and frequent contact with the different drivers. It was therefore decided to use the direct supervisors of the drivers, the transport foremen, to appraise the drivers' performance.

Having reviewed the purposes of performance appraisal and the options available to an organization for selecting raters, it seems appropriate to also pay attention to some of the problems or obstacles to the smooth functioning of a performance appraisal system. Some solutions for overcoming these problems have also been suggested.

Potential Performance Appraisal Problems and Proposals for Solving Them

A variety of problems may be encountered along the way of designing a performance appraisal system which would satisfy the needs of both the organization and employees. A problem-ridden system is likely to be rejected by the organization as well as by the employees being evaluated.

The problems associated with performance appraisal can be divided into two broad categories: those related to maintaining the performance appraisal system and rater problems. Following a description of these two categories, possible solutions are described.

Maintaining of Performance Appraisal Systems

Many performance appraisal programmes have died on the vine, because at the initial introduction of such a programme, no systems

were set up to keep the programme going. Managers returned to their jobs after the introduction of a new appraisal system and nothing more was heard about performance appraisal from the personnel department or any of the other people involved in the performance appraisal. To help ensure that performance appraisal programmes are utilized on a continuous basis, top management must provide encouragement and ongoing support for such programmes (Haynes, 1978).

There are a variety of methods available to help a performance appraisal system survive. The responsibility for this falls mainly on the human resources department who should ensure that other persons involved in the appraisal process, like line managers for instance, regularly conduct performance appraisals. Some proposals for keeping a performance appraisal system active are discussed below.

Firstly, it should be remembered that performance appraisal is not an unique event, but that it is an ongoing process. Therefore, to implement such a system and continue using it, management must support the system. A typical way to achieve this is for the personnel department to send out a reminder to managers and supervisors to notify them of the next due date for performance appraisal. Another way to ensure that superiors definitely do performance appraisal at the scheduled times, is to include a specific performance criterion on their performance appraisals

- the superiors are then rated on the frequency with which they do performance appraisals.

Secondly, as stated earlier, support from line managers and supervisors can also help to keep a performance appraisal system active (Sausser, 1980). Initially this can be achieved by involving them in the development of such a system. Thirdly, a performance appraisal system can also be kept active by making sure that employees get constructive feedback on their performance. They must derive some benefits from good performance, even if it is just verbal reinforcement (Strauss & Sayles, 1980).

In the fourth instance, management must also be prepared and committed to spend the time and money required to make an appraisal system work. Therefore various subsystems in the organization should be initiated to support a meaningful performance evaluation system (Schuler, 1981). This implies that in the case of low performing individuals some options for action should be open: The individual may undergo training to add new skills or be transferred to an area where his/her abilities and skills will be better utilized. Failing that, the person may be demoted or have his/her services terminated.

Rater problems with Performance Appraisal

According to Latham and Wexley (1981), rating errors are errors in judgement that occur in a systematic manner when individuals

observe and evaluate other individuals. Feldman (1979) in Latham and Wexley (1981), defined rating errors as a difference between the output of a human judgement process and that of an objective, accurate evaluation not influenced by bias, prejudice or other subjective, extraneous influences. Wexley, Sanders and Yukl (1973) stated that a factor which often complicates the problem with rater errors, is that raters or observers are usually unaware that they are making them. Therefore, it is also very difficult to correct these errors.

Robbins (1982) said that every evaluator has his/her own value system which acts as a standard against which appraisals are made. Relative to the true or actual performance an individual demonstrates, some evaluators may rate high and others low, thus committing a rating error. If all individuals in an organization were appraised by the same person, there would be no problem. Even though there would be an error factor present, it would be applied equally to everyone. The difficulty arises when different raters make different rating errors. Thus ratings of individuals will differ, some being high and others low. By comparing different appraisals of one individual, it is possible to determine whether rater errors have taken place (French, 1970). If a rater rates everybody in the same way (for example, a rater consistently rates low, with no discrimination between different items) there is also strong cause to believe that a rating error has taken place.

Rating errors thus occur as part of human nature, that is, they reflect the personality make-up of the rater. Therefore there is no single way to eliminate these problems, personality being such a complex concept. The more general rater errors will be discussed and proposals for solving these problems will be described when discussing each specific error. Some general solutions regarding most of the rater errors will be discussed once the errors have been identified.

Halo error

The Halo error occurs when a rater only has one characteristic of the employee in mind while rating. He/she then rates an employee high or low on all items because of this one characteristic. In other words, the rater's judgement on each aspect of performance is dependent upon or related to his/her overall judgement of the individual (Cascio, 1982; Dessler, 1983; Graves, 1982; Slivinski, 1975).

The following are indications of halo error being present in performance appraisals. According to Jacobs and Kozlowski (1985), the oldest indicator of halo error is the presence of moderately high to high values in the intercorrelation matrix of rating dimensions. Bernardin (1978) and Bernardin and Pence (1980) stated that another approach to determine whether halo error is present, is to focus on the variance or standard deviation

of a single rater's ratings of a ratee across performance dimensions. Smaller standard deviation or variance estimates are viewed to be as indicative of greater levels of halo error, since they reflect a unitary conception of a ratee.

Birkenbach (1984), a South African author, compared halo error in two types of performance appraisal systems, namely the behavioural observation scale and the graphic rating scale (these techniques and the development thereof will be discussed in the next chapter). The development of the behavioural observation scale was based on critical incidents obtained from seventeen line supervisors during interviews, whilst the graphic rating scale was developed by using the criterion headings of the observation scale as the rating dimensions. In the study, Birkenbach defined halo error in terms of the sizes of the inter-correlations between the dimensions of each format. In this instance, the greater the size of the correlation coefficient, the greater the halo was assumed to be.

The halo error can be illustrated with the following example: a rater had to rate a salesman on three criteria, namely initiative, punctuality and work knowledge, using a 5-point scale. Although the salesman attained a score of five on the criterion of job knowledge, he should actually have received a score of one and two for initiative and punctuality respectively. The rater's

judgement was influenced by the high score of five on job knowledge and the salesman was given a corresponding score of five on the other two criteria, instead of the score he actually deserved.

To eliminate halo errors is rather difficult. As was mentioned earlier, the training of raters may help to solve this problem. This aspect will be discussed in more detail later in this chapter. Another procedure for reducing the halo effect is to have all subordinates evaluated on one dimension before proceeding to the next dimension. This procedure forces the rater to think in terms of only one dimension at a time, rather than allowing the contaminating effect of other dimensions to influence his/her judgement (Ivancevich & Glueck, 1983).

Another method frequently used to deal with this type of error, is to stagger the questions on the evaluation so that a favourable answer for, say, question number 2 might be five on a scale of one to five, while a favourable answer for question number 10 might be one on a scale of one to five (Robbins, 1982). Structuring the questions in this way attempts to reduce the halo effect by requiring the evaluator to consider each question independently.

Recency of events

This problem occurs when recent performance is viewed as more important than performance which occurred long ago. For

example, if a subordinate did an outstanding job the previous day or week, it can offset a mediocre performance for the rest of the previous year. This error, caused by recency of events, is understandable because of the difficulty in remembering things that happened long ago (Mathis & Jackson, 1982; Slivinski, 1975; Spool, 1978).

According to Ivancevich and Glueck (1983) this problem may be solved by using specific performance appraisal techniques such as critical incidents or management by objectives. These techniques will be discussed in Chapter 3.

Central tendency

Such a distribution results when on a 5-point scale, only the three centre points on the scale are used when evaluating an employee. Raters are unable to commit themselves to high or low ratings in cases where individuals actually deserve such scores. By doing this, there is no discrimination between good and poor aspects of a certain employee's work and no or little distinction is made between a good and a poor worker (Ivancevich & Glueck, 1983; Wells, 1982). Employees are thus almost always rated as average workers. According to Birkenbach (1984) central tendency can be assessed by studying the kurtosis of the distributions produced by each performance dimension on an appraisal instrument.

A negative kurtosis is indicative of a flat curve which has ratings distributed across the range of possible scores (platykurtic curve), whereas a positive kurtosis indicates a tall curve with ratings distributed only within a narrow range of scores (leptokurtic curve). Central tendency as it may occur in performance appraisal is graphically illustrated in Figure 3.

Statistical corrections (Birkenbach, 1984; Robbins, 1982) or thorough training of raters (Alewine, 1982; Latham & Wexley, 1981) may alleviate this problem. Sometimes even just making the rater aware that this problem does exist may overcome it (Wells, 1982).

Leniency or harshness error

The different distributions resulting from leniency and harshness are shown in Figure 4. In this case raters tend to rate individuals higher or lower than they actually should.

Leniency error occurs during performance appraisal when the rater artificially assigns all employees high performance ratings, and all scores cluster at top levels of the measuring instrument. The harshness error in contrast occurs when the rater assigns all employees low performance ratings, and all scores cluster at the bottom levels of the measuring instrument (French, 1970; Robbins, 1982; Schuler, 1981; Slivinski, 1975).

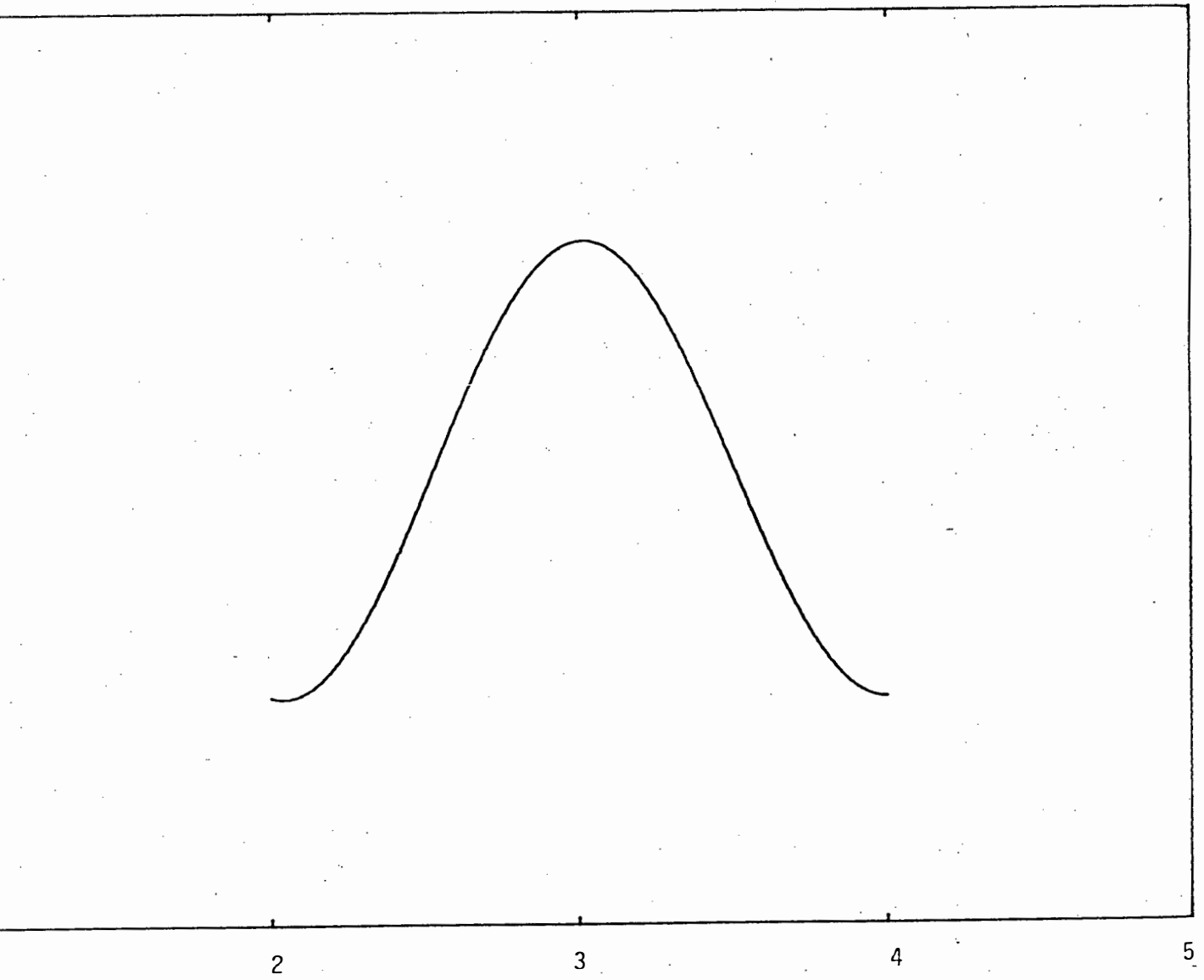
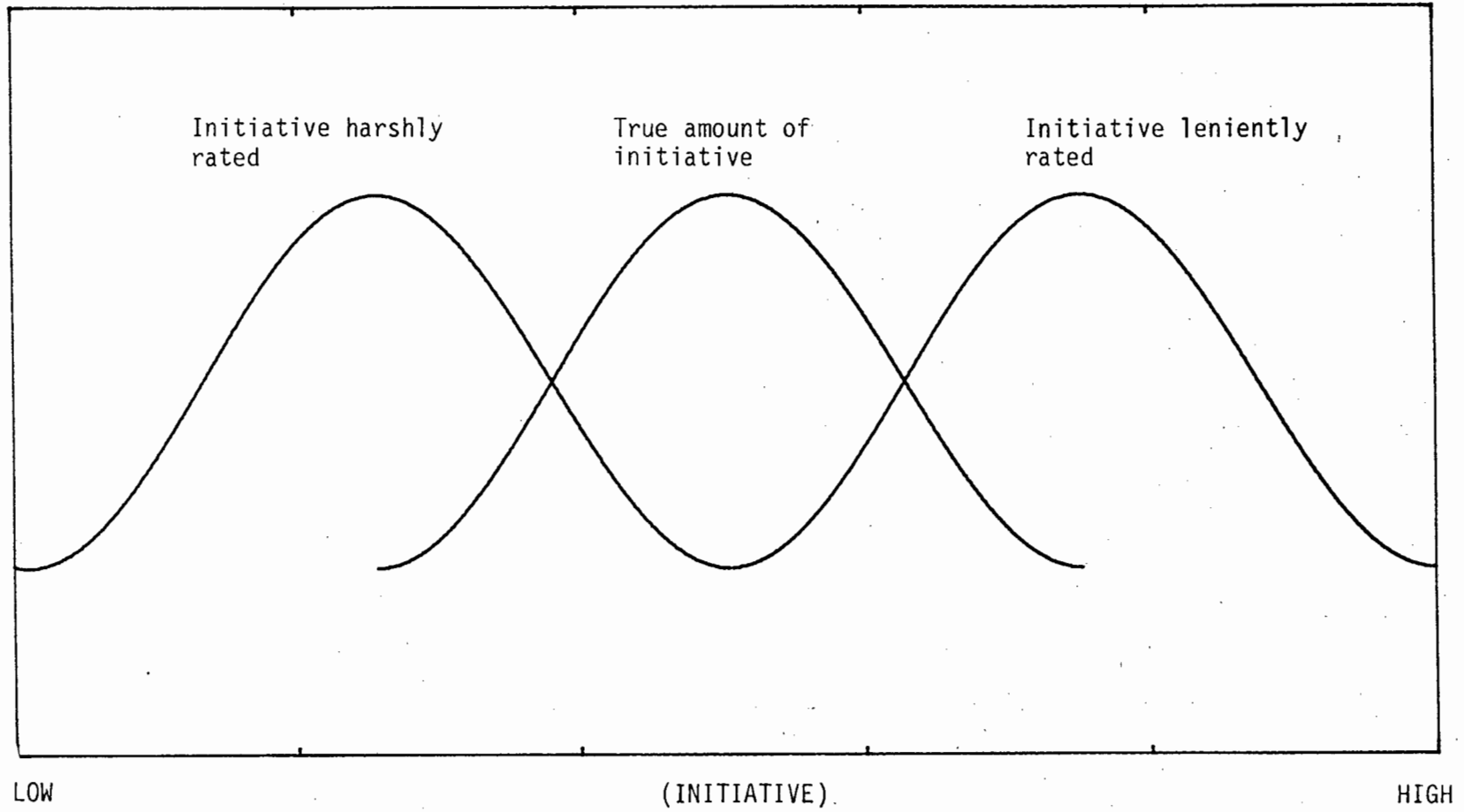


FIGURE 3: Central Tendency Error

FIGURE 4: Distribution of Lenient and Harsh Ratings



Birkenbach (1984) stated that leniency or harshness can be assessed by investigating the degree of skewness of the distributions for each performance dimension on a specific rating scale. A positive score implies that the distribution is positively skewed, which is to be expected at harsh ratings. A negative score implies lenient ratings while a skewness score of nil indicates that the ratings are spread in the form of a normal distribution.

To solve this problem, raters should examine their own ratings for harshness or leniency when evaluating employees. Another solution may be grading on a curve, where raters are forced to have a specified number of employees in each category on a normal distribution curve (Ivancevich & Glueck, 1983). This, however, does not necessarily imply that a more reliable evaluation of an employee will be obtained. By enforcing a normal distribution on the results, average performers may be pushed into a higher or lower category for lack of space in the middle category.

Added to that, there is always the possibility that a whole team supervised by one individual may be below average. By forcing a normal distribution, one loses out on the possibility of comparing more than one work team.

Personal bias

Raters may like or dislike certain employees more than others, and this can influence the ratings they give.

Rater bias is very difficult to overcome, especially if a manager is not aware of such a situation or if he/she does not admit the existence of such biases. Examination of the ratings by higher level managers may help to solve this problem (Mathis & Jackson, 1982). Certain evaluation techniques such as forced choice, performance tests and management by objectives tend to reduce this problem.

The one asset individual

According to Slivinski (1975) the person with the impressive appearance, the graduate of the boss' own alma mater, the good talker or the person with the advanced degrees, may tend to get a more favourable rating than the subordinate who often lacks these irrelevant attributes.

Contrast

The contrast error is the tendency for raters to compare subordinates with one another rather than to rate them in terms of the requirements of the job. This type of rating is unfair, especially if performance is supposed to be rated according to the objectives and standards that have been agreed upon by the management and subordinates.

Projection

The more closely an employee resembles the rater in attitudes, behaviour and background (e.g. from the same school), the stronger is the tendency of the rater to judge that individual favourably. It is also known as the similar-to-me effect (Slivinski, 1975). For example, a supervisor/manager who sees himself/herself being firm and willing to take a stand, values this characteristic and believes it makes him/her a good leader. When he/she sees a subordinate who is also willing to take a stand, he/she will automatically believe that the person is a good leader, whether this is actually the case or not.

Effect of past record

In this case the individual's past performance contaminates the rating of his/her present performance. If an individual did well in the past, a manager may make the assumption that he/she is performing well at the present time (Graves, 1982; Slivinski, 1975).

The most general rating problems encountered during a performance appraisal, have now been discussed. The following are some solutions to these problems.

Proposals for Solving Rater Problems

There appears to be no single solution to these rater problems, but to create an awareness of their existence may already contribute to some extent to a solution. Birkenbach (1984), a South African author, suggested three approaches to solve rater problems.

Firstly, it could be done statistically. This is particularly relevant in the case of problems such as central tendency errors, or leniency and harshness on the part of the rater as discussed earlier in this chapter.

Secondly, Birkenbach (1984) suggested the development of different formats of performance appraisals. Examples of such formats are behaviourally based appraisal instruments, such as the Behaviourally Anchored Rating Scale and the Behavioural Observation Scale. The main characteristics of these scales will be discussed in the next chapter. The main advantage of these systems is that rating errors are claimed to be reduced to a minimum. This can be achieved because levels of performance are better defined, performance dimensions are closely specified and raters tend to be more cooperative and attentive to the task of rating.

In the third instance, Birkenbach (1984) supported an opinion held by various other researchers, that the best way to eliminate

rating errors, is to train the raters (Alewine, 1982; Borman, 1979; Dessler, 1983; Eichel & Bender, 1984; Graves, 1982; Latham & Wexley, 1981; Sauser, 1980). However, this aspect of performance appraisal is still a very controversial one. Contradictory results have been obtained from various studies.

Several studies indicated that no significant improvement in results occurred after raters had been trained. Bernardin and Walter (1977) found that no significant improvement regarding accuracy in judgement, interrater reliability, assessing discrimination across raters and other rater errors, was found after raters had been trained.

In another study, Borman (1975) reported that the training of raters was only partially effective in reducing the halo error. This study investigated the effects of a short training session which was designed to reduce halo error. Ninety people in low and middle management positions were involved in the study. They had to rate hypothetical first-line supervisors on six performance dimensions prior to and following the five minute training session. Whereas the halo effect was reduced significantly, the validity of the ratings was generally unaffected. The performance ratings after the training session also showed lower reliability than previous ratings.

It may be argued that the lack of improvement in the accuracy of performance ratings as mentioned above, may be due to the

fact that the wrong training methods were used. Wexley, Sanders and Yukl (1973) found that specific training techniques were more successful in reducing rater errors than others.

Latham, Wexley and Purcell (1975) expressed the point of view that pure knowledge of rating errors alone will not lead raters to take effective steps to counteract these errors. In their study they compared several training methods on performance appraisal errors. Sixty managers in a large organization were randomly assigned to a control group, a group discussion or a workshop. The group discussion and the workshop involved training which was directed toward the elimination of rating errors which frequently appear in performance appraisals. These included contrast effects, halo error, similarity and first impressions.

The workshop training method included videotapes of hypothetical job candidates whose performance was being appraised by a manager. The trainees then had to indicate how they thought the manager in the videotape evaluated the candidate. Afterwards a group discussion followed regarding each trainee's decision. Thus, trainees had an opportunity to practice and to receive feedback.

In the discussion group, definitions of specific rater errors were presented to the trainees. Trainees were also given an example of errors which could occur in each one of the situations, namely the performance appraisal situation, the selection interview

and an off-the-job situation. The trainees then discussed amongst each other personal examples of rating problems.

The results of a follow-up study six months later, indicated that observations of trainees in the control group were characterized by similarity, contrast and halo errors. The trainees in the discussion group still tended to make impression errors, whilst the observations of trainees in the workshop were relatively free of the errors discussed previously.

Spool (1978) however criticized the study for the fact that simulations were used rather than measuring actual on-the-job-behaviour. Another limitation of the study may have been caused by discussions amongst the participants during the six months between the testings, thus influencing the results. Spool (1978) pointed out that even Latham et al. (1975) were unclear as to what features of the workshop were responsible for its success.

Spool (1978) also reviewed the literature on rater training for the last 25 years up to 1978. He indicated studies which improved rater accuracy, as well as studies where rater accuracy remained the same after training the raters. He found that research done on this subject was not really authoritative, because of the contradicting results of as many different studies. His

most important finding was that with some exceptions, training programmes reviewed were generally effective in increasing the accuracy of employee observation, and came to the conclusion that some training is better than none.

Recently, Banks and Roberson (1985) also compared three performance appraisal training systems, namely the group discussion, the workshop and rater accuracy training. Their results supported the research of Latham et al. (1975). They found that the workshop was the best method for training raters to reduce errors during performance appraisals.

It is thus obvious that the effectiveness of rater training is still a very controversial aspect surrounding performance appraisal. The present study acted on the premise (Spool, 1978) that some training is better than none. This will be discussed in more detail later.

Another aspect of performance appraisal which is very important to any such system, is the acceptability thereof to all those involved with it. To be acceptable to managers as well as employees, a performance appraisal has to meet certain requirements and criteria. A discussion of this aspect follows.

Criteria for Successful Performance Appraisal Systems

As stated before, if a performance appraisal system is to be accepted by all involved, it has to meet certain criteria to fulfill the expectations of those appraising as well as those being appraised. This is probably the most important aspect of any performance appraisal system, because if an appraisal is not generally accepted, it is bound to fail as a management tool, which is exactly what happened in the organization involved in this study. Nobody viewed the appraisal system which had been in use prior to this study as sufficiently useful to warrant time and effort.

There are various criteria which contribute to the acceptance of performance appraisal systems by all involved. The criteria generally accepted by personnel practitioners include job relatedness of the appraisal system, a clear definition of successful performance, reliability, validity and practicability.

Job Relatedness

In the first instance, a performance appraisal should be job related (Richardt, 1976). As mentioned earlier, the best performance appraisal is one which is tailormade for a specific job.

According to Richardt (1976), the appraisal system must be based on the behavioural aspects of the job, because this

increases the reliability of such a system. A behaviour-orientated appraisal system has some advantages over other appraisal systems (Borman in Landy, Zedeck & Cleveland, 1983; Latham & Wexley, 1981; Richardt, 1976). These advantages will be discussed when describing the different performance appraisal techniques.

Some questions which must be answered regarding a job related performance appraisal are: What are the goals of the activity? What level of performance should be obtained on these goals? What behaviour does an employee on this job have to engage in to achieve the desired goals and how important is each one of these behaviours (Clayton and Gatewood, 1981; Wells, 1982).

Definition of Success

A second criterion requiring attention is the existence of a clear definition of success. The system must differentiate between successful and unsuccessful behaviour and both these concepts must be described in clear understandable terms (Cascio, 1982).

Wells (1982) supported this when he stated that every employee must know the standards against which he/she will be evaluated, in other words, what the specific definition of success is. Therefore employees should have prior knowledge of what is expected of them on the job. Sauser (1980) confirmed this by stating

that standards, expected results and goals should all be identified or negotiated at the beginning of the appraisal period. If there are changes in the required goals during the appraisal period, managers should make these known and be sure they are understood.

Reliability and Validity

Another criterion for a good performance evaluation system is its reliability. The question to be answered is: Would two people having the same amount of knowledge about an individual's performance evaluate the individual in the same way? (Biesheuvel, 1984; Eichel & Bender, 1984; Graves, 1982; Landy & Farr, 1983; Latham & Wexley, 1981; Slivinski, 1975).

The criteria selected for a performance appraisal system should also be easily observable and measurable, to increase the reliability of the system. In other words, a supervisor must be able to observe an employee engaging in certain behaviours included on the appraisal form so that he/she could make a valid judgement about the employee's worth to the organization. For example: It may be easier to evaluate a person on his/her attendance of work than on his/her friendliness.

Regarding validity, it can be said that a performance appraisal should contain a representative sample of what an individual must do on the job. The appraisal should furthermore be able

to predict future job success. These two kinds of validity, the first content- and the second criterion-referenced, are necessary to ensure a proper performance evaluation system (Dessler, 1983; Eichel & Bender, 1984). The different aspects of reliability and validity regarding the appraisal instrument to be developed in this study, will be discussed later.

Standardization

If performance appraisal information is used to compare employees across jobs, or in other words, one appraisal system is used for various different jobs, the conditions and procedures surrounding the rating process must be the same for all employees.

According to Latham and Wexley (1981), standardization refers to minimizing differences between rating methods of different raters when administering and scoring the appraisal instrument. However, it has been indicated that an appraisal system developed for a specific job is superior to one developed for use across jobs (Richardt, 1976; Slivinski, 1975).

Practicability

Another criterion is practicability. For the performance appraisal system to be used it must be based on the needs of line management (Cascio, 1982; Hyde & Smith, 1982). Richardt

(1976) stated that the involvement and acceptance by line management, as well as the involvement and acceptance by subordinates could also determine how successful the performance appraisal system will be. Because the line managers are usually the people who have to do the appraisal, the system must be acceptable to them. If they do not agree with the system, or see serious flaws in the system, they will lose confidence in it. They will not see it as a management tool to be used to determine the merit of each employee or to develop the subordinate to be more productive. The danger always exists that the performance appraisal is developed by the personnel department without consulting the line managers first, and therefore the system may be unacceptable to them.

As in the case of line managers, the subordinates must also accept and believe in the performance appraisal system, especially where the subordinate receives feedback on his/her performance and the evaluation thereof. No employee would like a poor performance record and if he/she thus feels that the performance appraisal system does not give a good reflection of his/her work, the employee may reject it.

Formal Documentation

This documentation must describe the required information about the employee who is being evaluated. Information should include written identification of the responsibilities assigned

to the employee, an identification of the results to be attained, standard output in terms of quality and quantity as well as the way results will be viewed in making evaluative judgements (Graves, 1982).

Acceptability to Trade Unions

In the case of this study, trade unions did not influence the appraisal system directly, as only employees in jobs up to job grade 6 belonged to a trade union. Truck drivers were appointed in grade 7.

However, if an appraisal system is used for job grades lower than 6, trade unions and their policies will definitely have to be considered. Not much attention will be given to this topic here, because it has little relevance to this study. However, it may be important to mention that trade unions and organizations usually have a conflict of interest when it comes to employees. Trade unions are more interested in the job security of employees, whilst organizations stress the human capital investment, which infers the logical necessity of dispensing with people's services. Thus, performance appraisals cause conflict, because the inadequate worker may be asked to leave the organization, while this is just what trade unions strive to prevent. The attitude of a trade union to performance appraisal, therefore, may differ significantly from that of the organization. This is the kind of problem which

cannot be addressed without taking the interests of both the union and the organization into account.

Employee Participation

Another important aspect of performance appraisal is employee participation in the development of a performance appraisal system.

In general, the role of the appraisee in performance appraisal has been and continues to be predominantly passive. While it appears as if participation in the feedback discussion is the most frequent avenue of employee participation, employees can also participate in various stages of performance appraisal, namely : goal-setting, criteria development, data collection, problem solving, self-rating and feedback.

The important question to address is : What impact does employee involvement have on the appraisal process? In this regard Wexley, Singh and Yukl (1973) reported greater motivation by employees to improve. It was also reported that both managers and employees were more satisfied with the appraisal process after participation by the employees (Blake & Mouton, 1961; Fletcher, 1973; Wexley et. al., 1973). Blake and Mouton (1961) further reported greater feelings of unity and team spirit by managers and employees.

Latham and Wexley (1981) stated that the **setting of specific goals** to be achieved by the subordinate is also a very important

aspect of a performance appraisal system. This supported previous research by Smith and Brouwer (1977) who stated that: "goalsetting is at the heart of the performance appraisal process, for it gives meaning and relevance to appraisal and development activities" (Smith & Brouwer, 1977, p. 77).

Latham and Yukl (1975) reported inconsistent results after reviewing studies which measured the effect of participation in the actual goal setting process. However, a later study by Latham and Saari (1979) illustrated the importance of participation more positively. They found that individuals who were involved in the actual setting of goals and were given support by the experimenter, set high goals and actually performed at higher levels than those who were not involved.

Thus, involving the employee in the appraisal process is an investment with both costs and returns. The costs include sharing control of the performance appraisal process with employees and the required time to involve the employee. The return is likely to be a psychological one - both the manager and the employee will probably be more satisfied with the entire process, and more open communication could be created.

Over and above these generally accepted criteria as mentioned above, some authors also described further requirements which deserve mention.

In addition to goal setting subordinates should also be allowed to discuss their own feelings regarding their performance during the appraisal. According to Latham and Wexley (1977), this will lead to employees being more satisfied with the whole practice of performance appraisal.

According to Cascio (1982), the most important aspect of a good performance appraisal system, is its **relevance**. This implies that there should be a clear linkage between performance standards and organizational goals, and that there is a correspondence between the critical job elements identified through a job analysis and the dimensions to be rated on an appraisal instrument (Cascio, 1982). Management basically has to indicate what it wants from a performance appraisal. The basic purposes of the performance evaluation programme have to be clearly defined and the procedures and techniques employed adapted to them. The specific limited purposes must also be co-ordinated so that the total programme contributes to the achievement, both of major organizational and individual goals of those involved.

A realistic programme must therefore be established which contains policies, procedures, standards and practices that support the programme's objectives. Such a programme will include the design of the day-to-day mechanics of the administration of the appraisal, as well as the steps to involve all who are affected by it, in its final administration.

Relevance also implies that job analyses, performance standards and appraisal systems should be reviewed and updated periodically.

Banks and Roberson (1985) provided us with an overview of criteria for successful performance appraisal instruments. Based on their study, they compiled a list of criteria indicating whether a performance appraisal system could be considered as "good" or not. They categorized the criteria under four main headings, namely domain clarity and item development, item selection and analysis, item scoring and interpretation and contaminating effects. These categories are listed and defined in Figure 5.

The aforementioned criteria will all be kept in mind when specifying the approach to developing a new appraisal instrument and it will be attempted to follow the prescriptions of these criteria as closely as possible.

Summary

To summarize, this chapter dealt in the first instance with the purposes of performance appraisal. The discussion then focused on raters and some problems with regard to performance rating as well as some solutions for solving or overcoming these problems. The chapter concluded with some proposals for workable or "good" performance appraisal systems.

<u>CRITERIA OF GOOD APPRAISALS</u>	<u>DEFINITION</u>
A. DOMAIN CLARITY AND ITEM DEVELOPMENT	
Representative sampling of behaviour domain	The extent to which items cover the range of behaviours included within the construct.
Multiple Items	The degree to which construct is measured by several items.
Construct well defined	The degree to which it is clear what the construct does and does not cover.
Items developed from input of several experts	The degree to which items do not reflect a single ideosyncratic view.
B. ITEM SELECTION AND ANALYSIS	
Interitem correlation	Internal consistency of items
Item discriminability	The degree to which the item correctly separates people into high and low criterion groups
Item validity	The degree to which item is related to other meaningful constructs
C. ITEM SCORING AND INTERPRETATION	
Standardized scoring procedure	A standard approach to scoring items
Uniformity of test procedures	Consistency in the application of test procedures
D. CONTAMINATING EFFECTS	

FIGURE 5

Criteria for good Performance Appraisal Systems

(From: Banks & Roberson, 1985, p. 130)

The next chapter will deal with various performance appraisal techniques. These techniques will then be assessed in terms of their ability to meet the criteria set in this chapter, as well as those of the organization for appraising the performance of truck drivers.

CHAPTER 3

Different Approaches to and Techniques of Performance Appraisal

To develop any of the most widely used appraisal systems, three approaches are commonly used. These approaches will be discussed before going on to discuss specific performance appraisal techniques. At the end of the chapter, these different appraisal techniques will be compared on some of the criteria described in the previous chapter.

Approaches to Performance Appraisal

Broadly speaking, there are three approaches which can be followed when developing a performance appraisal system. In other words, each appraisal technique (as will be described later in this chapter) can be approached in three ways, namely the trait-oriented, the results-oriented or the behaviour-oriented approach. The approach used to develop an appraisal system will be determined by the specific situation for which such a system is to be developed, and therefore it is clear that different appraisal systems will be approached in different ways. Thus, the approach used for a specific performance appraisal system will determine the information which appears in the system.

Trait-oriented Approach

In the past, employee performance evaluation systems almost exclusively dealt with personal traits (Slivinski, 1975). Traits such as commitment, decision-making and initiative were included in many performance appraisal systems, but these words were surrounded by ambiguity. To tell an employee to show more initiative or leadership may sound like good advice, but it does not tell the individual how to deal with the advice. These concepts must thus be defined explicitly for the employee (Latham & Wexley, 1981). According to Graves (1982), if people are rated favourably on their personality traits, it does not necessarily imply that their work performance will also be of a high standard.

Smith and Brouwer (1977) confirmed this by stating that one of the major weaknesses in appraisal programmes has been that performance appraisal has been treated as a method to determine and rate personality traits. According to them, this fact of having to rate employees on their traits created immediate conflict in the rater, who resists being cast in the role of a psychologist. They further stated that trait rating is based on the erroneous belief that personality consists of discrete, easily identifiable characteristics.

The trait-oriented approach does have some advantages. In the first instance, these appraisal systems can be developed

very quickly, and in the second instance, they can be used across jobs (Latham & Wexley, 1981).

This, however, implies that they are not developed for a specific job and that the performance criteria will be too vague. In other words, actual performance will not be measured. For example: The word, initiative, could be used for various jobs, but it should be established whether this word really gives an indication of what an employee is worth to an organization, seeing that it could be interpreted in various ways by different people. Concepts of this kind remove the objectivity of a rater and force him/her to make subjective evaluations (Baron, 1983).

Results-oriented Approach

This approach to performance appraisal concentrates mainly on the results being achieved by the employee. However, it has recently been realized that the management by results approach is too rigid and shortsighted and that it basically falls into the same kind of trap as those focusing completely on personality (Slivinski, 1975).

Senior level management, stockholders, and consumers are generally concerned with the economic and cost-related outcomes of an organization. In other words, they are concerned with quantitative measures of performance outcomes, such as profits,

costs and returns on investment (Latham & Wexley, 1981; Slivinski, 1975). The measures mentioned above usually serve as excellent indicators of an organization's effectiveness. However, by themselves, they remain inadequate indicators of an employee's job effectiveness. There are several reasons for this.

In the first place, it is difficult to obtain cost-related measures for employees in different jobs. Some jobs may consist of more easily measurable results than others. For example, it is easier to measure results achieved by a vehicle salesman than those achieved by a personnel manager.

A second problem with the use of cost-related measures is that they often take factors into account for which the employee cannot be held responsible. Because of a machine breakdown, an employee's production figures may drop in one day and his/her monthly production figures may be influenced. Other examples of these factors may include the lack of tools or equipment, availability of materials and the amount of supplies, which could all contribute to the effective performance of employees.

In the third instance, a results-at-all-costs mentality may develop in an organization and this can run counter to corporate ethical policies, as well as to the overall productivity of the organization. Personal feelings and interests may be ignored in such an approach which may result in a decrease in productivity (Latham & Wexley, 1981).

Lastly, economic measures on performance outcomes by themselves may not inform employees what they need to do to maintain or increase productivity (Latham & Wexley, 1981). A good example of a results-oriented approach is the management by objectives technique, to be discussed later.

Behaviour-based Approach

As a result of the abovementioned problems with the two approaches, psychologists have increasingly stated their desire to measure and evaluate the employee in terms of observable behaviours that are critical to job success or job failure (Latham & Wexley, 1977). Behaviourally-based appraisal methods can account for far more job complexity, they can be related more directly to what the employee actually does and they are more likely to minimize irrelevant factors not under the control of the employee than the other two approaches (Alewine, 1982; Latham & Wexley, 1981). Added to this, behavioural criteria not only assess the employee in terms of factors over which he/she has control, but also specify what the person should do or not do to attain the required outcome. Thus, behavioural measures based on a job analysis indicate precisely what should be done by an employee to warrant recognition, discipline, transfer, promotion, demotion or termination (Beach, 1980).

According to the study done by Banks and Roberson (1985) previously mentioned, it was clear that behaviourally-based performance

appraisal systems were superior in many ways to other appraisal formats. They pointed out that because behaviourally-based measures are based on job analyses and are developed by "experts" representing relevant viewpoints, these measures most probably possess construct validity.

The biggest disadvantage of a behaviour-based approach to performance appraisal is the time and resources needed to develop meaningful appraisal systems (Graves, 1982; Robbins, 1982; Schuler, 1981). Usually, after a job analysis has been done, a job description is written, specifying behaviours to appear on the appraisal instrument. Behavioural-oriented appraisal systems are developed for specific jobs, therefore since the job analysis performed for one job is not valid for another, the entire process must be repeated for every single job. Furthermore, large samples of managers or experts are needed to make repeated judgements about desired and undesired behaviours, thus increasing the cost of developing such systems.

If all the abovementioned facts are taken into account, it seems as if the most meaningful approach to any performance appraisal technique would be the behaviour-oriented approach. This approach appears to concentrate more on the actual performance of an employee, in contrast with the other approaches which concentrate on the personality traits and the cost-related outcomes of an employee. However, these other two approaches also have some

advantages as indicated earlier, and will therefore be kept in mind when developing the new appraisal system for the organization.

The various performance appraisal techniques will now be discussed and the extent to which they meet some of the criteria for "good" appraisal systems previously described, will be indicated.

Different Performance Appraisal Techniques

According to Richardt (1976), performance appraisal is more a process than a technique. The critical element involved is not the specific technique used, but the amount of thought that has gone into the process. "The objective of appraisal is to elicit thinking in the superior about the performance of his subordinates, with the specific purpose of planning developmental activities which will improve that performance" (1976, p. 28). Because of this, there are probably just as many effective appraisal programmes, techniques and systems as there are companies using performance appraisal. The most popular of these systems will now be discussed briefly. It was decided to use Eichel and Bender's (1984) categorization of appraisal systems, because this method is quite clear and easily understood. As is illustrated in Figure 6, the different appraisal techniques can be divided into three categories, namely the comparative techniques, absolute techniques and the outcome-oriented techniques.

PERFORMANCE APPRAISAL TECHNIQUES	
Comparative Techniques	Ranking Paired Comparisons Forced Distributions
Absolute Techniques	The Essay Approach Weighted Checklists Critical Incidents Graphic Rating Scales Behaviourally Anchored Rating Scales Behavioural Observation Scale
Outcome-Oriented Techniques	The Direct Index Approach Standards of Performance Management by Objectives

FIGURE 6

Classification of different Performance Appraisal Techniques

Comparative Techniques

In the comparative appraisal technique, the evaluation is made by comparing an employee to other employees on the different dimensions included in the performance appraisal. This comparison is generally made on a global dimension which tries to determine the employee's overall effectiveness to the organization (Wells, 1982). Three comparative measures which are mostly used are ranking, paired comparisons and forced distributions. A discussion of each method follows.

Ranking

Two main types of ranking are used when evaluating performance, namely straight ranking and alternative ranking.

- Straight ranking: This is a very simple procedure and involves a comparison of appraisees. When a performance appraisal takes place, the evaluator is typically asked to consider all of the employees and to identify the best performer, the second best and so on (Flippo, 1980). People are frequently ranked on an informal basis, making this procedure a natural one for most evaluators (Cummings & Schwab, 1973).

- Alternative ranking: This is a more complex variant of straight ranking. According to Cummings and Schwab (1973), an alphabetical list of all employees to be ranked is given

to the evaluator who is asked to think of the best employee in the group on the dimension of interest. Following this, the evaluator must think of the poorest employee on the same dimension. When a person is identified as best or poorest, his name is removed from the list and recorded on a separate ranking sheet. The evaluator can now alternate between thinking of the best and poorest on a list which becomes increasingly shorter (Ivancevich & Glueck, 1983).

Paired Comparisons

This approach makes the ranking method easier and more reliable. The appraiser compares each person being evaluated with one another. This is done by using pairs. The rater chooses for each pair of persons the one who ranks highest in the pair. The number of times each person is preferred is calculated and a ranking of all the individuals is made on the basis of this preference frequency (Cascio, 1982; Cummings and Schwab, 1973; Eichel and Bender, 1984; Graves, 1982; Landy, 1985).

The paired comparison technique with four employees (M through P) is illustrated in Figure 7. The circled letters indicate the evaluator's choice of the better employee in each pair. If the number of times an employee is preferred is tallied, it is found that O would be ranked highest, followed by M, N and P.

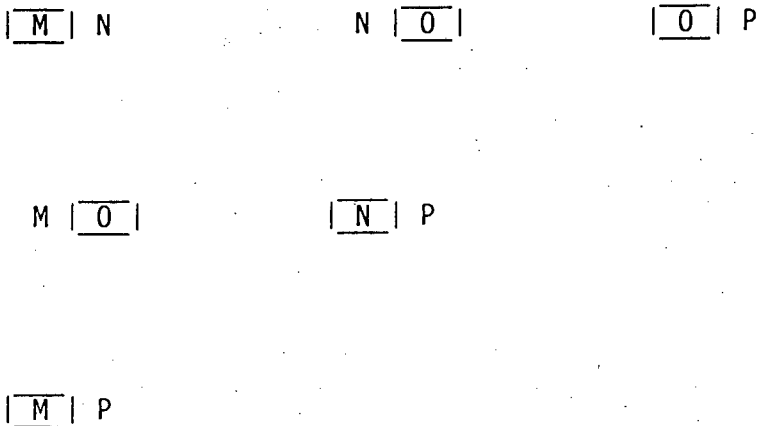


FIGURE 7

The Paired Comparisons Appraisal Technique

The principal advantage of the paired comparisons technique is the fact that the rater's judgement will be visible (Graves, 1982). For example, if a rater says that person M's performance is better than person N's, and N's is better than O. If the rater then adds that O is better than M, the rater is doing more than recording a preference - the rater is also documenting an inconsistency. The practical value of this process is that paired comparisons are far more difficult to manipulate, especially if the number of ratees increases.

The main disadvantage of this technique involves the number of ratees included. As the number of ratees increases, the number of possible pairs also increases very rapidly (Cummings & Schwab, 1973; Graves, 1982). Therefore a vast number of comparisons will have to be made in some instances. According to Graves (1982), the unreliability caused by fatigue and boredom would quickly erase any benefits of using paired comparisons to appraise employees' performance. This fact limits the use of the technique to fairly small units only.

Forced distribution

The forced distribution system is similar to grading on a curve. The rater is asked to place employees in some fixed distribution of categories, such as 10 per cent in low, 20 per

cent in low average, 40 per cent in average, 20 per cent in high average and 10 per cent in high. By doing this, a normal distribution is enforced. By forcing the distribution in this way, interindividual constant errors such as leniency are controlled (Cummings & Schwab, 1973). French (1970) also stated that this method appears to minimize the problems of the halo effect, discussed earlier.

The principal disadvantage of forced distribution, is the fact that the appraisees as a group do not usually conform to whatever distribution is established (Cummings & Schwab, 1973). Therefore, employees are forced into certain categories when appraised, thus not indicating the true worth of such an employee to the organization.

The forced distribution method is illustrated numerically in Figure 8 and the curve which results from this distribution is illustrated graphically in Figure 9.

A newer variation of forced distribution is the point allocation technique (PAT). In PAT, each rater is given a number of points per employee in the group to be evaluated and the total points for all employees evaluated cannot exceed the number of points per employee times the number of employees evaluated (Slivinski, 1975).

The forced distribution and PAT are most likely to be used by superiors, but could also be used by peers or subordinates (Schuler, 1981).

TOTAL NUMBER OF EMPLOYEES	LOW 10%	NEXT 20%	MIDDLE 40%	NEXT 20%	HIGH 10%
40	4	8	16	8	4

FIGURE 8 : Distribution of 40 Employees by Forced Distribution

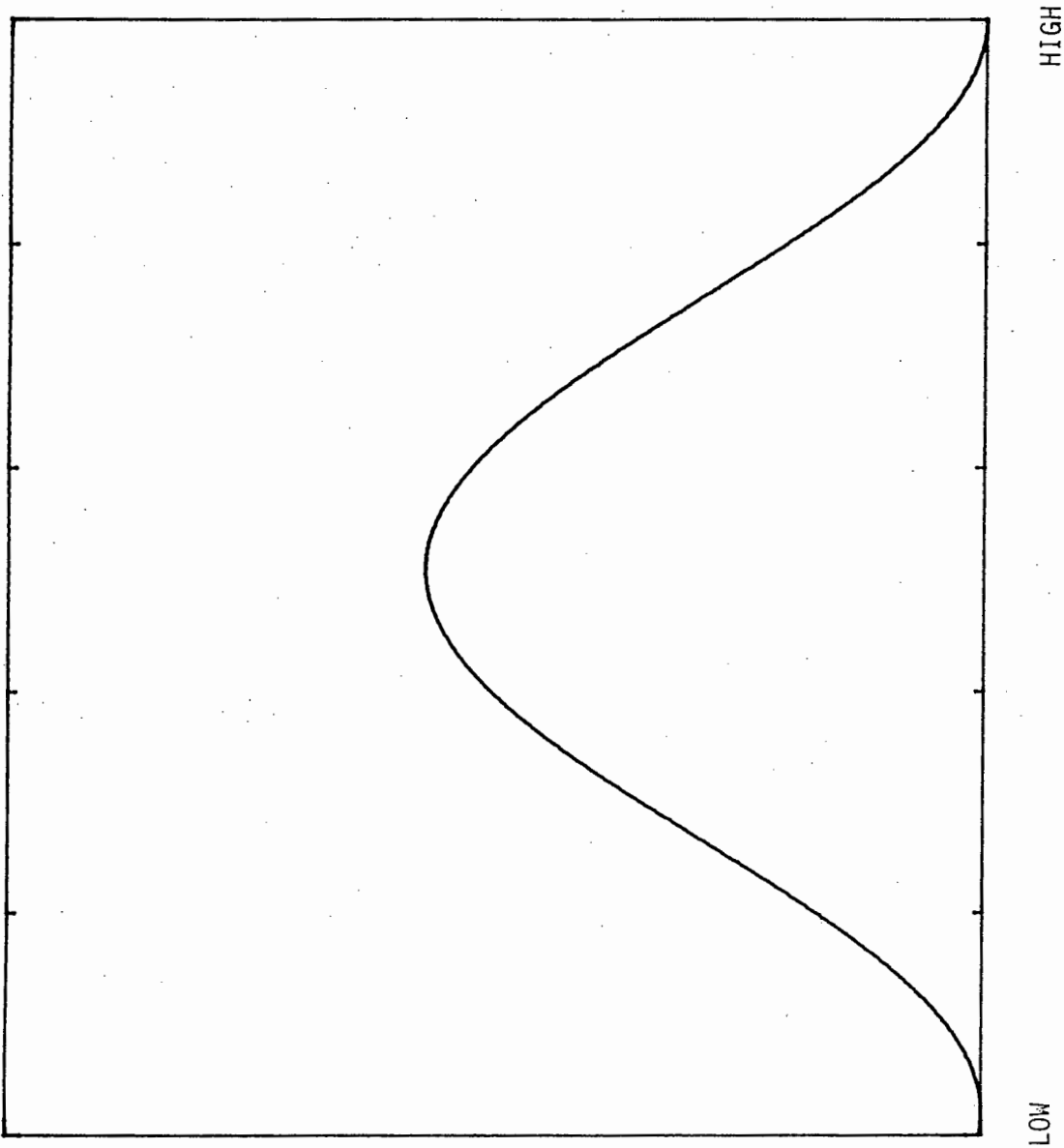


FIGURE 9 : The forced-choice appraisal technique

Evaluation of Comparative Techniques

There are two main advantages to comparative methods, namely simplicity and naturalness. The evaluation of persons by merely ordering them from best to worst in terms of some specific characteristics, is a rather simple process. Furthermore, to rank people is a very natural type of evaluation, because it involves a kind of judgement which is frequently made in everyday living (Eichel & Bender, 1984).

Another advantage is stated by Cummings and Schwab (1973). They maintained that ranking procedures are not subject to interindividual constant errors such as leniency, central tendency and strictness, because the evaluator is forced to array appraisees from high to low.

A major limitation of the ranking systems described, however, is the fact that employees are generally ranked on only one dimension, which is usually some global effective measure (Lopez, 1975). This one dimension usually consists of a personality trait, as discussed earlier. Thus, the trait-oriented approach may be a quite generally accepted method for comparative appraisal techniques. Dunnette (1966) pointed out that treating job success as consisting of only one general characteristic, is usually unrealistic. Problems are created because the definition of overall effectiveness may differ from one evaluator to another.

Another disadvantage is that rank order data do not reveal very much about the intervals between the individuals listed in the ranking (Haynes, 1978). According to Eichel and Bender (1984), comparative methods yield only an ordinal scale that permits rank ordering of individuals without knowing how much difference there is amongst the individuals.

A further problem may be that if a specific group's performance is low, the highest ranking member of the group may be low compared to another group with high performance. The number of employees to be appraised may also influence the ranking process. As the number of employees to be appraised increases, the task of ranking becomes more difficult, or even impossible for the evaluator (Cummings & Schwab, 1973; Eichel & Bender, 1984; Ivancevich & Glueck, 1983).

Absolute Appraisal Techniques

In the case of the absolute appraisal technique, the evaluator appraises employees without comparing them with other employees. Techniques in this category include the essay technique, weighted checklists, critical incidents, graphic rating scales, the behaviourally anchored rating scale, as well as the behavioural observation scale.

Essay evaluation/narrative technique

According to Wells (1982), the narrative technique of performance appraisal is the first technique which can realistically be considered as an appraisal technique, because it is evaluative as well as descriptive. Essentially, this procedure attempts to evaluate performance qualitatively, with an underpinning of factual information to support it. In the essay evaluation, the rater is usually asked to write an essay, describing the strong and weak aspects of an employee's behaviour. The rater writes a narrative description of the individual's performance and uses specific incidents where necessary. The appraisal could also contain such adjectives as "excellent" and "unsatisfactory" (Cascio, 1982; Eichel and Bender, 1984; Slivinski, 1975).

The essay evaluation can be used independently, or it can be combined with other techniques. It is usually combined with the graphic rating scale. In this case the essay summarizes the scale, elaborates on some of the ratings and discusses added dimensions which do not appear on the scale (Graves, 1982). In both cases the essay can be open-ended, or it may be based on guidelines which suggest the topics to be covered. The specific purpose of the essay may also be stated.

Superiors, peers or subordinates of the employee to be evaluated can use this method of appraisal. A typical essay appraisal question might include the following aspects: Describe in your

own words the individual's strengths and weaknesses, potential, overall performance, including quantity and quality of work.

A main disadvantage of the essay evaluation is that the evaluative words could take on a variety of meanings to those who read the appraisal. Confusion about the intended meaning of these words could develop (Slivinski, 1975; Wells, 1982). In addition, it is impossible to compare different employees, since different essays touch on different aspects of ratee performance or personal qualifications, even if the evaluations were written by the same appraiser (Eichel & Bender, 1984). The employee's rating may also depend as much on the writing skills of the evaluator as on the actual performance of the employee (Ivancevich & Glueck, 1983).

Weighted checklist

A checklist is developed by compiling a series of descriptive statements on job-related behaviour. Weights are assigned to the behaviours described on the checklists by a group of persons (job analysts or supervisors of the job to be appraised). The more important behaviours to be revealed by employees are assigned bigger weights than the less important ones. In other words, these weights are based on how favourable or unfavourable these behaviours are for successful performance, as determined by the organization.

The list of descriptive statements of job related behaviour is then given to the rater, who must indicate (by a checkmark, for example) which qualities or characteristics describe the subordinate most appropriately. The weights and the checkmarks are then added up for each subordinate (Eichel & Bender, 1984; Zippo & Miller, 1984), and the employee who received the biggest total of weights is seen to be the best employee.

An example of a weighted checklist is given in Table 4.

Critical incidents

This technique involves keeping a record of unusually good or undesirable incidents occurring in an employee's work, and provides a factual record for subsequent discussions and decision-making. Superiors must indicate which of a number of incidents of performance (successful and unsuccessful) were exhibited by each subordinate. Usually these incidents - regarded as critical by the organization, subordinates and superiors - are grouped into several categories representing separate dimensions of performance. Once the categories are developed and statements of effective and ineffective behaviour are provided, the rater prepares a log for each employee. During the evaluation period, the evaluator records examples of critical behaviours in each of the categories and the log is used to evaluate the employee at the end of the period (French, 1970; Graves, 1982).

TABLE 4

Illustration of a weighted checklist developed for kitchen managers

<u>ITEM</u>	<u>SCALE VALUE</u>
Criticizes employees unnecessarily	0,7
Enjoys contacting customers personally	6,9
Always keeps the kitchen clean	7,1
Does not know which supplies to order	0,1
Is an expert on meat cuts	5,7

Cummings and Schwab (1973) stated that the implementation step of the critical incidents technique involves giving each evaluator a list of the general categories developed. It is then expected of the rater to record any positive or negative incidents that occur pertaining to the general categories. The information received from this procedure may serve as a basis for evaluational as well as developmental actions.

Because the critical incidents technique is not evaluative and describes specific actions and behaviours, Graves (1982) suggested that it could be a powerful support to the performance discussion. The manager or supervisor and the subordinate are then able to discuss the actual behaviours which the employee exhibited when faced with problem situations (Dailey & Madsen, 1980).

According to Wells (1982), the critical incidents technique is just a formalized means of documenting performance expectations and outcomes, rather than an appraisal. This documentation is invaluable when discussing performance, but is inadequate when making and defending comparative decisions.

Robbins (1982) stated that the strength of the critical incident method is that it focuses on behaviours. Additionally, a list of critical incidents as compiled for a specific employee provides a rich set of examples from which the employee can be shown which of his/her behaviours are desirable and which ones call for improvement.

The main disadvantages of this technique are in the first instance the fact that appraisers are required to regularly write down these incidents which may be very burdensome and time-consuming. Secondly, the critical incident technique does not lend itself to quantification, which complicates comparison and ranking of different employees (Cummings & Schwab, 1973; Graves, 1982; Robbins, 1982).

Graphic rating scale

Rating scales are the most commonly used systems in performance appraisals (Siegel, 1980). Individuals are rated on a number of traits or factors. The rater must then judge "how much" of each factor the individual has. Usually performance is judged on a 5-point scale, and the number of factors appearing on this scale could range between five and twenty (Cummings & Schwab, 1973). Factors almost always appearing in rating scales are quality and quantity of work, judgement of the employee, job knowledge and motivation (Kellog, 1965).

Rating scales generally consist of several statements about an employee's behaviour or characteristics. The normal procedure is to establish a continuous or discrete scale for each item. A continuous scale does not have specific points where a rater should indicate the assessment of an employee. The mark could be placed anywhere on the scale, whereas with a discrete scale

there are several designated places where a rater has to indicate the score of an employee.

Figure 10 illustrates a few frequently used types of rating scales. In Figure 10 dimension A is scaled continuously. In the assessment of the appraisee, evaluators can place a checkmark anywhere on the scale to indicate high or low performance. The problem with this type of rating scale is the difficulty in assessing the extent of the difference between two checkmarks. All that can be deduced from this scale, is that one employee is better or worse than another employee, but the degree to which they differ, remains unknown.

Dimension B is indicated on a discrete numerical scale. Dimensions C and D are also scaled discretely, but these scales use adjectives instead of numbers to describe positions on the scale. The evaluator should check one of the points on the scale. Dimension D includes a 12-point scale.

According to Cummings and Schwab (1973) discrete scales generally result in greater interrater agreement and are therefore preferable to continuous scales.

A major drawback of the ranking scale is that descriptive words used may have different meanings for different raters. Factors or categories such as "initiative" and "co-operation" are subject to many interpretations, especially when used in

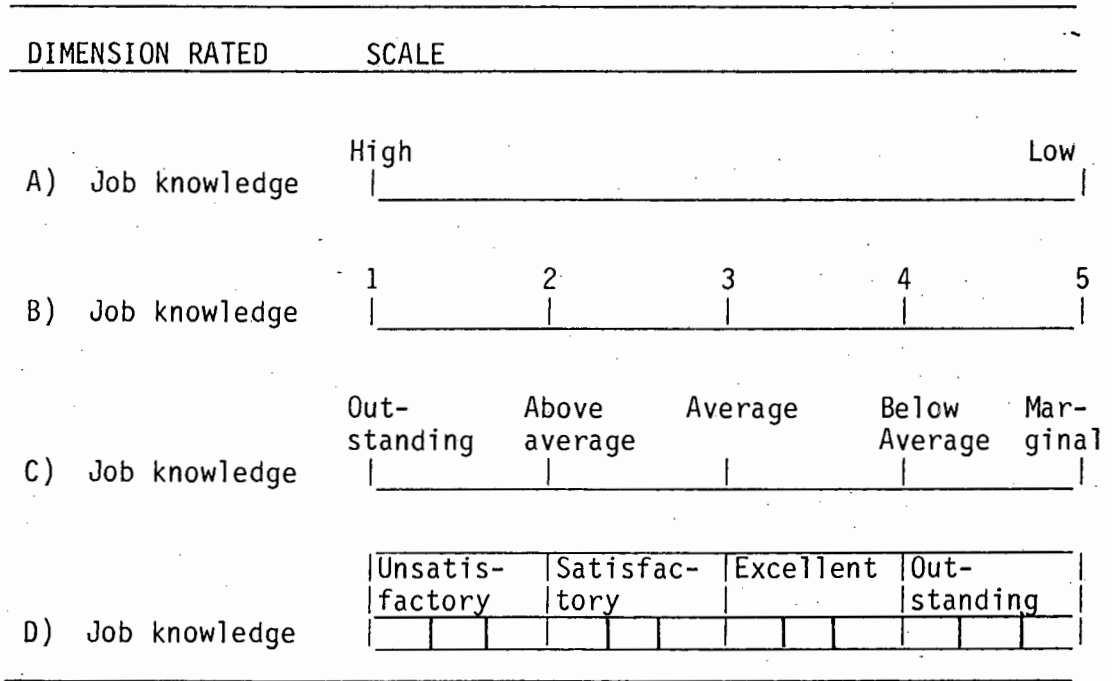


FIGURE 10

Different types of Graphic Rating Scales

conjunction with words such as "average", "outstanding" or "poor" (Mathis & Jackson, 1982).

According to Latham and Wexley (1981) the graphic rating scale is inferior to other scales if raters have been trained. Rater problems which occur frequently with graphic rating scales are leniency, harshness or central tendency.

Behaviourally anchored rating scale (BARS)

A behaviourally anchored rating scale in effect combines major elements from the critical incident- and graphic rating scale appraisal techniques. The main, although not the only difference between BARS and the graphic rating scale, is that rating points appearing on the BARS indicating good or poor performance have behavioural descriptions attached to them (Cummings & Schwab, 1973; Robbins, 1982). Steps in the development of a BARS are as follows: Firstly, persons with knowledge of a specific job or group of similar jobs (e.g. jobholders or supervisors), are requested to describe specific incidents critical to successful performance in these jobs. These incidents are then grouped into from five to ten general performance dimensions which are defined. One dimension may be quality of work and it may be defined as the factor which measures the quality of work done by the employee. Thoroughness is a prerequisite for quality.

The clearly defined dimensions and the critical incidents are then given to a second group of persons, as familiar with

the job as the previous group. These people must then assign each incident to the dimension it best fits. This process is called retranslation (Eichel & Bender, 1984). If the majority of both groups agree upon the dimension assignment, the incident is retained in the design.

A third group of individuals must then rate the behaviour described in each incident in terms of effectiveness or ineffectiveness on the appropriate dimension (Landy & Farr, 1983; Latham & Wexley, 1981). This is typically done by using 7- or 9-point scales. On a 7-point scale a 7 usually represents outstanding job performance, a 4 average and a 1 poor performance. Average effectiveness ratings for each incident are then determined and a standard deviation criterion (typically $SD \leq 1,5$ for a 7-point scale) is set for determining which incidents will be included on the final anchored scales.

A subset of the criteria meeting both the retranslation and standard deviation criteria are used as behavioural anchors for the final performance dimensions, hence the term behaviourally anchored rating scales (Eichel & Bender, 1984; Latham & Wexley, 1981). A final BARS instrument is typically comprised of a series of vertical scales (one scale perdimension) which are anchored by the included incidents.

Regarding the scale, Richardt (1976) stated that the number of points on the scale is arbitrary. Too few points to check

obviously result in too little information on differences between subordinates, and too many discrete points require finer discriminations than the evaluator can reliably make. Richardt further suggested that seven to nine points prove to be adequate for most purposes. It also appears that inexperienced evaluators have greater difficulty with more discrete points than do experienced evaluators.

Jenkins and Taber (1977) further stated that there is little utility in adding scale points beyond five. According to Guion (1965), the number of divisions in scales may vary widely, but that it is usually an uneven number so that "average" will have a central position in the scale. However, since more discrimination is often needed at the "above average" levels, an even number of divisions may be used so that the average spot is somewhat off-centre.

An example of one dimension in a BARS is shown in Figure 11. In this figure, the dimension to be measured is quality of work. A 7-point scale is attached to the dimension with the behavioural anchors indicated at points 1, 3, 5 and 7. The anchors describe behaviour which could be expected from employees doing their jobs. A rater must indicate on the 7-point scale how a specific employee could be expected to perform on a specific dimension.

A main advantage of the BARS is that descriptions of behaviour appearing on the scale are expressed in the rater's own terminology,

Dimension : Quality of Work

This factor measures the quality of work done by the employee.

Thoroughness is a prerequisite for quality.

-
- 7 All the employee's tasks excel the quality demands and his/her ideas and techniques add a positive contribution to the development and improvement of quality
 - 6
 - 5 Some of the tasks are better than the quality requirements and the ideas and techniques of the employee sometimes lead to the improvement of quality
 - 4
 - 3 Most of the employee's tasks meet the quality requirements. It is seldom necessary to repeat a task
 - 2
 - 1 Most of the employee's tasks do not meet the required quality. Some tasks have to be repeated

FIGURE 11

One Dimension on a Behaviourally Anchored Rating Scale

because raters have an important role in the construction of the BARS. This fact eliminates much of the ambiguity which is usually found in rating scales based on traits (Eichel & Bender, 1984; Miles & Snow, 1984).

Another advantage of BARS is that the scales lend themselves to employee counselling by providing the employee with specific feedback on areas in need of improvement as well as on strengths (Latham & Wexley, 1981).

Behaviourally anchored rating scales, however, also have several limitations. According to Graves (1982) the most obvious disadvantage of the BARS is the time and resources needed to develop meaningful behavioural anchors. Too many key people and too much time are used when developing BARS (Eichel & Bender, 1984).

Schwab, Heneman and De Cotiis (1975) stated that a substantial number of critical incidents generated in the job analysis are discarded when developing BARS. This could result in a loss of important information. Borman (1979) indicated that raters frequently are unable to match observed behaviour directly with the scale anchors. According to Borman, raters have difficulty in discerning any behavioural similarity between a ratee's general performance and the highly specific behavioural examples used to anchor the scales.

Another limitation of BARS as identified by Eichel and Bender (1984) is that the scales must be updated and validated continually to ensure that the behaviours specified are still relevant to the job.

Latham & Wexley (1981) described a further disadvantage of BARS when stating that the endorsement of an incident above the neutral point of BARS implies endorsement of all the other incidents appearing between the incident checked and the neutral point.

The Behavioural Observation Scale (BOS)

The behavioural observation performance appraisal technique focuses primarily on the observable behaviour of employees. Therefore the BOS actually consists of several behavioural statements which are critical to the success of a specific job. These statements (earlier described as critical incidents) are then grouped together to form a behavioural dimension. A BOS can thus consist of a few dimensions, where each dimension includes a list of critical incidents regarding this dimension. An example of one of these behavioural dimensions appears in Figure 12.

A supervisor or rater must then rate employees on each of these critical incidents appearing on the scale. By comparing the different scores received by employees, a distinction can be made between the effective and the not so effective employees.

A. PLANNING/FORECASTING

1. Operates on a crisis basis

Almost always 1 2 3 4 5 Almost never

2. Sets goals that are difficult, but attainable

Almost always 1 2 3 4 5 Almost never

3. Establishes a realistic timetable to get the job done

Almost always 1 2 3 4 5 Almost never

4. Planning/Forecasting is based on investigation of facts

Almost always 1 2 3 4 5 Almost never

FIGURE 12

Example of one BOS Performance Dimension for Evaluating Managers

(Adapted from Latham & Wexley, 1981, p. 56).

The specific steps for developing a BOS as described by Birkenbach (1984) as well as Latham and Wexley (1981), are discussed below.

In the first instance, critical incidents are collected from specified job incumbents. The critical incidents technique implies the following: Individuals who are aware of the aims and objectives of a given job, who frequently observe employees performing their functions and who are capable of determining whether the job requirements are being performed satisfactorily, are interviewed about the most important functions of the job. In the case of the present study, the transport manager, the transport foremen as well as ten truck drivers were involved in this process.

Statements about the critical aspects (functions) of the specific job at hand are received from the people interviewed and then categorized. This implies that incidents, essentially describing the same behaviour, are grouped together into one cluster. In other words, similar if not identical critical incidents are grouped together to form one behavioural item. Several items, also similar in context, are then grouped together to form a behavioural dimension. The distinction between critical incidents, behavioural items and behavioural dimensions are illustrated in Figure 13.

The importance of rater participation regarding the abovementioned processes has already been discussed in a previous chapter but

CRITICAL INCIDENTS:

1. Drive the truck to customers
2. Drive the truck in a responsible manner

BEHAVIOURAL ITEM:

Ensure safe driving of truck

BEHAVIOURAL DIMENSION:

SAFETY

FIGURE 13

The Difference between Critical Incidents, Behavioural Items
and Behavioural Dimensions

it may be useful to stress this fact again. Rater participation ensures that a representative sample of critical job behaviours are included on the appraisal instrument. It also means that behavioural descriptions included on the BOS are written in a form which is clear and unambiguous to the raters (Latham & Wexley, 1981).

To determine if another individual or group of individuals would have developed the same behavioural dimensions from the critical incidents obtained in the job analysis, interjudge agreement is assessed. This is achieved by placing the critical incidents in random order and giving them to a second individual or group to reclassify these incidents into the behavioural dimensions specified by the first group of people. According to Latham and Wexley (1981) the ratio derived from the assessment of interjudge agreement should be ,80 or higher for a behavioural dimension to be acceptable for inclusion on the scale. Should the ratio be below ,80, the items included under this dimension should be re-examined to establish whether they should be reclassified under another dimension, or whether they should be rewritten to increase specificity.

At this stage the relevance or content validity of the BOS criteria are examined. Anastasi (1982) stated that relevance or content validity is concerned with the systematic evaluation of appraisal instruments. This evaluation must be done by people

who are familiar with the job (Baker & Morgan, 1984). According to Latham and Wexley (1981) the content validity of the BOS can be determined by putting ten per cent of the critical incidents derived from the sample aside before categorizing these items. After the categorization of the remaining 90 per cent of critical incidents is done, the other ten per cent is examined to see if any of the incidents describe behaviours that have not yet appeared. If this is the case, it is assumed that the number of incidents which had been collected was not sufficient for the appraisal instrument to possess content validity.

The last step in the development of the appraisal instrument is to attach a 5-point Likert scale to each behavioural item (Latham & Wexley, 1981). Jenkins and Taber (1977) indicated that there is little utility in adding scale values beyond five. Observers (raters) now have to indicate the frequency with which they have observed a job incumbent engage in each behaviour. A total score for each employee is then determined by summing the observer's responses to all the behavioural items. Employees can then be compared on various dimensions by just comparing the scores (or the total scores) they received on these behaviours. They also could be ranked from good to poor depending upon their total scores.

Percentages can also be added to the scale values. For example, employees receive a score of 1 on a 5-point scale if they have been observed engaging in a behaviour 0 - 64 per cent

of the time. A 2 is scored when employees engage in a behaviour 65-74 per cent of the time, a 3 for 75 - 84, a 4 for 85 - 94 per cent of the time and a 5 for 95 to 100 per cent of the time (Latham & Wexley, 1981). These percentages corresponding to the 5 points on a Likert Scale may differ depending on the job and the organization involved.

At this stage an analysis should be conducted to determine whether some items included on the scale should be eliminated because of their inability to discriminate between good and poor performers (Latham & Wexley, 1981). The actual statistical procedure involves correlating the scores on each behavioural item with the total of all items appearing on the scale. A decision to remove some of the items may rest on these correlations.

According to Latham and Wexley (1981), if the total of individuals to be rated is approximately three to five times more than the items included on the performance scale, a factor analysis may be conducted. In the case of this study, there were only two times more individuals than items, therefore a factor analysis was considered inappropriate.

The organization using this appraisal system may group the different total scores into separate classes if it is deemed necessary. This implies that a descriptive label can be attached to employees receiving a specific score. Classes usually have labels such as "poor", "average" or "excellent". Thus, depending

on the classification of the scores by an organization, a score of 20 may indicate a "poor" employee, a score between 20 and 40 an "average" employee, and a score between 40 and 60 an "excellent" employee.

There are various inherent advantages of the BOS as an appraisal technique. In the first place, it is developed by a systematic job analysis supplied to a large extent by employees for employees. This means that understanding of and commitment to the use of the appraisal instrument are greatly facilitated (Latham & Wexley, 1981) (The importance of this aspect has already been discussed in Chapter 2). Both managers and subordinates can thus understand the criteria on the appraisal form because all of those in a specific job were involved in developing the appraisal instrument.

Secondly, the BOS states explicitly what behaviours are required from an employee in a given job. Therefore, it can either serve alone or as a supplement to existing job descriptions. When used as a job description, the BOS can be used as a job preview for potential employees by indicating to them what will be expected from them in the job (Latham & Wexley, 1981). Wanous (1973) stated that realistic job previews are an effective means of reducing employee turnover and job dissatisfaction.

In the third instance, the BOS also facilitates explicit performance feedback, because it encourages meaningful discussions between the supervisor and the employee about the employee's

strengths and weaknesses. General remarks about the employee's work are avoided in favour of specific behaviours for which the employee is praised or encouraged to demonstrate on the job (Latham & Wexley, 1981). Regarding this aspect of performance appraisal, Dosset, Latham and Mitchell (1979) indicated that explicit performance feedback using BOS combined with the setting of specific goals, is an effective means for bringing about or maintaining a positive behaviour change. In their study, they used 60 female clerical personnel. Some of these females were assigned specific goals, while others were just told to do their best on certain tasks. It was found that specific goals led to higher performance than did the "do your best" goals.

According to Kane and Bernardin (1982) the BOS method of performance appraisal includes a "crucial flaw" (Kane & Bernardin, 1982, p. 636). On the BOS, a given occurrence rate interval does not connote a constant level of performance satisfaction for all job behaviours. An example of a policeman's job was used to explain this problem. A policeman who arrests people and receives a good score for this on the scale, namely a 4 or 5, may receive an inappropriate score, because the BOS does not indicate **how** the arrests were made, and whether undue force was used to make the arrests. According to them, the policeman could appear before court as a result of the force used in making the arrests, but he would still receive a high score on the BOS because of the number of arrests he made.

This author does not consider this aspect to be a major problem because a rater should be trained to take everything into account when making a rating. Even an untrained rater should be able to distinguish between appropriate and inappropriate behaviour and to rate accordingly. However, it must be admitted that this problem could be valid where no controls or checks on an employee's job are available. In this instance, if an employee seems to engage in the desired behaviour with a high frequency and it is not possible to determine **how** the work is being done, a problem may arise. It is of course important to note that the BOS is a behavioural appraisal scale where the behaviour should be observable. This implies that behaviour which is not readily observable, or cannot be checked or controlled easily, should not be included on the scale.

Outcome Oriented Techniques

This type of appraisal technique concentrates mainly on the specific accomplishments or results achieved by the employee. Examples include the direct index technique, the standards-of-performance technique and the management by objectives technique.

Direct index approach

With this technique, individuals are evaluated solely on the basis of the results they have achieved in their jobs. The direct index measures subordinate performance by objective, impersonal

criteria such as productivity, absenteeism and turnover (Schuler, 1981).

The actual collection of data is determined largely by the specific objective to be measured. An example of this method could be in the evaluation of a manager's performance, counting the number of employees terminating their services, or by looking at the absenteeism figures for the employees (Slivinski, 1975).

The direct index technique can be divided into measures of quality as well as measures of quantity. Quality measures may include, for example, customer complaints, scrap rates and the number of defective units or parts produced, whereas quantity measures may include units of output per hour, new customer orders and sales volumes (Schuler, 1981).

Cummings and Schwab (1973) stated that an advantage of the direct index appraisal technique is that it avoids many of the rater problems usually present in performance appraisal systems. This is true, because direct measures avoid the need to have an appraiser do a subjective evaluation of an employee's performance, especially if the procedure to obtain the direct measure is well defined (Eichel & Bender, 1984). It is easier to determine how frequently an employee was absent from the job, than to determine the amount of initiative revealed by that same employee.

A major drawback of the direct index approach is that direct measures are obtainable only when the employee produces a distinguishable

output, for example, the quantity of crates loaded on a truck per hour.

Standards of Performance

The performance standards approach compares actual accomplishments with a detailed set of expectations (Levinson, 1976).

The organization determines daily or weekly work standards in advance. Each employee is given a more or less complete set of job duties and the supervisor will eventually relate the performance appraisal and feedback interview to the attainment or non-attainment of these work standards (Slivinski, 1975).

The principle disadvantage of the performance standards method is the amount of time required to define job priorities and work standards for each job (Eichel & Bender, 1984).

An advantage of the performance standards approach is the fact that the approach is participative, because the subordinate defines the work standards in conjunction with the superior (Eichel & Bender, 1984). Feedback and appraisal interviews are also more objective because they are based on specified outcomes in the principal job segments, rather than on personal attributes.

Management by Objectives (MBO)

Management by objectives can be described as a "process whereby the superior and subordinate managers of an organization

jointly identify its common goals, define each individual's major areas of responsibility in terms of the results expected and use these measures as guidelines for operating the unit and assessing the contribution of each of its members" (Odiorne, 1979, p. 52).

In other words, it relates organizational goals to individual goals. This helps to increase an employee's understanding of where the organization is at the moment and where it is heading (Burack & Smith, 1977).

To implement an MBO appraisal technique in an organization, four developmental steps are required. The first step is to establish the goals each subordinate is to attain. Superiors and subordinates can work together to establish goals for the subordinates. The goals can refer to desired outcomes to be achieved, means (activities) for achieving the outcomes, or both (Odiorne, 1979).

The second step involves the subordinate's performance within a previously arranged time period. As subordinates perform, they know fairly well what there is to do, what has been done and what remains to be done (Odiorne, 1979).

The third step is a comparison of the actual level of goal attainment against the agreed-upon goals. The evaluator explores reasons for goals not being met or for goals being exceeded. This step helps to determine possible training needs. It also alerts the superior to conditions within the organization which

may affect a subordinate's performance and over which the subordinate may have no control (Eichel & Bender, 1984; Graves, 1982; Odiorne, 1979).

The final step is to decide on new goals and possibly new strategies for goals not previously attained. At this point, subordinate and superior involvement in the goal setting process may change. Subordinates who successfully reached the established goals may be allowed more participation in the goal-setting process in the future (Odiorne, 1979; Schuler, 1981).

Some problems regarding MBO have been identified by various authors. In the first place, Cummings and Schwab (1973) stated that MBO was initiated in organizations primarily as a developmental tool rather than an evaluative one. Graves (1982) supported this view by stating that MBO is not an appraisal system because it is not evaluative. He further stated that it is contrary to the policy of MBO for individuals to be compared or evaluated. The accomplishment of a stated objective is nothing more than just an accomplished objective. Before this accomplishment can be useful for decision-making by management, the accomplishment must be evaluated or appraised. The problem lies in comparing a challenging objective almost attained to the overattainment of an easy objective. The resolution of such a situation is a judgemental one, and these judgements must always follow MBO to provide managers with the information they need for making decisions. MBO provides none of this necessary information, and this appears to be its major weakness as an appraisal method.

Furthermore, Ivancevich and Glueck (1983) also indicated some deficiencies regarding MBO as a performance appraisal technique. According to them, it includes too much paperwork, which results in a substantial loss of working time. This system also lends itself to malpractice by the users thereof. In many instances too many goals are usually set and this results in confusion between employee and supervisor with regard to goals. MBO is sometimes also forced into jobs where it is extremely difficult to establish objectives. At times, too much emphasis is placed on short term planning. Original goals are also seldom modified (Lawler, Mohrman & Resnick, 1984).

In conclusion and on the basis of the above review, it appeared that the **behavioural observation scale would meet the requirements of this study**. The arguments for this choice are discussed below.

Motivation for selecting the Behavioural Observation

Scale as Performance Appraisal Technique

Apart from the advantages mentioned earlier in this chapter, some other reasons contributed to the decision to use the behavioural observation scale as a performance appraisal system for truck drivers in this particular organization.

In some instances, the success of an employee in a job depends almost entirely on the behaviour demonstrated by the employee while carrying out his work. Therefore, to rate such an employee on personality traits would leave one none the wiser regarding his/her real performance. This was particularly relevant in

the case of the truck drivers in this wine manufacturing organization. The job was relatively straightforward and did not demand a great deal of thinking and reasoning from any driver. That which was required was of a fairly basic nature. The most important aspect of the job of truck driver, was the behaviour demonstrated in doing the job. The driver had to physically drive the truck to customers, supervise and help with loading and unloading of cargo, keep the truck in running order and liaise with clients (A more detailed job description is given in Chapter 4). All of these duties were performed by engaging in some type of behaviour or another.

A prerequisite for a BOS is that behaviour should be observable. This was the case in this instance. The truck driver's supervisor was in constant contact with him and even when the driver and his truck had left the depot, radio contact was maintained. Reports on the behaviour of the driver were received from customers and the driving habits of the driver were recorded by a tachometer installed in each truck.

An added prerequisite was that the performance appraisal system to be developed should not be too complex to understand, because of the limited educational levels of both drivers and their supervisors. The BOS caters for this need because the involvement of both drivers and supervisors leads to a more understandable system.

Another advantage to using BOS lies in the fact that it includes more items than is usual for other appraisal systems.

Thus, a wider scope of behaviour may be assessed. Furthermore, the system is not too difficult to develop and does not require a large amount of time to complete. As stated earlier, it also provides means for feedback to employees and for setting future goals to be attained.

In the previous chapter the criteria for good performance appraisal systems as described by Banks and Roberson (1985) were reported. They compared the various techniques reviewed above in terms of these criteria. The results obtained in their study are indicated in Table 5. In Table 5, a Y indicates a positive answer and the N a negative response on a specific criterion. A dash (-) indicates that that specific criterion could not be determined or that it did not apply. When judged on these criteria, it becomes clear from this comparison that the BOS appears to be the superior appraisal instrument.

In conclusion, the abovementioned review confirms that a BOS would be the most suitable performance appraisal technique to be used in this study. It is easily developed, it measures what should be measured and it provides a means for feedback to employees on their performance. It also describes what is needed for improved performance.

Thus, although a behavioural approach to performance appraisal is time consuming to develop as discussed earlier, it was decided that the benefits to be derived from such a system would justify the time spent in the development thereof.

TABLE 5

Criteria for good appraisal systems included in different appraisal formats

CRITERIA OF GOOD TESTS	RANKING	PAIRED COMPARISON	FORCED DISTRIBU- TION	ESSAY	WEIGHTED CHECKLIST	CRITICAL INCIDENTS	GRAPHIC RATING SCALE	BARS	BOS
A. DOMAIN CLARITY AND ITEM DEVELOPMENT									
Representative sampling of behaviour domain	N	N	N	N	Y	Y	-	Y	Y
Multiple Items	N	N	N	N	Y	Y	N	N	Y
Construct well defined	-	-	-	N	Y	Y	-	Y	Y
Items developed from input of several experts	N	N	N	N	Y	Y	-	Y	Y
B. ITEM SELECTION AND ANALYSIS									
Interitem correlation	N	N	N	N	-	-	N	N	Y
Item discriminability	N	N	N	N	-	-	N	-	Y
Item validity	-	-	-	N	-	-	-	-	-
C. ITEM SCORING AND INTERPRETATION									
Standardized scoring procedure	Y	N	Y	N	Y	Y	Y	Y	Y
Uniformity of test procedure	N	N	N	N	Y	Y	Y	Y	Y

(Adapted from Banks & Roberson, 1985, p. 134).

CHAPTER 4

Methodology

Although the problem to be investigated in this study has been mentioned on several occasions so far, it seems appropriate at this stage to restate the problem as it existed in this particular wine manufacturing company. On the basis of this analysis, some decisions can be taken for solving it. These include the choice of a sample and the steps to be taken in developing a performance appraisal system for the truck drivers involved.

Definition of the Problem

As was mentioned in Chapter 1, the Transport Department of this wine manufacturing company did not view their performance appraisal system as important for various reasons. This implies that prior to the execution of this study, no reliable system existed to appraise the performance of the truck drivers. Therefore, no formal means existed to discriminate between good and poor drivers and to determine who should be promoted or who should receive a salary increase. No methods were available to determine which aspects of a driver's performance should be developed by means of training or other appropriate steps. In addition, the drivers did not know how to improve their own performance. They could not set developmental and career goals for themselves, which had a rather negative influence on worker morale.

In the previous chapter, it was mentioned that the appraisal system to be developed should be easily understandable to the drivers, as some of them did not have a very high degree of formal education and some of them also had communication problems because of English and Afrikaans being their second or third languages.

The problem can thus be summarized as follows: A performance appraisal system had to be developed for the truck drivers in the Transport Department. This system had to be easily understandable to everyone using it; it had to provide a means for improving the performance of truck drivers; and it also had to be able to identify training needs. Furthermore, the system had to provide a means for giving employees feedback on their performance as well as to counsel employees on problems regarding their work.

The procedure followed in the execution of this study will now be discussed.

Procedure

During the course of the study, four main problems had to be addressed. In the first place, a preliminary study had to be executed to determine the need (if any) for a new performance appraisal system for truck drivers. In other words, the satisfaction of the drivers with the previous performance appraisal system, as well as their expectations of a new system, were determined.

Secondly, after the need had been ascertained, a performance appraisal instrument had to be developed. This was done after comparing several appraisal techniques. It was mentioned previously that one of the primary purposes of the new performance appraisal system was to improve the performance of the truck drivers after implementation of the system. Thus, the third problem to be addressed, was to determine whether the performance of the drivers involved in this study actually did improve.

In the fourth instance, the satisfaction of the drivers with the old and new appraisal systems was compared to determine if the drivers were more satisfied with the new system than with the previous one. At the same time, drivers again had to indicate their expectations regarding performance appraisal, thus ensuring that areas still needing improvement, were identified.

These four problems will now be discussed in more detail. For the sake of greater clarity, they will be discussed as separate entities, although information gathered during the procedure could influence all three problems simultaneously. For example, the reliability of the performance appraisal system developed could only be determined after the appraisal scores of the drivers had been obtained (the third problem) but it will be discussed when describing the development of the appraisal instrument (the second problem).

During the study, two main groups of people were involved, namely the truck drivers and their direct supervisors, the transport foremen. Various aspects regarding these groups will now be discussed, before addressing the problems mentioned above.

The Sample

The sample consisted of 80 truck drivers and six transport foremen at the same wine producing company. All were employed in the Transport Department of the company. The sample was divided into two groups, each consisting of 40 truck drivers and three transport foremen. One group situated in the Eastern Cape served as a control group and the other group situated in the Western Cape was the experimental group.

The organizational position of the truck drivers and the supervisors in the Transport Department is illustrated in Figure 14.

To determine whether a direct comparison of the control with experimental groups was possible, some analysis of the functioning of these two Transport Departments was required. Job descriptions were obtained for both groups, but no difference between the two groups could be found in these descriptions. The combined job description is included later in this chapter.

The main difference between the experimental and control groups was their geographical setting. This, however, could make a positive contribution to the study in that the control group could not be influenced or contaminated by the activities of the experimental group.

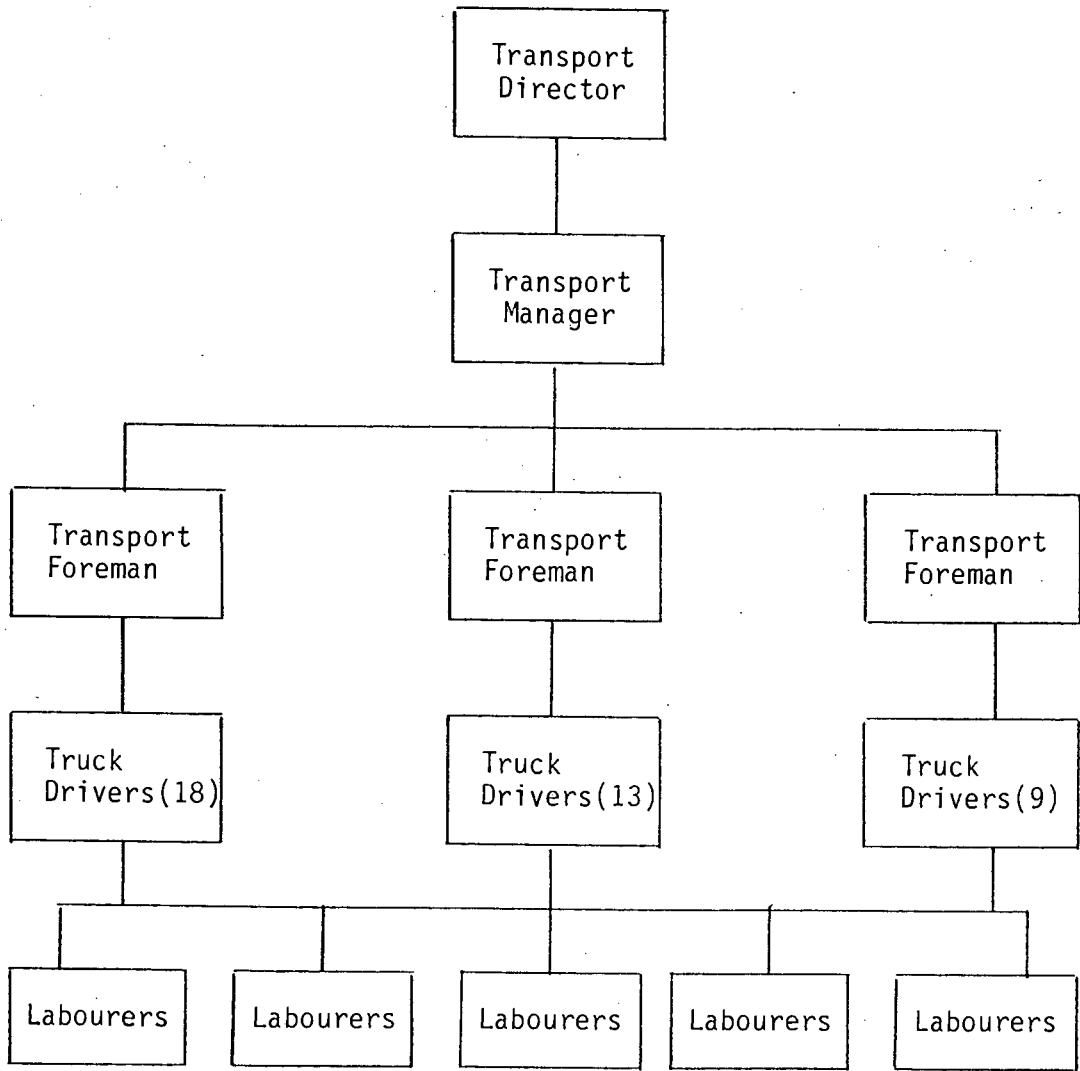


FIGURE 14

The Transport Department

The drivers in the experimental group had an average service period of 10 years, an average age of 32 years and an average educational level of standard 6. Those in the control group had an average service period of seven years, an average age of 27 years and an average educational level of standard 7. The fact that the experimental group had a longer service period, a lower educational level and an older age, may be attributed to the fact that the branch of the company situated in the Western Cape is much older than the one in the Eastern Cape. Therefore, it was possible for drivers to be employed longer in the Western Cape than in the Eastern Cape.

A more detailed description of the truck drivers and transport foremen follows.

Truck Drivers

The truck drivers in the experimental group were all coloured males, whilst 27 of the drivers in the control group were blacks and 13 coloureds. All of them were appointed in job grade 7 (the company used the Castellion job grading system where employees are graded from 1 to 16. A grade 1 usually describes the general labourer, whilst grade 16 indicates high management level).

Each of these truck drivers had to supervise their own working team which usually consisted of four, but never more than five general labourers. These labourers were appointed in job grade

1. The labourers assisted the truck drivers with the loading and unloading of cargo. Each driver was directly responsible for the general appearance and conduct of his team.

A specific truck was assigned to each driver by the company and the maintenance and general appearance of this truck was the sole responsibility of the driver.

To be appointed by the company as truck driver, the applicant had to meet certain general requirements. In the first place, he had to possess the prescribed licence for the driving of heavy vehicles. Secondly, the applicant also had to be physically fit, because the driving of these heavy vehicles for hours required strength, stamina and endurance. In the third instance, a driver also had to pass a driving test initiated by the company.

Before being appointed permanently, each driver was appointed for a two-month probational period. During this period, their general abilities as well as driving skills were monitored and if they were found to meet the requirements set by the company, they were appointed permanently.

Transport Foremen

Five of the respondents were coloureds, while one was an Asian. All of them were truck drivers before being promoted to transport foremen. They all had at least a standard 10 school qualification (12 years of schooling) and their average years

of service in the company was 13 years. They were all appointed in job grade 8.

Each transport foreman had at least 14 truck drivers, but never more than 18 under his direct supervision. The transport foremen had almost direct control over each aspect of the work done by the truck drivers, as the drivers had to report all their activities directly to the foremen. All complaints about truck drivers (e.g. by customers) were also to be dealt with by the foremen.

Preliminary Study : Determining the need for a new
Performance Appraisal System

As indicated earlier in this chapter, a preliminary study was done to assess the need for a new performance appraisal system for truck drivers in the Transport Department. In doing this, the attitude of the drivers towards the existing performance appraisal system, as well as their expectations of such a system, had to be assessed. A questionnaire developed by the company involved in this study in conjunction with an insurance company for the same purpose, the Performance Appraisal Satisfaction Questionnaire (Van Wyk, 1983), was used. The questionnaire consisted of 17 questions and the people completing it had to indicate for each question what their experiences and expectations regarding

performance appraisal on that specific question were. An example of the questionnaire appears in Appendix 2.

The questionnaires, in both official languages, were handed out to be completed by each truck driver in the experimental group. No questionnaires were handed out to drivers in the control group, to ensure that their performance and views regarding performance appraisal were not influenced in any way. This was believed to contribute to the reliability of the appraisal scores of these drivers derived from the newly developed performance appraisal system.

The instructions on how to complete the questionnaire appeared on the questionnaire itself and the supervisors (transport foremen) were available to help with the completion of the questionnaires if any problems were encountered. It was pointed out to the truck drivers that the information wanted from them would be treated confidentially and that they did not have to identify themselves when completing the questionnaire.

The purpose of this measurement was to determine the discrepancy between what truck drivers really experienced from the existing performance appraisal system and what they would expect from a new system. A t-test for dependant samples was used to determine the significance of this discrepancy (if any). The same measurement

was used to compare the satisfaction of the drivers with the "old" and "new" performance appraisal systems, to be discussed later in chapter 4.

In analysing the data, multiple t-tests were not used, because when there are many samples and hence more than one comparison, the sampling distribution of t is no longer appropriate (Pagano, 1981). It would also increase the actual probability of making a Type I error. According to Campbell (1966) and Miller (1975) the t-test is a powerful technique for comparing related samples and it is likely to detect significance if present in the data.

Miller (1975) stated that the assumptions underlying a t-test are that the difference scores (one for each pair of observations) may be regarded as a random sample of differences from a normal population, and that the measurements are on an interval scale. Provided that the sample of differences obtained does not suggest glaring deviations from normality in the population, the t-test may be used.

Pagano (1981) also stated that the t-test for dependent groups allows utilization of both the magnitude and direction of the difference scores. According to Pagano, it essentially treats the difference scores as raw scores and tests the assumption that the difference scores are a random sample from a population of difference scores having a mean of zero.

However, Cronbach and Furby (1970) and Johns (1981) are of the opinion that by comparing difference scores, unreliable results will be obtained. According to Johns (1981) the biggest problems with difference scores include potential unreliability, systematic correlation with their components and spurious correlations with other variables.

Another point of criticism is that in many instances measuring instruments are constructed ad hoc for use in a particular study. The reliability of these instruments is virtually never reported. Since these instruments have no published history of reliability, one may be unsure about the validity of the measurement. Johns (1981) further stated that the most likely problems occur when difference scores are correlated with each other, when they are correlated with non-difference measures in some multivariate manner, or when scores derived from various sources are used in a single study (A possible solution to the abovementioned problems may be to correct for attenuation before the analysis).

Not one of the four abovementioned problems applied to the study at hand. Although the measuring instrument used in this study, the Performance Appraisal Satisfaction Questionnaire, had no published history of reliability, reliability was determined for both the expectations and experiences of the truck drivers

in this study and it was found to be satisfactory (the reliability of this instrument is discussed in chapter 5).

The purpose of this preliminary study was only to determine whether a need for a new performance appraisal system existed in the company. Therefore a test of significance had to be performed between the expectations and experiences of the drivers. According to Cronbach and Furby (1970) a significance test in a one-group design need only ask whether μ_y is reliably different from μ_x . This difference in sample means for x and y is the best available estimate of the mean difference.

Construction of the Behavioural Observation Scale

In order to develop as comprehensive a performance appraisal system as possible, it is essential to recognize the positive aspects of other systems. Even though the behavioural observation scale, a behaviour oriented approach, was used, some borrowing took place from the trait and results approaches as discussed in the previous chapter. In this way, an assessment of the total person was obtained. Slivinski (1975) supported this idea when he stated that "... the trend today in assessment is towards a more qualitative and analytical assessment of the total person

to determine his/her long range potential for positions of higher responsibility" (1975, p. 6).

The development of a BOS was discussed in Chapter 3 and in the case of this study the advice given was closely followed. The BOS used in this study was developed as follows.

Development of the Performance Appraisal Instrument

A job analysis of the job of truck driver was done according to the critical incident technique. The six transport foremen, ten truck drivers as well as the transport managers of both Transport Departments were all involved in writing down incidents which were viewed as critical to the job. (A list of these critical incidents appears in Appendix 3). An industrial psychologist employed by the company, in conjunction with the transport foremen, used this information to write a preliminary job description which would serve as a guideline when writing the different BOS items. This job description is given below to familiarize the reader with the job of the truck driver and with the items appearing on the BOS.

Job description

The main purpose of the truck driver is to transport the company's products as speedily and safely as possible to the

customers. However, to be successful in his job, a driver must also be able to perform certain other functions. These functions include liaising with customers, supervision of the loading and unloading of cargo, managing of administrative work, maintenance of the truck, as well as the supervision of labourers.

The different functions to be performed by the driver in the abovementioned areas of his work, will now be discussed.

Driving Truck to Customers

- The company vehicle should be driven in a safe and courteous manner to promote the company-public image.

- The driver must receive verbal and written instructions and documents from the transport foreman or the fleet co-ordinator. These delivery instructions must be carefully studied and the location of each customer must be known before leaving the depot. Drivers are expected to take an intelligent interest in the routing of their trucks, so as to secure the maximum deliveries in the shortest possible distance.

- A pre-driving inspection of the truck must be done by the driver which includes checking oil, tyres, fuel, warning

lights, et cetera and it must be refuelled if necessary.

If there are any problems with the truck, it should be reported to the transport foreman immediately.

- The loaded trailer must be coupled to the truck. The driver must ensure secure coupling and check aspects such as the trailer's tyres and lights. Should the trailer not have been loaded, the driver must take it to the loading bay, supervise the loading thereof and check the cargo. He must then report the trailer number and deliver the invoices to the operations room.

- The driver must call the labourers together to inspect the canvas and ropes holding the cargo. In the case of a labourer being absent, the driver has to approach the transport foreman for a substitute.

- He must notify the operations room of his departure or arrival in the truck.

- When driving, he must adhere to traffic rules and regulations.

- A driver's log sheet has to be completed on arrival and departure from every customer.

- Radio contact with the operations room should be maintained at all times, especially when arriving and departing from customers.

- The vehicle must be manoeuvred into marked parking bays or the customer's yard, whichever may be required. Back at the depot, the trailer as well as the truck must be parked safely.

- The tachograph and logsheet must be handed in to the transport foreman at the end of the day and other documents to various people.

Supervising loading and unloading of cargo

- The driver should see that his truck is loaded properly and that goods are accessible for offloading. Goods to be delivered first should be loaded last, but when arranging this, due regard must be paid to the correct distribution of the weight of the load over the body of the truck. A truck should not be loaded in excess of the load capacity painted on the vehicle.

- The driver must ensure the presentability and quality of products loaded and delivered (a bottle with a torn label should not be delivered).

- He must check that the cargo is unloaded according to invoice and to the customer's instructions.

- The driver must check the loading of empty containers and bottles. He must also check for foreign and chipped bottles which may not be returned. If necessary, the empty bottles must be sorted on the customer's premises.

- The receipt of the amount of containers and bottles collected from a customer should also be completed by the driver.

- In case of breakages by the driver, he must give the customer credit on the back of the invoice. Credit for empty bottles returned and cash on deliveries should also be taken into account.

- When returning to his depot, the empty containers should be off-loaded at the dry goods department and the full containers at the warehouse.

Liaise with customers

- The driver must keep in mind that he is also a salesman and as a representative of his company he must at all times

be courteous to the customer and not use any bad language or make offensive remarks.

- Furthermore, the driver must negotiate with customers in terms of when to start unloading, in what order and where the liquor must be unloaded.
- Ensure bank guaranteed cheques for the correct amount before off-loading.
- Receive cheques and cash from customers and issue COD receipts. Reconcile the cash received on delivery with the transport foreman on return to depot.
- Ensure that the customer signs the invoice presented by the driver.
- Handle customer complaints and if there was a delay in the delivery of goods to a customer, explain the reasons for this delay.

Administrative work

- Handle accidents as per prescribed procedure.

- Complete first section of driver's log sheet before departure.
- Complete tachochart.
- Fill in log sheet at every stop and departure.
- Complete invoices and indicate breakages, empties and containers. Receive the customer's signature and pass credit.
- Complete the empty return slip after adding empties appearing on back of invoice.
- Complete log sheet and tachochart at the end of the day.
- Fill in the monthly operation sheet (this indicates a summary of the day's work).
- Complete COD books and check the cash received.
- Ensure that all documentation is correct and handed in.

Maintenance of truck

- Drivers are responsible for the vehicles they drive. They must realise that the efficiency and costs of the company's

transport service are to a great extent their responsibility. Therefore, they should at all times do everything in their power to ensure that the vehicles in their charge give the most satisfactory and economical service. It should also be remembered that the sales and the service given by the company to their customers depends entirely on the vehicles always being serviceable and available.

- Drivers should familiarise themselves with the service chart. The periods when the engine, sump, gearbox and differential lubricants must be flushed, drained and refilled and the grade of lubricant required will be detailed on the chart for the vehicle concerned and drivers must adhere to the details on the specific chart. The scheduling for servicing the vehicle is the driver's responsibility.

- As trucks have to travel over public roads and through residential and business areas, the driver should remember that his truck is an advertisement for the company. It is therefore essential that it be kept clean and be operated with every courtesy and regard for the safety of pedestrians and the motoring public.

- The driver and his assistants are responsible for seeing that their vehicles are in a clean condition. Everything possible must be done to preserve the good appearance of the vehicles.

- An overall inspection of the vehicle should take place every day. This includes checking brakes, tyre pressure, oil, fuel, licences, insurance papers, permits, etc. An example of such a checklist is included in Appendix 4.

Supervise labourers

The driver must:

- See to the safety of labourers.

- See that the labourers are presentable in general. The driver and his assistants must wear the applicable company uniform or protective clothing. The personal apparel should be in a clean and hygienic condition.

- See that liquor is handled safely and speedily.

- Ensure that labourers are punctual and on the truck in time for departure.

- Represent the labourer to the transport foreman on personal matters.

Now that the reader is familiar with the job of the truck driver, the different behavioural dimensions derived from the job description, will be described.

Description of Behavioural Dimensions

Apart from the job description, the critical incidents had to be classified again. Incidents which were similar in context were grouped into different behavioural items. Interjudge agreement was calculated to determine which items should be included on specific dimensions. The same people who were responsible for the critical incidents originally, were divided into groups A and B. Each group consisted of three transport foremen, one transport manager and five truck drivers, all randomly assigned. This was to ensure the acceptance of the appraisal scale to all the different parties in the Transport Department.

In this process, six behavioural dimensions, differing slightly from those included in the job description, were identified. The six dimensions were: Responsibility, Safety, Supervision, Liaison with customers, Maintenance of truck and Performing administration. Groups A and B then had to classify the critical incidents into the six different dimensions. Interjudge agreement on this classification was assessed. The ratio of agreement was calculated by counting

the number of behavioural items which both groups agreed should be placed in a specific dimension divided by the combined number of items placed by both groups in that dimension. This process is described in detail by Latham and Wexley (1981). The ratios derived in this way are indicated in Table 6. The interjudge agreement regarding each dimension will now be discussed.

Responsibility

A ratio of 0,77 was calculated in this dimension. This was not the minimum prescribed ratio of 0,80 described by Latham and Wexley (1981), therefore the items included in this dimension had to be re-examined. It was decided that this low correlation could probably be attributed to the fact that this dimension had more items than the other dimensions. Therefore it was decided to include this dimension on the scale without any changes.

Safety

A score of 0,81 received on this dimension confirmed its inclusion on the BOS without any changes.

Supervision

A very high ratio of interjudge agreement was obtained on this dimension. The ratio of 0,96 was the second highest ratio determined. This dimension was also included unchanged on the BOS.

TABLE 6

Ratio of Interjudge Agreement for each Dimension

DIMENSION	RATIO
Responsibility	0,77
Safety	0,81
Supervision	0,96
Liaison with Customers	0,87
Maintenance of truck	1,0
Perform administration	0,81

Liaison with customers

The ratio determined on this dimension was 0,76 after the first classification. After rediscussing this dimension, it was again classified. An item previously included on the dimension "Perform Administration" by Group A, was reassigned to this dimension, increasing the ratio to 0,87, which was high enough to receive a place on the BOS.

Maintenance of truck

The ratio on this dimension was 1,0, the highest ratio determined on any dimension. The reason for this was probably that there were fewer items included on this dimension than on any of the six dimensions. Another reason could be the technicality of the items which made these items easily discernable.

Perform administration

The ratio of interjudge agreement determined on this dimension was 0,81, which indicated that this dimension could also be included on the BOS.

Thus, six dimensions, each consisting of a differing number of behavioural items, constituted the behavioural observation scale as developed in this study.

To determine the importance of each item appearing in the different dimensions, the remaining truck drivers as well as the transport foremen had to rate each item on a 5-point scale (The scale appears in Appendix 5). Not one item received a score of lower than 3, and all of them therefore remained in the list. The mean ratings given to the items was 4,3. This process established the content validity of the BOS.

The attachment of a 5-point Likert scale to each item concluded the rating scale. Percentages were also added to the scale values. Supervisors were instructed to score a 1 if an employee engaged in a certain behaviour 0-20 per cent of the time, a 2 for 21-40 per cent, a 3 for 41-60 per cent, a 4 for 61-80 per cent and a 5 for 81-100 per cent of the time. These are the same percentages used by Latham, Mitchell and Dosset (1978). It was believed that this was an easier method than the one discussed by Latham and Wexley (1981) where more difficult percentages were defined. In this case employees received a 1 if they were observed engaging in a behaviour 0-64 per cent of the time, 2 for 65-70 per cent of the time, 3 for 75-84, 4 for 85-94 and a 5 for 95 to 100 per cent of the time. The degree to which observers can actually distinguish between these percentages is sometimes questioned (Latham & Wexley, 1981.)

Reliability

To determine the internal consistency of the instrument developed, Cronbach's alpha coefficient of reliability was calculated

for all the dimensions on the BOS. These coefficients appear in Chapter 5.

Validity

As discussed earlier, the content validity of the appraisal instrument was determined by putting ten per cent of the critical incidents derived from the sample aside before categorization. After categorizing the remaining 90 per cent of critical incidents, the ten per cent of incidents originally left out were examined to see whether any of them described behaviours which had not yet appeared. If this was the case, it would be assumed that the number of incidents which had been collected were not sufficient for the appraisal instrument to possess content validity (Latham and Wexley, 1981). According to Guion (1980), an important aspect of any determination of content validity, is the definition of a job content universe based on a job analysis. In this study a job analysis was done which defined the content of the job in question. The importance of the final behavioural items appearing on the appraisal instrument derived from this job analysis, was also determined. By doing this it was hoped that the validity of the contents of the appraisal instrument would be ensured.

Performance appraisal at three-monthly intervals

The next phase of this study involved the actual assessment of performance. The appraisal system developed in this study

was used to appraise the performance of the drivers in the experimental as well as the control group. Thus, it would be possible to assess to what extent the new instrument was successful, because it provided for goalsetting, feedback and had credibility. However, as discussed earlier, to obtain the best possible results from a performance appraisal, raters had to be trained. Therefore, the training of the raters will be discussed first, before discussing the actual appraising of the performance of the truck drivers.

Training of the raters

The supervisors of both the control and experimental groups were trained on several aspects with regard to performance appraisal. The training was done by the training department of the company. The training programme used was called the Interaction Management Programme.

The programme covered aspects such as, firstly, how to prepare an employee for a performance appraisal, secondly how to discuss satisfactory performance with an employee and in the third instance, methods for giving feedback on unsatisfactory performance. Raters were also advised on how to discuss issues related to salaries with employees.

The purpose of the training programme was to ensure that the supervisors were familiar with the concept of performance appraisal, as well as with the different rater errors which could be made. As already indicated, simply making raters aware of

these errors may contribute to a more reliable appraisal (Spool, 1978).

According to the Interaction Management Programme, the ultimate measure of success of any performance appraisal is whether or not the appraisal resulted in improved employee performance. To meet this objective, the programme trained supervisors to prepare for performance appraisal and to adequately use the performance appraisal system which reviews satisfactory as well as unsatisfactory performance in the light of previously agreed-upon goals and standards.

The Interaction Management Programme is based on interaction modelling. This modelling usually takes place in the form of role playing or the screening of films or videos. According to Byham and Robinson (1976) no theory is taught, only positive steps for handling each situation. During the training, situations which may arise when supervisors are executing a performance appraisal are identified, and these situations are modelled by specifying positive steps for handling each situation. This corresponds closely with behaviour modelling training, which has been shown to be a successful method for improving interpersonal skills of supervisors (Birkenbach, Kamfer & ter Morshuizen, 1985; Latham and Saari, 1979). A short discussion of the training technique follows.

Six supervisors are usually trained at a time (in the case of this study, three Transport Foremen and three supervisors

from other departments were trained together). The training is structured in such a way that one aspect of human interaction is learned at a time. A step-by-step approach for handling each situation is also provided and a positive model using this approach shows learners how each difficult situation can be handled successfully. Each learner has the chance to practise handling the difficult situation in the classroom.

Raters also received thorough training on the different rating errors which can be made (as discussed earlier) and suggestions for avoiding these errors were put forward. These rater errors were identified and explained to the supervisors. Examples of each rater error were given, and afterwards supervisors had to identify errors in their own ratings and discuss the occurrence thereof.

It was hoped that the training programme would result in a consistent approach to performance appraisal, full documentation of appraisal discussions, more positive employee attitudes towards the appraisal process and greater supervisor confidence in the system. A typical Interaction Management Module is provided in Appendix 6, with the different times allocated to each exercise also indicated.

After the raters had been trained in different aspects regarding performance appraisal, the three different performance appraisals of the control as well as the experimental groups, took place over a six month period.

The actual Appraisal and Feedback Interviews

The newly developed performance appraisal instrument was used to evaluate the performance of the drivers in both the experimental and control groups three times in a six-month period. In other words, three appraisals were done with two three-monthly intervals.

The first appraisal was done before the drivers had any experience with the new system at the beginning of the six-month period. The second appraisal was done after three months and the third one at the end of the six months.

After each appraisal, truck drivers in the experimental group were called in by their supervisors (transport foremen) for feedback interviews on their performance. At this meeting the supervisors also set new goals for the different drivers, especially regarding behaviours on which the driver received a low score. These drivers each received a copy of their performance appraisal form to keep for future reference. Positive aspects of their work were complimented and constructive criticism was given when necessary.

In other words, these feedback interviews with the experimental group took place three times during the six months and it was expected that the drivers would receive better scores on the second and third appraisals than they did on the first one. To ensure that supervisors could not be influenced by previous

scores, these performance appraisal scores were kept by the personnel departments.

The performance of the drivers in the control group was also appraised three times but in this case they received no feedback on their performance, nor were any goals set for them. In fact, they were not even aware that their performance had been appraised, as was the case with the previous appraisal system.

The process of appraisal as described above is illustrated in Figure 15.

Apart from the newly developed appraisal system, unobtrusive measures were also used to determine whether the performance of the truck drivers in the experimental group had improved. According to Luthans (1977) unobtrusive measures are a naturalistic observational technique which minimizes subject awareness, and could contribute to the reliability of other measuring techniques.

Because a truck driver actually has specific delivery appointments each day, it is not always possible for him to make more deliveries than prescribed. The number of deliveries is not determined by the driver but usually by his direct supervisor. It is possible for him, however, to make deliveries in a much shorter time, so that he could arrive back at the depot sooner in case he should be needed again. It is also possible for a driver to make deliveries in a much more efficient way and to promote the company's image at all times.

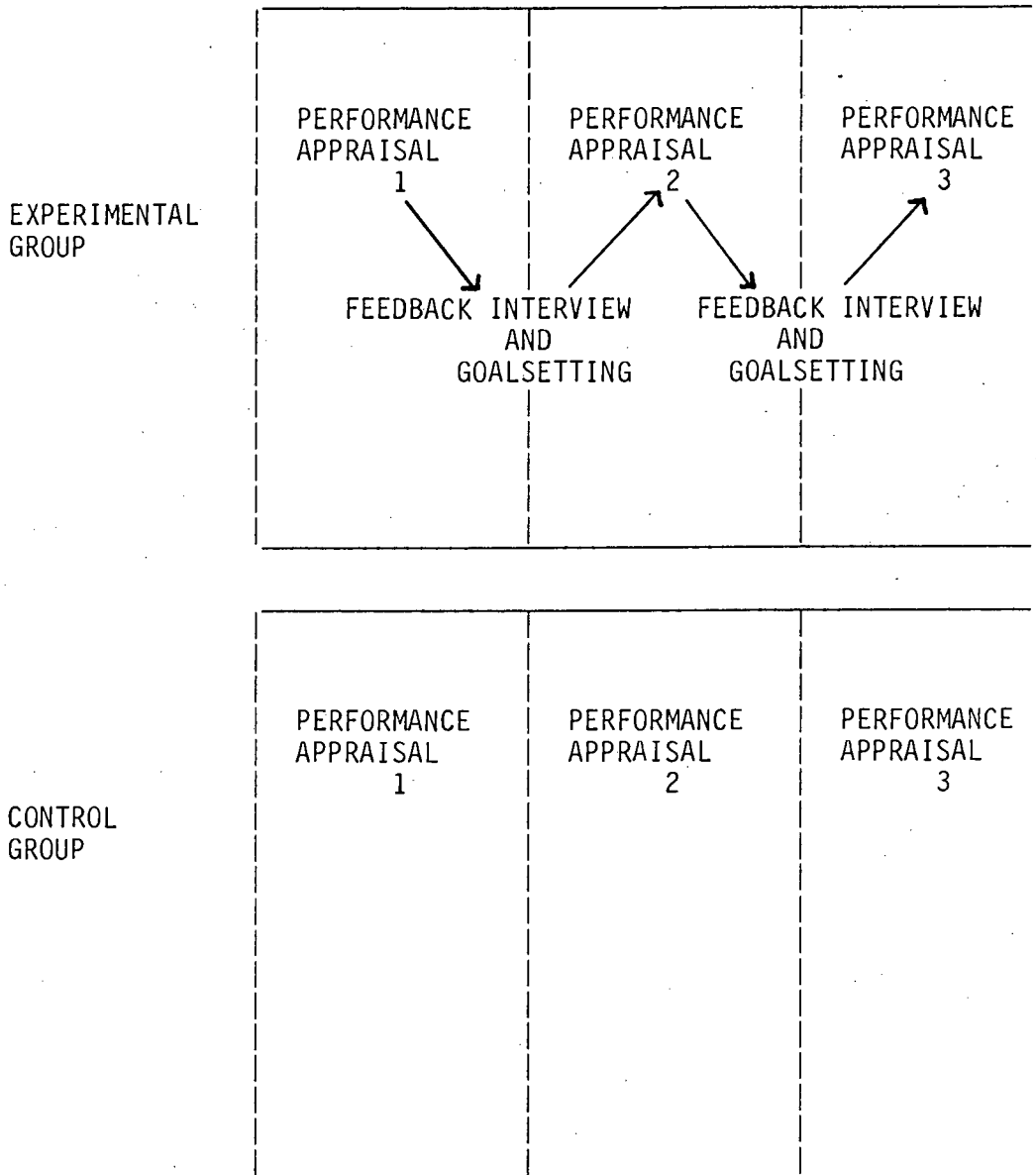


FIGURE 15

Performance Appraisal at three monthly intervals

Therefore, attention was specifically given to unobtrusive measures, such as the appearance of the drivers and their crew, customer complaints, absence of work, accident rates and disobeying traffic regulations.

The rationale for each of the abovementioned unobtrusive measures will now be discussed.

According to Guion (1965), all other things being equal, the employee who attends work regularly is more valuable to any organization than one who frequently misses work. In the case of the truck drivers, one driver being absent influenced the functioning of the whole Transport Department, because a team of workers (general labourers) is left without a supervisor and a loaded truck without a driver. In this instance, the scheduling of work for a whole day has to be rearranged to accommodate the extra truck and working team. Obviously, the reasons for the absenteeism of the drivers were also noted.

Accidents are very costly. The value of a loaded truck may easily exceed one hundred thousand rand. From the organization's point of view, accidents also result in lost working time, treatment costs for the injured, insurance costs, morale problems of employees and poor public relations.

The public image of the company may also be influenced negatively by traffic violations. Therefore it was decided to study the pattern of traffic violations of each driver.

As can be expected, a customer is one of the most important persons to any organization. Therefore, the way in which the drivers liaised with customers had to be acceptable. Thus, the number and nature of complaints received from customers on each driver were also noted.

In the last instance, the appearance of the driver and his crew are also very important to the company's image. Therefore, supervisors were also instructed to observe the personal appearance of these employees very carefully.

Some of the abovementioned aspects were included on the final BOS, including the appearance of drivers and customer complaints. Because separate records were kept of accident rates, traffic violations and absenteeism, it was deemed unnecessary to include these aspects in the BOS.

It was standard practice in the Transport Department to record customer complaints, accident rates, traffic violations and the absentee rate of each driver on a monthly basis. It was therefore a relatively simple matter to study these records to determine whether changes in the abovementioned areas had taken place or not. The supervisors were not aware of the intention of the study; they were only instructed to observe each driver carefully on the job, especially on the dimensions included on the final performance appraisal instrument.

Statistical Analysis

To determine whether there would be any increase in the performance of the truck drivers over the six month period, planned comparisons were used.

A planned comparison may only be performed when there is some a priori hypothesis about what the pattern of means (derived from some sort of experiment) would look like (Loftus & Loftus, 1982). To test a hypothesis for the existence of a specific pattern of population means, a set of numbers (called weights) has to be specified. The pattern of these weights should correspond to the hypothesized pattern. These numbers or weights must add up to zero (eg. if one expects a pattern of means to increase linearly, the weights -1, 0, 1 may be used. On a graph this specifies a linear increase, and if added the total is zero).

These weights are then used to generate a sum of squares due to the hypothesis which is part of the sum of squares between conditions (Loftus & Loftus, 1982). This sum of squares due to the hypothesis is based on 1 degree of freedom and is therefore a mean square as well as a sum of squares. This mean square can be F-tested against the appropriate error term, thus assessing the significance of a particular hypothesis. According to Hays (1973), this is a statistically powerful method to test hypotheses.

Furthermore, the sum of squares due to the hypothesis may be divided by the sum of squares from where it came (sum of squares

between conditions) to determine the percentage of variance among the various conditions that is accounted for by the hypothesis.

If there is more than one hypothesis to be tested, appropriate sets of weights should be generated and sums of squares for the total amount of hypotheses to be tested should be calculated. The significance of these hypotheses may then be tested.

This discussion of planned comparisons was quite brief. A more detailed description of this statistical technique can be found in Hays (1973) and Loftus and Loftus (1982).

Preference was given to planned comparisons for analyzing the data rather than to use one- or two-way analyses of variance. Planned comparisons seem to provide more information to the researcher. According to Pagano (1981), in analysis of variance a significant F-value indicates that all the conditions do not have the same effect on the dependent variable. In other words, a significant F-value only implies that at least one condition differs significantly from at least one of the others, but it does not specify which. It is also possible that they are all different, or that any combination of the conditions could show differences. When using planned comparisons, it is not only possible to determine **whether** conditions differ, but also **which** conditions differ.

Pagano (1981) further stated that planned comparisons are more powerful than other post-hoc tests (eg. Tukey's HSD and

Newman-Keul's tests) and that it is the method of choice when applicable (for a discussion of these post hoc tests, see Pagano, 1981).

The different hypotheses formulated and the rationale for each hypothesis will now be discussed. These hypotheses should not be viewed as separate aspects of the thesis, but as integral parts of the statistical technique. Without some a priori hypotheses, planned comparisons could not have been used.

The Hypotheses

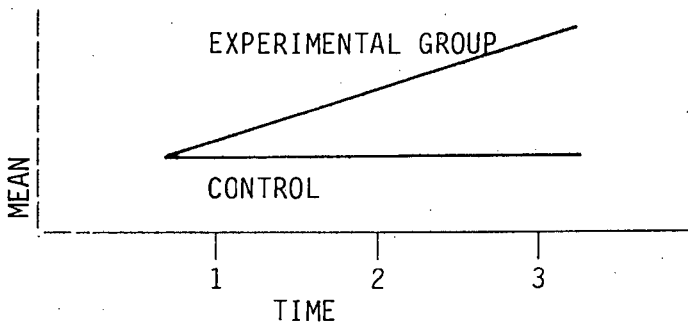
The different hypotheses as discussed below are graphically illustrated in Figure 15.

Hypothesis 1

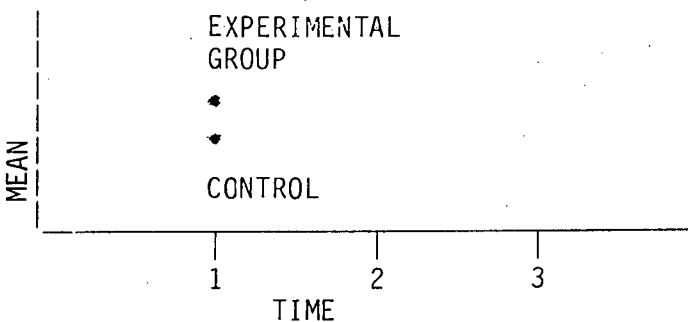
The performance of the experimental and control groups was exactly equal at the outset of this study. The performance appraisal scores obtained in the experimental group would increase linearly towards and up to the third appraisal, whilst the scores of the control group would remain unchanged for the three appraisals.

In Chapter 1 it was stated that the main objective of this study was to develop a performance appraisal system for truck drivers, with the **inherent potential of improving the performance of these drivers**. By setting goals for each truck driver and

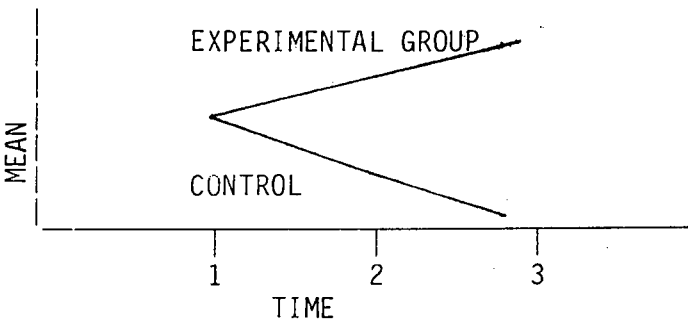
H1



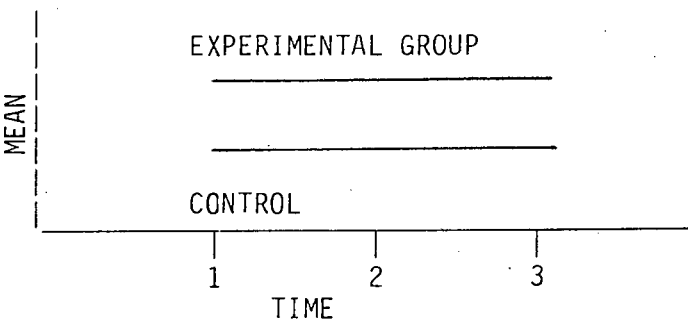
H2



H3



H4



H5

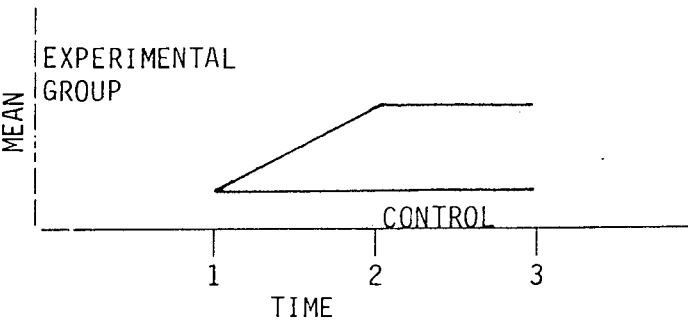


FIGURE 16

Hypotheses Graphically Illustrated

implementing performance appraisal feedback interviews, it was expected that the performance of the drivers would increase, and subsequently their performance appraisal scores. In the literature study, reasons were found to believe that the abovementioned actions would account for improved performance (Glueck, 1979; Latham & Wexley, 1981; Smith & Brouwer, 1977). The first hypothesis, therefore, was formulated to account for the possible improvement of the appraisal scores of the drivers in the experimental group. Because no intervention had taken place in the control group, it was expected that those scores, and thus the performance of the drivers, would remain unchanged during all three appraisals. Hypothesis 1 can therefore be viewed as the most important hypothesis in the sense that it accounted for the main objective of this study.

Alternative hypotheses were also specified to account for other possible outcomes of the study.

Hypothesis 2

The appraisal scores of the experimental and control groups were not equal at the outset of the study.

It was expected that the appraisal scores of the experimental and control groups would be the same for the first appraisal. As confirmed by the job description in Chapter 4, the functions to be performed by the truck drivers in both groups were the

same. Added to that, the transport foremen doing the ratings underwent the same training programme. Therefore, a hypothesis had to be specified which would determine whether the two groups were equal at the outset of the study, and furthermore, whether they were comparable.

Hypothesis 3

The scores of the experimental group would increase towards the third performance appraisal, whilst the scores of the control group would decrease.

The possibility also existed that the scores of the experimental group would increase from the first to the third appraisal, whilst those of the control group would decrease. As already mentioned, the intervention was expected to have a positive effect on the scores of the experimental group. Certain unexpected external factors, such as an increased work load or destabilizing social circumstances, could however have been the cause for drivers in the control group being rated lower than usual, thus negatively influencing this group's score.

This hypothesis was also formulated to ensure that a possible improvement in the scores of the experimental group would not be ascribed to the intervention, without assessing whether these results may appear to be inflated because of a decrease in the

appraisal scores of the control group. By formulating this hypothesis, it was ensured that the results were seen in perspective.

Hypothesis 4

The scores of the drivers in the experimental group were higher compared to those of the control group at the outset of the study. The scores of both these groups would remain unchanged during all three appraisals.

It was possible that due to some effect of the intervention, the appraisal scores of the experimental groups would be higher than those of the control group during the course of the whole study. The attention given to the experimental group, and the involvement of drivers from the experimental group in developing the BOS, may have contributed to improved performance and higher appraisal scores. Hypothesis 4 was thus formulated to determine whether improved performance could be attributed to attention received (Hawthorne effect) (Freedman, Sears & Carlsmith, 1978). Hypothesis 4 differs from hypothesis 2 in the sense that all three performance appraisals are considered, whereas hypothesis 2 only accounted for differences in the experimental and control groups at the first appraisal.

Hypothesis 5

The scores of the experimental group would increase linearly towards the second appraisal, but remain unchanged towards the

third one. The scores of the control group would remain unchanged during all three appraisals.

It was possible that continuous improvement in performance would not take place during the course of the study, due to unattainable or too easy goals set for the drivers as discussed earlier (Latham & Wexley, 1981; McClelland, 1961). Therefore, hypothesis 5 specified improvement in appraisal scores during the first appraisal period, but no change during the second period. The initial effect of the intervention (Hawthorne effect due to attention received) may also have worn off after the second appraisal and drivers were thus not motivated any more for improved performance. As in hypotheses 1 and 4, it was expected that no change would occur in the appraisal scores of the experimental group for all three performance appraisals.

The abovementioned hypotheses accounted for most of the probable outcomes of this study.

As discussed earlier in chapter 4, the drivers' satisfaction with the old and new appraisal systems was also compared, and areas of the appraisal system still needing attention were identified. It was believed that drivers would be more satisfied with the newly developed appraisal system, because recommendations made by them regarding performance appraisal were incorporated in the development of the BOS.

In this chapter, the extent of and the procedures followed in this study were discussed. The results obtained from the various exercises mentioned above are presented and discussed in Chapter 5.

CHAPTER 5

Results and Discussion

The first stage of this study initially emphasized the workers' reactions to the old performance appraisal system, thus identifying certain needs of the employees regarding this aspect. The second stage of the study emphasized the need of the organization by evaluating the extent to which the performance of truck drivers had improved after the intervention. This dual approach would give a clearer indication of the extent to which the behaviourally-based appraisal system developed in this study would meet the demands of both the organization and the employees.

Initially, only the results of the first measurement of the experimental group's reactions to the old appraisal system will be reported. The control group did not participate in this exercise, because they were not aware of either the development of the new appraisal system or of the fact that the experiment had taken place. Furthermore, because they were not aware of the new system, they could not judge the extent to which this system satisfied their expectations.

Secondly, the development of the Behavioural Observation Scale for the truck drivers will be described before discussing the effect of this appraisal system on the performance of these drivers.

In the third instance the experiences of the drivers with the "old" and "new" appraisal systems will be compared to determine whether the drivers were more satisfied with the "new" system.

In the last instance, other possible expectations of the drivers after having been exposed to the new appraisal system, will be discussed. This would provide information regarding areas of the performance appraisal system which may still need some attention and improvement.

Preliminary Study : The Need for a Performance Appraisal System

The need for a new performance appraisal system was established by means of the Performance Appraisal Satisfaction Questionnaire (Van Wyk, 1983). This instrument assessed the experiences of the truck drivers with the "old" performance appraisal system and identified their expectations of a new system. These expectations were incorporated in the development of the Behavioural Observation Scale.

The reliability of this measurement was determined by calculating Cronbach's coefficient alpha. Because two responses had to be given to one item, reliability was calculated for both responses, namely experiences and expectations. The coefficient alpha for experiences was 0,81 and for expectations 0,69. The

reliability of the instrument was considered satisfactory. The lower reliability of the expectations could possibly be attributed to the fact that 40 **different** people indicated 40 **different** sets of expectations, whilst the experiences of these people were based on **one** specific situation.

The mean scores obtained from the truck drivers in the experimental group on the 16 different questions in the Questionnaire are reported in Table 7. The numbers of the questions appearing in this table correspond with the numbers appearing in the questionnaire in Appendix 2. The experiences of the drivers with the old system are indicated by X_E , and their expectations of a new performance appraisal system by X_{Ep} . The discrepancies between drivers' experiences with and their expectations of a performance appraisal system are indicated by $X_{Ep} - X_E$.

A t-test for dependent samples was performed to determine the significance of this difference, resulting in $t = 9,72$ ($p < 0,001$). This indicated a significant difference between what drivers experienced with the existing performance appraisal system and their expectations of a new system. Considering the reliability of the measuring instrument, it was clear that the truck drivers expected far more of an appraisal system than they were receiving.

The abovementioned results stressed the importance of a performance appraisal system to the organization as well as to

TABLE 7

Discrepancy between Experiences and Expectations of Truck Drivers regarding Performance Appraisal

	EXPERIENCED	EXPECTED	DIS-CREPANCY
N = 40	\bar{X}_E	\bar{X}_{EP}	$\bar{X}_{EP} - \bar{X}_E$
QUESTION NUMBER			
1	1,1	9,1	8,0
2	1,2	10,5	9,3
3	1,3	10,8	9,5
4	1,6	10,9	9,3
5	2,3	9,9	7,6
6	1,0	10,9	9,9
7	1,5	10,9	9,4
8	1,9	10,9	9,0
9	2,0	10,5	8,5
10	1,9	11,0	9,1
11	1,2	10,8	9,6
12	1,9	10,5	8,6
13	1,4	10,4	9,0
14	0,9	10,2	9,3
15	2,6	10,3	7,7
16	1,6	10,8	9,2
	$\bar{X} = 1,6$	$\bar{X} = 9,9$	$\bar{X} = 8,9$

the employees. Employees are generally interested in how they are doing on their jobs, and how they are viewed by the organization (Slivinski, 1975). Therefore the organization should try to cater for these needs.

Discrepancies of varying degrees were indicated on all sixteen questions. The most important discrepancies as obtained from the abovementioned exercise will be discussed before describing areas in which smaller discrepancies were recorded.

From Table 7 it is evident that the largest discrepancy ^{was} existed with regard to question 6. This implied that the drivers were of the opinion that their performance was not measured against factors previously agreed upon by them and their supervisors. In other words, drivers were not aware which aspects of their performance were to be appraised. This is in direct opposition to the criteria for good performance systems which were discussed in Chapter 2. Wells (1982) stated that an employee should know the standards against which he/she will be evaluated. Sauser (1980) supported this view when he stated that standards, expected results and goals should all be identified at the beginning of the appraisal period. Cascio (1982) further said that a performance appraisal system should be able to differentiate between successful and unsuccessful behaviour and that both these concepts must be described in clear, understandable terms.

It thus appeared as if the drivers were unaware of which aspects of their job the organization viewed as important. They performed their jobs without knowing what a "good" driver had to do, and because of this shortcoming, the need to be informed of all aspects of their job was identified. The large discrepancy on question 3 also supports the above. Drivers did not view the factors against which their performance was appraised as important to their jobs. They would have preferred to, in conjunction with their superiors, set standards or goals to be reached and against which their performance could be measured afterwards (the discrepancy on question 11 is also one of the highest recorded).

The drivers were also very dissatisfied with the lack of frankness of discussions with their supervisors after performance appraisals (question 7). This is an unfortunate situation, as a discussion of the results of an appraisal helps employees to achieve a clear understanding of their value to the organization (Cascio, 1982; Graves, 1982; Landy, 1985). The appraisal interview is an ideal situation for counselling of employees. During the interview, employees can be told how they are rated and why (Schuler, 1981). The importance of the appraisal interview was stressed by a low discrepancy score on question 15. This indicated that drivers felt each time they discussed their performance with their supervisors, their relationship had been influenced positively. This happened in spite of the fact that they

experienced the discussion as not being frank and open. In other words, a need for personal contact with the supervisors could be identified amongst the drivers. This confirmed the views of Burke, Weitzel and Weir (1978) who found that participation of employees in the appraisal interview correlated highly with the employees' satisfaction with the appraisal process. A high discrepancy score recorded on question 4 also implied that drivers expected to air their own views regarding their performance during feedback interviews.

Other problem areas identified during this exercise will now be discussed.

In the first instance, drivers felt that they did not have enough time to prepare themselves for possible appraisal interviews (Question 2). On closer examination it was found that the drivers had to complete a great deal of administrative work (as discussed in the job description in the previous chapter) in preparation for performance appraisal interviews. This need had to be accommodated in the new appraisal system.

Secondly, drivers felt that the way performance appraisal was handled in the company definitely did not have a positive influence on their performance. This is contrary to opinions voiced by researchers (Eichel & Bender, 1984; Graves, 1982; Latham & Wexley, 1981; Richarddt, 1976; Sauser, 1980; Wells,

1982) that a performance appraisal system should make a positive contribution to the organization, as well as the employees. The drivers further indicated that performance appraisal on the whole was not practised effectively in the company (question 16).

As mentioned earlier, another question which yielded a high expectation score was the need of drivers to be assisted by supervisors when planning for improvement in their performance (question 11). Drivers clearly needed co-operation and support from their supervisors while planning for future development. This may be due to the fact that supervisors and their employees did not set mutually agreed upon goals in the "old" system as specified by Latham and Wexley (1981). Smith and Brouwer (1977) also stated that goal setting gives purpose and direction to performance appraisal and to the developmental processes.

The lowest discrepancy score (question 5) indicated that the drivers felt that their performance was evaluated objectively. In other words, they were satisfied with the way in which their performance was being evaluated by the different supervisors. This could be viewed as a vote of confidence in the appraisal skills of their supervisors, and could result in the supervisors being more at ease during the performance appraisal feedback interview. The new appraisal system could thus build on this trust between employees and their supervisors.

Another question (number 17) determined the frequency with which drivers wanted their performance appraised. It was indicated that the mean time-period between performance appraisals should not exceed 3,5 months. In other words, regular performance appraisals, three to four times a year, would seem to be in order. This again emphasized the need for performance appraisal to be a continuous process rather than a once a year event (Haynes, 1978).

On the grounds of the abovementioned results, it was evident that a definite need for a new performance appraisal system existed for truck drivers. The drivers were not satisfied with the previous appraisal system and they really expected much more of another system. In order to develop an acceptable appraisal system for the truck drivers, their expectations regarding performance appraisal were incorporated in the new system.

The development of the BOS as well as the resulting changes which occurred are presented below.

Results of Performance Appraisal

The results of the various exercises connected with the three-monthly performance appraisals will first be presented and discussed afterwards. The statistical characteristics of

the BOS will be discussed, which include the means, standard deviations, reliability and validity of the measuring instrument. The actual comparison of the three different performance appraisals, will be discussed later in this chapter.

Means and Standard Deviations

The means and standard deviations for each item on all three performance appraisals are indicated in Table 8. These results influenced other statistical techniques still to be described. It appears from Table 8 that some changes in the mean scores of both the experimental and control groups had taken place. The significance of these changes was determined by means of planned comparisons. The results are discussed later.

Reliability

The reliability of the Behavioural Observation Scale constructed in this study was determined by calculating coefficient alpha for each dimension. The overall reliability of the measuring instrument, including all the different items from each dimension, was also calculated. These results are reported in Table 9.

The overall reliability of the appraisal instrument developed was ,98. The dimension with the highest coefficient alpha was

TABLE 8

Means and Standard Deviations for Different Performance Appraisals

	PERFORMANCE APPRAISAL 1		PERFORMANCE APPRAISAL 2		PERFORMANCE APPRAISAL 3	
	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD
<u>EXPERIMENTAL GROUP</u>						
DIMENSIONS:						
Responsibility	3,12	0,81	3,41	0,62	3,56	0,56
Safety	3,18	0,90	3,38	0,66	3,54	0,61
Liaison with Customers	3,03	0,77	3,35	0,61	3,51	0,53
Supervision	3,29	0,58	3,33	0,52	3,47	0,52
Maintenance of Truck	3,05	0,66	3,20	0,52	3,41	0,55
Administration	3,20	0,68	3,44	0,53	3,63	0,53
<u>CONTROL GROUP</u>						
DIMENSIONS:						
Responsibility	3,33	0,57	3,34	0,53	3,32	0,51
Safety	3,40	0,65	3,42	0,58	3,46	0,61
Liaison with Customers	3,32	0,69	3,35	0,65	3,33	0,61
Supervision	3,26	0,46	3,26	0,51	3,24	0,53
Maintenance of Truck	3,11	0,55	3,20	0,48	3,15	0,44
Administration	3,29	0,61	3,31	0,59	3,29	0,53

Skakelincj

TABLE 9
Reliability of Measuring Instrument

DIMENSION	COEFFICIENT ALPHA
Responsibility	0,94
Safety	0,92
Liaison with Customers	0,93
Supervision	0,86
Maintenance of Truck	0,75
Administration	0,91
Overall Reliability	0,98

Responsibility (.94). This dimension included more items than any other dimension and the high reliability coefficient may possibly be attributed to that. The dimension with the lowest reliability coefficient was Maintenance of Truck (.75). This dimension had the lowest number of items included compared to the other dimensions and this is believed to be the cause of the lower coefficient alpha.

To determine whether each item deserved a place on the scale, coefficient alpha of the scale after the removal of each item was calculated. In other words, for each individual item the reliability coefficient, coefficient alpha was computed from all the other items (Nie & Hadlai Hull, 1981). This process refers to the internal consistency of the scale, which provides a measure of the extent to which the instrument is free of content sampling error (Latham & Wexley, 1981). Lemke and Wiersma (1976) stated that when items measure similar concepts, the reliability of the total test increases. When different concepts are measured, the reliability decreases. Ghiselli, Campbell and Zedeck (1981) supported this when stating that to the extent that items included on a scale are not homogenous, the internal consistency estimate will be lowered, thus decreasing the reliability of the measurement.

White (1982) confirmed that when developing a measuring instrument, items may be retained or discarded on the basis of their impact on overall reliability. He further mentioned that

if the removal of any particular item from a measuring instrument resulted in an improvement in the reliability of the instrument, the item should be discarded.

The abovementioned procedure was thus performed to assess whether items which did not contribute to the overall reliability and internal consistency of the appraisal instrument, were included on the observation scale.

The results of this procedure are reported in Table 10 (If only two decimal figures are considered, no change is recorded in the reliability. Thus, reporting only two decimals as is usual would be futile. Therefore three decimal figures are reported in Table 10. This will be done for the rest of the chapter when considered necessary).

It was found that the removal of any item from the scale did not increase the overall reliability, but rather decreased it marginally. It was further found that the removal of any item in a specific dimension also did not increase the reliability of that dimension but rather decreased it somewhat. In the dimensions Responsibility and Supervision, the removal of two items in each dimension decreased the reliability of the measuring instrument more markedly (0,980 to 0,979). If the BOS developed in this study is studied (Appendix 7), it is obvious that these four items represent very important aspects of the drivers' functions, for

TABLE 10

Reliability of Scale if each item is Deleted

		SCALE MEAN IF ITEM DELETED	SCALE VARIANCE IF ITEM DELETED	CORRECTED ITEM TOTAL CORRELATION	ALPHA IF ITEM DELETED
RESPONSIBILITY	RE 1	120,150	429,751	,683	,980
	RE 2	120,008	427,062	,757	,980
	RE 3	120,071	426,342	,760	,980
	RE 4	119,967	424,517	,798	,980
	RE 5	119,933	427,293	,757	,980
	RE 6	119,925	423,032	,829	,979
	RE 7	119,954	424,470	,762	,980
	RE 8	119,988	419,376	,854	,979
	RE 9	120,071	422,334	,782	,980
	RE10	120,062	429,648	,768	,980
SAFETY	SA 1	120,008	422,628	,801	,980
	SA 2	120,008	426,150	,756	,980
	SA 3	119,954	421,583	,794	,980
	SA 4	119,937	424,736	,758	,980
	SA 5	119,933	423,870	,791	,980
LIAISON WITH CUSTOMERS	L 1	120,058	428,214	,710	,980
	L 2	120,104	429,917	,690	,980
	L 3	120,000	421,740	,817	,980
	L 4	119,941	423,912	,768	,980
	L 5	120,150	420,546	,814	,980
	L 6	120,029	419,945	,830	,979
	L 7	120,054	423,130	,796	,980
	L 8	120,050	417,763	,849	,979
SUPERVISION	S 1	120,062	430,393	,665	,980
	S 2	120,095	432,547	,636	,980
	S 3	120,041	431,789	,621	,980
	S 4	120,087	431,427	,662	,980
	S 5	120,000	430,452	,680	,980
MAINTENANCE OF TRUCK	M 1	120,175	430,772	,663	,980
	M 2	120,154	430,925	,661	,980
	M 3	120,196	428,375	,756	,980
ADMINISTRATION	A 1	120,020	427,426	,734	,980
	A 2	119,983	425,079	,781	,980
	A 3	120,013	427,593	,746	,980
	A 4	119,991	428,292	,730	,980
	A 5	119,979	424,999	,796	,980
	A 6	120,041	424,768	,780	,980

RELIABILITY COEFFICIENTS

N OF CASES = 240,0

N OF ITEMS = 37

ALPHA = ,981

example, cash reconciling, handling road accidents as prescribed, off-loading empty bottles and crates and maintaining good customer relations. It thus stands to reason that by removing any of these items from the scale, the reliability of this appraisal measurement could be questioned.

The fact that the removal of no item increased the reliability of the scale, led to the conclusion that all the items already appearing on the scale should be retained.

An item analysis was also conducted. The resulting inter-item correlation matrix appears in Appendix 8. This procedure involved correlating the scores on each behavioural item with the sum of all the items, the purpose thereof being to ensure that each item on the BOS was unambiguous to the rater. According to Ghiselli, Campbell and Zedeck (1981) the items on the final scale should have high item-total correlations. This would indicate that the items are measuring the same dimension.

The results of the item analysis (Appendix 8) indicate that all the items included on the final BOS indeed correlated highly with the sum of all the other items. All the correlations were significant on the 0,01 per cent level, the highest correlation being 0,78 and the lowest 0,33. According to Latham and Wexley (1981), concerns over high intercorrelation scores amongst criteria on the appraisal form should be dealt with judiciously. They stated that extreme redundancy should be avoided, but by discarding

different items with high intercorrelations, accountability and control by the organization may be reduced. Furthermore, feedback to and development of the employee may be impeded.

As all the items on the scale specified certain behaviours to be performed by employees, the supervisors felt that to remove any item, a certain function on which the drivers were to be appraised, would be removed. Therefore, for face validity it had been decided to retain all the items on the scale, irrespective of the fact that various items correlated to a higher extent with the sum of all the other items.

The high reliability coefficient of the behavioural observation scale could probably be attributed to the way in which the instrument was constructed. In the first place, a thorough job analysis was performed, involving some highly experienced employees (Latham & Wexley, 1981). In the second place, the appraisal system was developed for a specific group of employees, thus achieving a tailormade system. According to Richardt (1976) such a step improves the objectivity and reliability of any performance appraisal system.

Performance was also stated in terms of measurable results to be achieved. In this process the truck drivers and their supervisors were all involved in the development of the system. They specified critical incidents, thus phrasing sentences in a way understandable to all using the system. According to Halloran (1981) the supervisor and the jobholder are the most suitable

persons to define behaviour to be evaluated, as they probably have the best knowledge of the job in question.

Interjudge agreement was also assessed by determining which item belonged in which dimension. Thus, general consensus had been achieved in categorizing the items to be included on the scale.

Furthermore, the importance of each item was also determined by having it rated on a 5-point scale. Thus, the content of the BOS developed, could also be viewed as valid.

To summarize, it can be stated that the general guidelines for the development of a behavioural observation scale, as described by Birkenbach (1984) and Latham and Wexley (1984) were followed closely in constructing the appraisal instrument. This resulted in the high reliability of the measuring instrument (The final BOS appears in Appendix 7).

Validity

The methods by means of which the content validity of the BOS in this study was determined, were described earlier.

In the case of this study, five items of the BOS which represented a percentage of 12,8 had been set aside. These were re-examined to determine whether any of these items described behaviours

which had not yet appeared. No new behavioural descriptions were identified, and it was assumed that the content validity of the appraisal instrument was satisfactory.

As discussed in the previous chapter, Guion (1980) stated that the important aspect of any determination of content validity, is the definition of a job content universe based on a job analysis. In this study a thorough job analysis was performed by highly skilled people (an industrial psychologist, for example). The identified functions of truck drivers were also rated on their importance to successful performance on the job. Thus, a job content domain was defined (Guion, 1980) which consisted of the most important functions to be performed by truck drivers.

Responses received on the Performance Appraisal Satisfaction Questionnaire (Van Wyk, 1983) implied that the new appraisal system was acceptable to those involved with it, (eg. supervisors, line managers) and that the drivers were quite satisfied with the system. Therefore, the face validity of the appraisal instrument developed was also considered as satisfactory.

Results of the Implementation of the BOS

The results obtained from the three different performance appraisals of both the experimental and control groups will be presented here and discussed afterwards. The scores of these two groups were compared by means of planned comparisons.

Planned Comparisons

As already discussed in the previous chapter, planned comparisons may only be performed when there is some a priori hypothesis about what the pattern of means (derived from some sort of experiment) should look like. Therefore, five hypotheses were defined, describing the expected outcomes of this study. One hypothesis defined the main objective of this study, whereas the other specified alternative outcomes. These hypotheses have been given in Chapter 4. By performing planned comparisons, it will be determined whether the different hypotheses can be accepted or rejected.

The process of planned comparisons involves deriving a set of numbers whose pattern corresponds to the hypothesized pattern with the only constraint that this set of numbers (also called weights) must add up to zero (Loftus & Loftus, 1982). The question to be addressed is : To what degree is the pattern of population means that corresponds to the conditions in the study similar to the pattern of weights previously specified. In other words, how well do the weights correlate with the population means? If the correlation is high, then the hypothesis is supported. If the correlation, on the other hand, is not very high, then the hypothesis can be rejected. By performing the F-test (as discussed in the previous chapter) the significance of a particular hypothesis may be assessed.

In addition to testing for significance the percentage of variance among the various conditions that is accounted for by

the hypothesis, can be determined. This percentage is equivalent to the Pearson r^2 between the weights and the sample of means (Loftus & Loftus, 1982). Thus, if more than one hypothesis is accepted, the percentage of variance explained by each hypothesis can be compared to determine which one of the hypotheses received stronger support.

The results of the planned comparisons regarding each hypothesis (as discussed in the previous chapter) will now be reported and discussed afterwards. As mentioned earlier, more than one hypothesis may receive substantial statistical support. Therefore, a general discussion will follow after the discussion of each separate hypothesis, wherein the hypothesis which received the strongest support will be described.

Hypothesis 1

The performance of the experimental and control groups were exactly equal at the outset of this study. The performance appraisal scores obtained in the experimental group would increase linearly towards and up to the third appraisal, whilst the scores of the control group would remain unchanged for the three appraisals.

The results obtained from the planned comparisons regarding hypothesis 1 are presented in Table 11. The averaged F-score

TABLE 11

Results of Hypothesis 1

VARIABLE	F-SCORE	SIGNIFICANCE OF F
Responsibility	8,92	**0,003
Safety	1,61	0,206
Liaison with Customers	4,69	*0,031
Supervision	4,85	*0,029
Maintenance of Truck	7,62	**0,006
Administration	11,94	**0,009
AVERAGED	6,34	**0,000
D.F. = 6, 1404		

** p < 0,01

* p < 0,05

was $F(6, 1404 \text{ DF}) = 6,34$; $p < 0,01$. The hypothesis received highly significant support and was thus accepted. The appraisal scores of the drivers in the experimental group had increased significantly, whereas those of the drivers in the control group had not.

To determine the variability in these scores between the experimental and the control group, as well as over performance appraisals one, two and three (in other words over a six month period), the proportion of variance accounted for by the hypothesis in each dimension had to be calculated. This was done by dividing the sum of squares due to the hypothesis (SSH) by the sum of squares between groups (SSB). The percentage of variance can thus be illustrated as $\frac{SSH}{SSB} \times 100$. The percentages of variance explained for this hypothesis, as well as for the other hypotheses are presented in Table 12.

From Table 11, it is evident that five of the six dimensions received significant support, whereas only one dimension, namely Safety, did not seem to follow the hypothesized pattern of means. The dimension in which the highest percentage of variance was explained, was Administration (93,9%). This implies that almost 94 per cent of the variability in Administration scores between the experimental and control groups over performance appraisals one, two and three was accounted for by the pattern specified in hypothesis 1. Other dimensions for which a large percentage

TABLE 12

Percentage of Variance Explained

DIMENSIONS	HYPOTHESIS 1			HYPOTHESIS 2			HYPOTHESIS 3			HYPOTHESIS 4			HYPOTHESIS 5		
	SSH	SSB	% VARIANCE	SSH	SSB	% VARIANCE	SSH	SSB	% VARIANCE	SSH	SSB	% VARIANCE	SSH	SSB	% VARIANCE
RESPONSIBILITY	3,35	4,64	72,2	0,84	4,64	18,1	0,79	4,69	17,0	0,14	4,64	3,0	3,7	4,64	5,82
SAFETY	0,74	2,77	26,7	0,93	2,77	33,5	0,00	2,77	0,03	0,24	2,77	8,6	0,42	2,77	15,2
LIAISON WITH CUSTOMERS	2,02	4,92	41,1	1,69	4,12	34,3	0,09	4,92	1,8	0,09	4,92	1,8	1,53	4,92	31,1
SUPERVISION	1,34	1,43	93,7	0,13	1,43	0,9	0,92	1,43	65,0	0,70	1,43	48,9	0,99	1,43	69,2
MAINTENANCE OF TRUCK	2,23	2,98	74,8	0,5	2,98	1,6	0,82	2,98	27,5	0,22	2,98	7,4	1,51	2,98	50,7
ADMINISTRATION	4,20	4,47	93,9	0,13	4,47	2,9	1,99	4,47	44,5	0,96	4,47	21,5	3,53	4,47	78,9

of variance was explained, included: Supervision (93,7%); Maintenance of Truck (74,8%) and Responsibility (72,2%). The dimensions Liaison with Customers (41,1%) and Safety (26,7%) followed the hypothesized pattern to a lesser extent.

Discussion

Hypothesis 1 received substantial statistical support. It thus appears as though the appraisal scores of the experimental group had indeed increased linearly during the three performance appraisals, whilst those of the control group remained relatively constant over the six-month period. It may thus be assumed that the performance appraisal instrument developed in this study, namely the BOS, contributed to an improvement in the performance of the drivers in the experimental group.

This increase in the appraisal scores can possibly be attributed to the way in which the BOS was developed, for example, personal goals were set for each driver and each driver also received feedback on his performance during an interview with his supervisor.

In the first instance, the BOS was developed as prescribed by Birkenbach (1984) and Latham and Wexley (1981), experts in this field. One of the most important aspects identified by Latham and Wexley (1981) was employee participation in the develop-

ment of the system. Therefore it was decided to involve the truck drivers and their transport foremen as much as possible in the development of the system. They were involved in writing the job description (Halloran stated in 1981 that the job holder and his/her direct supervisor are the most suitable people to write the job description). Furthermore, the supervisors were involved in determining the behavioural dimensions on the BOS, and the truck drivers had to rate the importance of each item included on the appraisal instrument. By involving the drivers, and especially the transport foremen in the development of the BOS, it was hoped to gain their acceptance of the new system. Richardt (1976) stated that because the line managers are usually the people who have to do the appraisal, the system must be acceptable to them. If they are not consulted during the development of the system, they may see serious flaws in the system and therefore reject it.

Wexley, Singh and Yukl (1973) also reported greater motivation to improve if employees are allowed to participate in the appraisal process. Employee involvement in the form of personal goal setting or discussions regarding performance also creates more satisfaction on the side of the employee, as well as the manager (Blake & Mouton, 1961; Fletcher, 1973; Wexley et al., 1973). Latham and Wexley (1981) further stated that by involving employees in the development of an appraisal system, the job functions are stated in such a way that they are unambiguous to the rater, as well as the ratee.

In other words, after the development of the system, the truck drivers were aware of exactly what was expected from them on the job (seeing that they themselves determined the importance of each item included on the scale), and raters had no doubt as to what aspects of the truck drivers' performance they had to appraise, which gave them a sense of involvement.

In the second instance, an additional manipulation, namely the setting of personal goals during performance appraisal interviews, may also have had a positive influence on the performance of the truck drivers. During the performance appraisal periods, poor aspects of the drivers' performance were identified and they were instructed to give more attention to those areas of their jobs. In other words, personal goals for each driver were set towards which his behaviour could be directed. This is a very important aspect of any performance appraisal system (Latham & Wexley, 1981), because it gives meaning and relevance to appraisal and development activities (Smith & Brouwer, 1977). Truck drivers were also allowed to participate in the setting of goals by expressing their own ideas about goals to be reached in the future. Latham and Saari (1979) indicated that participation in goal-setting could make a positive contribution to the striving of employees towards these goals.

In the third instance, performance appraisal feedback interviews were held after each performance appraisal. During these

interviews the drivers were allowed to voice their opinions regarding their performance and discuss their work with the supervisors. Mathis and Jackson (1982) stated that the feedback interview is vitally important, because this provides an opportunity for the organization to communicate the results of a performance appraisal to employees, thus helping them to achieve a clear understanding of how they are viewed by their supervisors, as well as by the organization (Alewine, 1982; Cascio, 1982; Wells, 1982). This enables employees to determine which areas of their work are viewed as good or poor by the organization and to take the required steps to improve.

A definite improvement in the performance of the truck drivers in the experimental group had thus taken place, and the BOS as developed in this study may have been the cause of this improvement. The performance of the drivers in the experimental group had remained unchanged during the course of this study, as expected.

The results of the planned comparisons regarding the six different dimensions of the BOS, indicated an improvement in appraisal scores on all of these dimensions. Some dimensions seemed to follow the hypothesized pattern of means more than others, as already indicated in Table 11.

The reason for the dimensions "Liaison with Customers" and "Safety" not following the hypothesized pattern to the extent

of the other dimensions, may be due to the fluctuations in the means of the control group (refer Table 8). After discussions with the supervisors in the experimental group, it was believed that the reason for the dimension Safety not following the hypothesized pattern, was the fact that there was not that much room for improvement on this dimension. Drivers usually prided themselves on the fact that they did not cause any accidents or get traffic fines. A truck is also a very expensive item (as well as the cargo on the truck) and the driver therefore would not want to be responsible for damaging his vehicle in fear of disciplinary actions by the company or of a possible contribution to the reparation thereof. This provided enough motivation to ensure the safe driving of a truck. Drivers also received rather clear feedback from society in the form of traffic fines or complaints to their supervisors on the way they were driving their trucks.

The different items included under the dimension Safety, were very important prerequisites to the successful functions of a driver. These items in the first place included a predriving inspection of the truck and the coupling of the trailer (refer job description in Chapter 4). Cargo also had to be loaded very securely and the weight thereof distributed evenly over the body of the truck and trailer to ensure better roadholding. The safety of the labourers on the back of the truck was also a very important aspect, as the driver was responsible for the lives of these

people. Drivers also had to adhere to traffic regulations when driving to and from the depot.

It is obvious that the abovementioned functions were the most essential aspects of the job of truck drivers. Therefore these functions had to be performed almost perfectly, thus not leaving much room for improvement. It is important to keep in mind that the dimension "Safety" did follow the hypothesized pattern of means to some extent (percentage of variance explained was 26,7%) and that an improvement on this dimension, however small and insignificant, had indeed taken place.

A very positive aspect was the improvement of the scores of the experimental group regarding the dimension "Administration" (percentage of variance explained was 93,9%). After discussing this with the transport foremen, it became evident that this was the one area of the drivers' job where a real need for improvement existed. Although drivers had to make daily entries on their respective log sheets, and had to hand in various documents (eg. driver log sheets and empty return slips) on a daily basis, these documents were only controlled and checked by the transport foremen at the end of each month. Drivers were aware of this fact and as a result sometimes handed in documents later than the target dates. However, when they realized that their performance was also being appraised on their punctuality with the handing in of these documents, this aspect of their job improved significantly.

In other words, administrative work was dealt with much more quickly than previously. As administrative work was such a large aspect of a driver's job, the improvement on this dimension not only positively influenced the functioning of the transport department as a whole, but also that of the organization.

The supervisory skills of the drivers also appeared to have improved quite substantially (percentage of variance explained was 93,7 per cent). For the first time these drivers were personally confronted with people trained in supervisory skills (transport foremen in feedback interviews) and behaviour modelling could thus have taken place. The improvement in this dimension implied improved relations of the drivers with their labourers and better supervision regarding the presence and appearances of their crews. Supervision in general (supervising loading and unloading of truck, and ensuring the presentability of products delivered) also seemed to have improved. Better supervision could also result in the general labourers being more satisfied with their jobs and therefore contribute to the improvement of their performance as well.

The appraisal scores on the dimension "Responsibility" also improved notably, implying a more positive and responsible attitude of drivers towards their job. Drivers appeared to maintain time schedules to a greater extent than before and also tried to improve

on their usual delivery times. Instructions from transport foremen were adhered to much more quickly, thus also positively influencing the discipline of the department. It was further reported that more regular radio contact by drivers was maintained with the operations room when leaving the depot to make deliveries. This enabled the operations room to keep better track of their drivers and their respective positions when they were out on the road, and to plan accordingly if extra deliveries had to be made.

The improvement of the scores in this dimension further implied that cash was treated more carefully and that the reconciliation thereof with the transport foremen took place with greater care. This was a very important aspect, because if the vast amounts of money involved in the drivers' job were considered, it was probably one of the most important jobs in the organization.

An improvement in the appraisal scores on the dimension "Maintenance of Truck" further implied that trucks were washed and cleaned more regularly and that mechanical maintenance was performed more often. This could probably contribute to a lengthened service period for each truck.

Hypothesis 1 further confirmed that the experimental and control groups were equal at the outset of this study. However, added to the fact that hypothesis 1 tried to account for possible differences in the experimental and control groups at the outset of the study, it also had to account for increased appraisal scores of the experimental group. To make absolutely sure that

the experimental and control groups were comparable and equal at the outset of the study, hypothesis 2 was formulated.

Hypothesis 2

The appraisal scores of the experimental and control groups were not equal at the outset of the study.

The results obtained from the planned comparisons regarding hypothesis 2 are presented in Table 13. The averaged F-score was $F(6, 1404 \text{ D.F.}) = 1,66$; $p > 0,05$. None of the separate dimensions received significant support.

Discussion

This hypothesis can be rejected because it did not receive significant support. It therefore implied that performance of the experimental and control groups were not significantly different at the outset of the study, or at the first performance appraisal.

This hypothesis confirmed the expectation that the two groups performed equally at the outset of the study. As mentioned earlier, the transport departments in both the experimental and control groups functioned in more or less the same way. A job description in each department confirmed this. Thus, the functions to be performed by all the drivers involved in this study (both

TABLE 13

RESULTS OF HYPOTHESIS 2

VARIABLES	F-SCORE	SIGNIFICANCE OF F
Responsibility	2,29	0,14
Safety	1,99	0,16
Liaison with Customers	3,92	0,49
Supervision	0,05	0,83
Maintenance of Truck	0,17	0,68
Administration	0,36	0,55
AVERAGED		
D.F. = 6, 1404	1,66	0,13

control and experimental groups) were basically the same. Raters in both groups underwent the same training and raters in the control group were instructed not to discuss any aspect of performance appraisal with the drivers under their supervision. Therefore, no reason existed to suggest that the experimental and control groups would differ at performance appraisal one. This hypothesis supported hypothesis 1, which specified that the two groups were equal at the outset of the study.

Hypothesis 3

The scores of the experimental group would increase towards the third performance appraisal, whilst the scores of the control group would decrease.

The results obtained from the planned comparisons regarding hypothesis 3 appear in Table 14. The averaged F-score obtained on this hypothesis was $F(6, 1404 \text{ D.F.}) = 2,11$; $p < 0,05$. The only dimension that received significant support was "Administration" ($p < 0,05$).

Discussion

This hypothesis was partially supported. Only one dimension, namely Administration, received significant support. The percentage of variance explained by this dimension was 44,5 per cent.

TABLE 14

Results of Hypothesis 3

DIMENSION	F-SCORE	SIGNIFICANCE OF F
Responsibility	2,11	0,15
Safety	0,02	0,96
Liaison with Customers	0,21	0,65
Supervision	3,35	0,07
Maintenance of Truck	2,81	0,09
Administration	5,65	*0,02
AVERAGED		
D.F. = 6, 1404	2,11	*0,49

* $p < 0,05$

The dimension Supervision explained 65 per cent of the variance over performance appraisals one, two and three but did not receive significant support. This high percentage of variance can probably be attributed to a slight decrease in the mean scores of the control group on this dimension (Refer Table 8).

Regarding the decreases in the scores on the dimension "Administration", the following is relevant. As discussed earlier, it was easy for drivers to disregard some of their administrative work and to hand it in to their foremen later than specified target dates. They were aware of the fact that this administrative work was only controlled once a month, and therefore saw no reason to complete it on a daily basis. Supervisors thus only realized that a job was not done correctly at the end of the month. Therefore, in difficult situations (such as increased work load or social unrest as discussed below) this aspect was usually the first to receive less attention from the drivers.

A possible explanation for the decrease in the scores of the control group, could have been the social unrest in black and coloured townships, which increased towards the end of the study. Soon after completion of the study, a state of emergency was declared in this area. This was sure to influence the performance of employees to some extent. In contrast the experimental group in the Western Cape did not experience the same levels of unrest at that point in time. The unrest only erupted during the second

half of 1985 in the Western Cape, by which time the data for this study had been collected.

Of course, there may be other contributing factors which were not known to the experimenter. One of these could have been that the new performance appraisal system created higher expectations with regard to performance on the part of the supervisors in the control group. Because these supervisors had such a clear understanding of exactly what was expected of drivers, it may have caused them to rate more harshly than before. They also experienced no improvement in the performance of their drivers as time went on, and they had no opportunity to brief their drivers on poor aspects of their performance. As already discussed, the transport foremen in the control group were instructed not to discuss any aspect of drivers' performance with them. This may have frustrated the transport foremen to such an extent, that they rated the drivers' performance lower than before.

However, hypothesis 1 explained a larger proportion of variance on both the dimension "Administration" and "Supervision" (93,9 and 93,7% respectively) compared to hypothesis 3. In view of the fact that hypothesis 3 only received partial statistical support, in contrast with hypothesis 1 which received overwhelming support, it was believed that hypothesis 1 explained the pattern of means on these dimensions to a larger extent. This implied an increase in the scores of the experimental group, but little or no change in those of the control group.

Hypothesis 4

The scores of the experimental group were higher compared to those of the control group at the outset of the study. The scores of both these groups would remain unchanged during all three performance appraisals.

The results obtained from the planned comparisons regarding hypothesis 4 appear in Table 15. The averaged F-score was $F(6, 1404 \text{ D.F.}) = 1,08$; $p > 0,05$. Not one F-score on any dimension was significant.

Discussion

This hypothesis was rejected. As already indicated, it was expected that the BOS and the way in which it was developed in this study would have had some effect on the appraisal scores of the experimental group. This hypothesis confirms that changes in those scores indeed had taken place over the six month period, and that the two groups were not different at the first performance appraisal.

This hypothesis further confirmed that the improvement in the scores of the experimental group could not only be attributed to the Hawthorne effect. If this was the case, it would have

TABLE 15

Results of Hypothesis 4

VARIABLES	F-SCORE	SIGNIFICANCE OF F
Responsibility	0,36	0,55
Safety	0,52	0,47
Liaison with Customers	0,23	0,63
Supervision	2,55	0,11
Maintenance of Truck	0,77	0,38
Administration	2,72	0,10
AVERAGED	1,08	0,37
D.F. = 6, 1404		

been expected that the experimental group's performance would initially be higher than that of the control group and remain that way for the duration of the study. However, as already indicated by the results of hypothesis 1, these scores had increased steadily from the first to the third appraisal.

Hypothesis 5

Hypothesis 5 specified that the mean scores of the experimental group would increase linearly up to the second appraisal and remain unchanged towards the third one. The scores of the control group would remain unchanged during all three performance appraisals.

The results of hypothesis 5 are illustrated in Table 16. The averaged F-score obtained on this dimension was $F(6, 1404 \text{ D.F.}) = 4,88$; $p < 0,001$. Three separate dimensions received significant statistical support, namely Responsibility ($p < 0,05$); Maintenance of Truck ($p < 0,05$) and Administration ($p < 0,01$).

Discussion

Hypothesis 5 received significant statistical support. Therefore it can be accepted that in some instances the performance appraisal scores of the drivers initially increased towards the second appraisal, but remained constant towards the third one.

TABLE 16

Results of Hypothesis 5

VARIABLES	F-SCORE	SIGNIFICANCE OF F
Responsibility	7,92	*0,01
Safety	0,91	0,34
Liaison with Customers	3,56	0,06
Supervision	3,58	0,06
Maintenance of Truck	5,19	*0,02
Administration	10,04	**0,00
AVERAGED	4,88	**0,00
D.F. = 6, 1404		

** p < 0,01

* p < 0,05

However, only the dimensions Administration (78,9%), Supervision (69,2%) and Maintenance of Truck (50,7%) to a lesser extent followed the hypothesized pattern of means.

Although it appears as if hypothesis 1 and 5 received the same amount of statistical support (both being significant at the 0,01% level), the amount of variance explained by the dimensions regarding the two hypotheses should be compared, to determine on which hypothesis the largest percentage of variance was recorded.

If the dimensions which received significant support in hypothesis 5, namely "Administration" (78,9% variance explained), "Supervision" (69,2% variance explained) and "Maintenance of Truck" (50,7% variance explained) are compared with the support received by the same dimensions in hypothesis 1, it is evident that a larger percentage of variance regarding these dimensions was explained in hypothesis 1 (Administration, 93,9%; Supervision, 93,7%; Maintenance of Truck, 74,8%).

However, the statistical support received by hypothesis 5, will have to be considered. It thus appears as if to some extent, the scores of the three dimensions mentioned above, did increase towards and up to the second appraisal, but remained constant towards the third appraisal. It is understandable that the scores could not increase much more after the second appraisal,

the mean scores of these dimensions at performance appraisal two being: Administration (3,31), Supervision (3,26) and Maintenance of Truck (3,20). This indicates that on a 5-point scale, little room for improvement still existed and further implied that in the future other standards would have to be used when rating employees (this will be discussed in Chapter 6).

Hypothesis 5 further supported hypothesis 1 and 4 which indicated that no significant changes took place in the performance appraisal scores of the control group over the six-month period.

Comparison of five hypotheses

As discussed earlier, hypothesis 1 accounted for the main objective of the study, whereas hypotheses 2 to 5 accounted for alternative outcomes. It became evident when discussing the results of these hypotheses, that hypothesis 1 received more statistical support than did any of the other hypotheses. The largest percentage of variance on all six dimensions was explained by hypothesis 1 (refer Table 12). It thus appears as if an improvement in the appraisal scores of the experimental group on all six dimensions included in the BOS, had taken place. The dimension in which the most significant increase had been identified, was Administration (93,9% variance explained). The importance of this aspect had already been indicated. Not only the functioning of the transport department could have been positively

influenced, but also that of the organization as a whole. On the dimensions Supervision (93,7% variance explained), Maintenance of Truck (74,8% variance explained) and Responsibility (72,2% variance explained), a substantial improvement in appraisal scores was also recorded. The appraisal scores on the dimensions Liaison with Customers (41,1% variance explained) and Safety (26,7% variance explained) also improved, but to a lesser extent than those on the other dimensions.

Thus, although other of the hypotheses received some statistical support and explained some proportion of variances between the two groups during the three performance appraisals, the results of hypothesis 1 were far more significant. If Loftus and Loftus' (1982) statement that planned comparisons is a powerful means of assessing the significance of the hypothesis is considered, it can be concluded that the intervention in this study was highly successful.

Use was also made of unobtrusive measures to achieve a qualitative means of determining whether the drivers in the experimental group were functioning on a higher level than before. (The term qualitative refers to the quality of work, although it will sometimes be expressed in numerical terms). At the outset of the study, it was believed that this could indicate some significant patterns, but the frequency with which drivers were issued traffic fines, had accidents and were complained about was so small that this idea was discarded. However, after discussions with super-

visors, it was believed that some changes in the performance of the drivers in the experimental group indeed had taken place. Therefore, the results of these unobtrusive measures, will now be discussed.

Unobtrusive Measures

In many instances the general appearance of drivers as well as in some cases, that of their vehicles, had improved. In discussions with the supervisors of the drivers in the experimental group after the implementation of the new performance appraisal system, the following aspects came to the fore.

Drivers now had clarity on what was expected from them in their jobs. They actually stated this fact quite openly. For the first time, especially after feedback interviews on this aspect, they realized the importance of their own general appearance, as well as that of their crew and the truck. One supervisor reported a specific incident, in which a driver who was usually rather untidy when he came to work, suddenly started to appear at work neatly shaved and well dressed. Another driver who rarely saw to the washing and cleaning of his truck, started washing it regularly.

Drivers also seemed to pay more attention to the appearance of their crews. According to one supervisor, even the crews

of some of the drivers seemed neater in appearance than before. Another driver had problems with the punctuality of his crew. After discussing this in a feedback interview, the driver solved this problem successfully. He now felt, more than ever, responsible for the behaviour of his crew and also felt that his supervisor would support him if he confronted his crew on this matter. Thus, it seems as if drivers were more successful in gaining co-operation from their crews.

One supervisor felt that he received fewer complaints from customers on the behaviour of drivers than before. In the period immediately before the new appraisal system was implemented, he received about four complaints in three months. In the three months after the first feedback interview he received only two complaints and in the following three months only one. The supervisor was of the opinion that this could be attributed to the positive effect of the performance appraisal system, but this may be an oversimplification. Because of a variety of circumstances, such as changing attitudes of customers, it is not possible to ascribe the reduction of complaints directly to the effect of the Behavioural Observation scale developed in this study. Comparisons of other periods also indicated that the frequency of complaints tended to differ from one month to another. The other two supervisors in the experimental group mentioned no significant decrease in complaints on the behaviour of drivers.

Absenteeism on the part of drivers was also noted and studied. No major changes could be found during the study. This had not been a problem in the case of the truck drivers.

Attention was also given to the number of traffic violations reported, but no specific pattern could be identified because of the low frequency of violations. Supervisors also believed that in certain cases drivers did not always report being reprimanded or fined by traffic officers.

Thus, although some of the supervisors of the experimental group were of the opinion that the BOS developed in this study had contributed to fewer customer complaints and better personal appearances of the driver and his crew, no significant conclusion regarding the abovementioned unobtrusive measures could be drawn.

In the remaining part of this study, the reactions of the drivers in the experimental group to the new appraisal system were assessed. Some attention was also given to further changes needed to increase the acceptability of the BOS to an even larger extent. This aspect will now be discussed.

Acceptability of the BOS

After the implementation of the newly developed performance appraisal system, the Performance Appraisal Satisfaction Ques-

tionnaire was again handed out to be completed by the truck drivers in the experimental group. The purpose of this exercise was to determine whether the drivers were more satisfied with the "new" appraisal system than with the "old" one. By again requesting the drivers to indicate their expectations regarding performance appraisal, the weak aspects in the new system were identified for future attention and consideration. At this stage the performance of these drivers had been appraised three times over a six-month period with the newly developed BOS. This implies that drivers had had ample time to familiarize themselves with the functioning thereof.

The following discussion will firstly compare the experiences of the drivers with the "old" and "new" appraisal systems, and secondly the expectations for improvement of the new system will be identified.

Comparison of experiences with the two systems

The experiences of the truck drivers in the experimental group with the "old" and "new" appraisal systems were compared. The same experience scores derived during the first completion of the Performance Appraisal Satisfaction Questionnaire were used. This was believed to be fairly accurate, seeing that these scores were obtained while the "old" appraisal system was still in use. These scores were compared with the scores applying

to the "new" appraisal system. The results obtained from this exercise are indicated in Table 17.

A t-test for dependent samples was done, resulting in $t = 7,63$, $p < 0,01$. This indicated that the drivers in the experimental group were significantly more satisfied with the "new" appraisal system than with the previous one. In other words, these drivers had a more positive orientation towards the newly developed appraisal system compared to the previous one. Drivers also felt more involved in the development of the system. This may be due to the fact that their complaints regarding the previous system were considered in the development of the BOS.

From Table 17, it is evident that the biggest discrepancies recorded were on questions 8, 11, 13 and 14. This indicates that drivers were much more satisfied with the fact that they received feedback on their performance from their supervisors (question 8). It thus appears that drivers welcomed the performance appraisal feedback interviews which were being held after each performance appraisal. This supports Latham and Wexley's (1981) statement, that appraisal interviews should be an inherent aspect of any appraisal system.

It further appears that drivers were very impressed with the fact that plans were made for further development and that goals were set to be attained in the future (questions 11 and

TABLE 17

Discrepancy between the Experiences of the Drivers with the First and Second Performance Appraisals

	EXPERIENCED 1	EXPERIENCED 2	DISCREPANCY
N = 40	\bar{X}_E	\bar{X}_P	$\bar{X}_P - \bar{X}_E$
NUMBER OF QUESTION			
1	1,1	5,3	4,2
2	1,2	3,2	2,0
3	1,3	5,1	3,8
4	1,6	4,7	3,1
5	2,3	5,4	3,1
6	1,0	5,6	4,6
7	1,5	5,8	4,3
8	1,9	6,9	5,0
9	2,0	6,2	4,2
10	1,9	6,3	4,4
11	1,2	6,2	5,0
12	1,9	5,8	3,9
13	1,4	6,9	5,5
14	0,9	6,5	5,6
15	2,6	7,0	4,4
16	1,6	6,1	4,5
	$\bar{X} = 1,6$	$\bar{X} = 5,8$	$\bar{X} = 4,2$

13). Drivers immediately felt more involved with the appraisal process, and with the overall functioning of the transport department. This could lead to increased job satisfaction and an improvement in the functioning of this department.

The biggest improvement in the experiences of drivers with the new appraisal system, was recorded on question 14. This implied that drivers experienced the new system to contribute to better overall performance, which in fact was the case if the results of the planned comparisons as discussed earlier were considered. A significant improvement in appraisal scores of the drivers were recorded, which implied an improvement in the performance of the drivers. This improvement was recorded on all six dimensions of the drivers' job, thus representing an overall increase in performance.

It further appears that relations between the drivers and their supervisors, the transport foremen, had improved since implementing the new system. This created a milieu which could only positively influence the functioning of the transport department. However, it appeared that drivers still needed more time to prepare themselves for feedback interviews (question 2), and that they should be allowed to express more opinions and ideas during the interview (question 4).

To determine which aspects of the newly developed BOS could still be improved, drivers again had to indicate their experiences

of a performance appraisal system, after having been exposed to the functioning of the BOS for six months. The results of this exercise are discussed below.

Experiences versus Expectations of New Performance Appraisal System

At the same time that drivers had to indicate their experiences of the new appraisal system, they were again asked to state what their expectations regarding performance appraisal were. This provided an opportunity for management to reconsider the needs of the drivers regarding appraisal, and to take appropriate action to make the new appraisal system even more acceptable to them. The results obtained from this exercise are presented in Table 18.

It is apparent from these results that the biggest discrepancy between the experiences and expectations of the drivers still existed with regard to the time available for employees to prepare themselves for a performance appraisal (question 2). Especially in the job of truck driver, a great deal of administrative work was involved (see job description - previous chapter). The drivers needed time to refer back to specific documents to keep them informed of any irregularities during the previous performance

TABLE 18

Results obtained from the Second Questionnaire determining the Experiences and Expectations of the Truck Drivers regarding the new Performance Appraisal System

N = 40	EXPERIENCED	EXPECTED	DISCREPANCY
	\bar{X}_E	\bar{X}_{EP}	$\bar{X}_{EP} - \bar{X}_E$
NUMBER OF QUESTION			
1	5,3	9,7	4,4
2	3,2	10,7	7,5
3	5,1	10,8	5,7
4	4,7	10,9	6,2
5	5,4	10,6	5,2
6	5,6	11,0	5,4
7	5,8	10,8	5,0
8	6,9	10,9	4,0
9	6,2	10,6	4,4
10	6,3	11,1	4,8
11	6,2	11,2	5,0
12	5,8	10,6	4,8
13	6,9	11,5	4,6
14	6,5	10,2	3,7
15	7,0	10,3	3,3
16	6,1	10,7	4,6
	$\bar{X} = 5,8$	$\bar{X} = 10,01$	$\bar{X} = 4,8$

appraisal period. This confirms that employees could only fully participate in the feedback interview if they were given sufficient time to prepare themselves for all the problems which may arise during such an interview (Latham & Wexley, 1981; Richardt, 1976).

The lowest discrepancy (question 15) indicated that truck drivers really experienced that their relationship with their supervisors had improved. This could probably be attributed to the fact that the supervisors were trained to create a pleasant atmosphere during the appraisal interview. Counselling and development of employees were emphasized, thus confirming the views of Schuler (1981) and Wells (1982) that this contributes to a positive feeling towards the appraisal interview by the employee.

The most positive experiences with regard to the BOS reported by the truck drivers, were the feedback they received on their performance and the goals that were set for them (questions 8 and 13). This supported the idea that feedback interviews help employees to achieve a clear understanding of how they are viewed by the organization and their supervisors, in this case, the transport foremen (Alewine, 1982; Eichel & Bender, 1984; Graves, 1982; Latham & Wexley, 1981). This positive reaction to the setting of goals is very important, because goal setting (or direction) is a central aspect of any performance appraisal system, as was pointed out earlier (Latham & Wexley, 1981; Smith & Brouwer, 1977).

With regard to expectations, the highest expectations identified by the truck drivers came to the fore in questions 11 and 13. Question 11 specified the extent to which plans were made for future development, whilst question 13 indicated the extent to which goals for truck drivers were set to be reached in the future.

Thus, a specific need was identified by the truck drivers. This need was to be assisted by their supervisors in future planning and career development. This may be attributed to the fact that the truck drivers' job followed the same routine each day, without the driver being able to show much initiative. The truck drivers operated mainly on instructions from their supervisors, and considering the fact that few opportunities existed for the promotion of drivers, the job of truck driver could become boring after a while (regarding promotion, only three jobs existed for transport foremen, while 40 truck drivers are employed, each waiting to be promoted to transport foreman). Therefore a driver may feel that after a while, his job holds no challenges and excitement for him, and thus expects assistance from his supervisor to create a new interest in his job.

In view of the abovementioned results, certain deficiencies and certain positive aspects regarding the performance appraisal system developed, were identified. This leads us to certain conclusions and recommendations, which will be discussed in Chapter 6.

CHAPTER 6

Conclusion

The conclusions derived on the basis of the results presented in the previous chapter, will now be discussed. Afterwards a few recommendations to the organization involved regarding this study will be made.

As indicated in chapter 1, the main objective of this study, was to develop a reliable performance appraisal instrument for truck drivers in the transport department. As previously described, this instrument had to meet certain prescribed requirements. In the first instance, it had to be more acceptable to the drivers than the previous appraisal system. Thus, drivers should feel more satisfied with the "new" performance appraisal instrument than with the "old" one. Secondly, this newly developed instrument also had to be able to provide a means for improving the performance of the drivers.

The system was developed as specified by Birkenbach (1984) and Latham and Wexley (1981), resulting in a reliability coefficient (coefficient alpha) of 0,98. Thus, the instrument proved to be highly reliable, implying that it could be used with great confidence for future performance appraisals of truck drivers in this organization. The instrument also met the requirements for face- and content validity as discussed earlier.

Furthermore, in the development of the new appraisal instrument, complaints and recommendations of drivers and their supervisors in connection with the previous appraisal instrument were considered. Having incorporated some of these recommendations in the new appraisal system, it was found that the satisfaction of the drivers with the "new" system had increased significantly compared to their satisfaction with the previous system. Drivers indicated that they felt much more dedicated to the new system, because they were involved in the development thereof. In discussions with several drivers after the implementation of the new system, it came to light that for the first time certain drivers became aware of exactly what was expected from them on the job.

They now knew on what aspects of their jobs they were being appraised, and the feedback interviews also provided the drivers with a means to query some of their supervisors' decisions, or to discuss their performance in general.

As already indicated, another requirement of the system developed was that it had to be able to improve the performance of the truck drivers. The results obtained in this regard have provided support for the effectiveness of the BOS developed in this study. By performing planned comparisons, a significant increase in the appraisal scores of the drivers in the experimental group was identified from the first to the third performance

appraisal. This implied that the performance of the drivers in this group had improved over a six-month period, whereas little or no change had been recorded in the control group. It thus appears as if the BOS developed in this study, may have contributed to the improvement of the performance of the drivers in the experimental group. By indicating to the drivers in this group exactly which aspects of their performance could be improved, it became possible for them to exhibit behaviour which would achieve this goal. Thus goals relating to the weaker aspects of a driver's performance, were set for him. It was expected that these would be reached before the next appraisal. The drivers indicated in discussions that these directed goals had a positive influence on their performance. They had clarity on which aspects of their performance was not up to standard, which enabled them to put effort into bringing these areas up to the standards expected from them.

Another requirement of the performance appraisal system was to provide a means for drivers to discuss their performance with their supervisors. This was considered to be particularly important because during these appraisal interviews, objectives were set for the drivers to achieve during the next appraisal period and ways in which they could succeed were identified. Drivers could discuss their performance frankly and openly with supervisors and together they could plan for future improvement. The weak aspects of the driver's performance were identified

and discussed and the better aspects verbally rewarded. A positive aspect of the feedback interviews was that according to the drivers, each interview contributed to an improvement in the relationship with their supervisor, thus ensuring a better understanding of each other in the transport department.

In conclusion, it can thus be stated with some degree of confidence that the original goals of the study were achieved, and that the intervention proved to be highly successful. However, some areas still needing attention had been identified. These will now be discussed. Some recommendations for the improvement of these areas will also be made.

Recommendations to the Organization

The BOS was developed for truck drivers in this particular organization. It was found to be highly reliable and it seemed as though most of the requirements set by the organization were met. Therefore, in view of the results achieved by this system, the organization may want to consider similar systems adapted to meet the requirements of other jobs in different departments. Considering the fact that in most instances, thorough job descriptions already existed for other jobs in the organization, the development of a BOS for these jobs may not be as time-consuming as may be believed.

The fact that the performance of the truck drivers had improved, implied that the drivers received higher appraisal scores. This

resulted in most drivers being rated as average (a 3 on the 5-point scale) or above average (4 to 5 on the same scale). In other words, after the completion of this study, the BOS did not discriminate between good and poor performers to the same extent as it did before. A solution to this problem may be to increase the standards according to which drivers were rated during the course of this study. Drivers initially received a score of 1 on the BOS when engaging in a specific behaviour 0-20 per cent of the time, a 2 for 21-40 per cent, a 3 for 41 to 60 per cent, a 4 for 61 to 80 per cent and a 5 for 81-100 per cent (refer chapter 4). Although complicating the task of the rater (as discussed earlier), by altering these different percentage intervals, it is possible to obtain a better discrimination between good and poor performers. In the future, drivers could for example receive a 1 if they behaved in a certain manner 0-64 per cent of the time, a 2 for 65-74 per cent, a 3 for 75-84 per cent, a 4 for 85-94 per cent and a 5 for 95-100 per cent of the time, as specified by Latham and Wexley (1981). By increasing the standards of ratings, a better distribution of performers will be obtained over the scale points of the BOS. This process is the responsibility of the transport department, which has to change the standards of ratings according to the performance of the drivers employed in this department.

It was further found that major needs which will develop in the near future amongst the truck drivers, will be for job enrichment because of the limited career opportunities available. The organization may therefore have to investigate means for

developing managerial skills in these drivers and thus create the possibility of a career path. Of course, job enrichment does not exclude the identification and development of managerial skill. In fact, the two can go hand-in-hand. This will be a major contribution to the development of other race groups in South Africa.

REFERENCES

- Alewine, T.C. (1982). Performance appraisals and performance standards. Personnel Journal, 61(3), 210-213.
- Anastasi, A. (1982). Psychological Testing (5th ed.). New York: MacMillan Publishing Co., Inc.
- Baker, H.K., & Morgan, P.I. (1984). Two goals in every performance appraisal. Personnel Journal, 63(9), 74-78.
- Banks, C.G., & Roberson, L. (1985). Performance appraisers as test developers. Academy of Management Review, 10(1), 128-142.
- Baron, R.A. (1983). Behaviour in Organizations. Massachusetts: Allyn and Bacon, Inc.
- Barret, R.S. (1966). The influence of the supervisor's requirements on ratings. Personnel Psychology, 19(4), 375-387.
- Basset, G.A., & Meyer, H.H. (1968). Performance appraisal based on self-review. Personnel Psychology, 21(4), 421-430.
- Beach, D.S. (1980). Personnel: The Management of People at Work (4th ed.). New York: MacMillan Publishing Co.
- Beatty, R.W., Schneier, S.E. & Beatty, J.R. (1977). An empirical investigation of perceptions of ratee behavior, frequency, and ratee behavior change using behavioral expectation scales. Personnel Psychology, 33(4), 647-658.
- Beer, M., & Ruh, R.D. (1976). Employee growth through performance management. Harvard Business Review, 54(4), 59-66.
- Bernardin, H.J. (1978). Effects of rater training on leniency and halo errors in student ratings of instructors. Journal of Applied Psychology, 63 (3), 301-308.
- Bernardin, H.J., & Walter, C.S. (1977). Effects of rater training and diary-keeping on psychometric error in ratings. Journal of Applied Psychology, 62, 64-69.
- Bernardin, H.J., & Pence, E.G. (1980). The effects of rater training : Creating new response sets and decreasing accuracy. Journal of Applied Psychology, 65 (1), 60-66.
- Biesheuvel, S. (1984). Work Motivation and Compensation : Volume 1. Johannesburg: McGraw-Hill Book Co.
- Birkenbach, X.C. (1984). Halo, central tendency and leniency in performance appraisal: a comparison between a graphic rating

- scale and a behaviourally based measure. Perspektiewe in die Bedryfsielkunde, 10(1), 16-34.
- Birkenbach, X.C., Kamfer, L., & ter Morshuizen, J.D. (1985). The Development and the evaluation of a behaviour-modelling training programme for supervisors. South African Journal of Psychology, 15, 11-19.
- Blake, R.R., & Mouton, J.S. (1961). Power, people and performance reviews. Advanced Management Journal, 26(3), 13-17.
- Borman, W.C. (1975). Effects of instructions to avoid halo error on reliability and validity of performance evaluation ratings. Journal of Applied Psychology, 60, 556-560.
- Borman, W.C. (1979). Format and training effects on rating accuracy and rater errors. Journal of Applied Psychology, 64, 410-421.
- Burack, E.H., & Smith, R.D. (1977). Personnel Management - A Human Resource Systems Approach. New York: West Publishing Company.
- Burke, R.J., Weitzel, W., & Weir, T. (1978). Characteristics of effective employee performance review and development interviews: Replication and extension. Personnel Psychology, 31(3), 903-919.
- Byham, W., & Robinson, J. (1976). Interaction Modelling. A new concept in Supervisory Training. Training and Development Journal, February.
- Campbell, D.T. (1966). Experimental and Quasi-Experimental designs for Research. Chicago: Rand-McNally College Publications.
- Campbell, J.P., Dunnette, M.D., Lawler, E.E. & Weick, K.E. (1970). Managerial behavior, performance and effectiveness. New York: McGraw-Hill Book Co.
- Cascio, W.F. (1982). Applied Psychology in Personnel Management (2nd ed.). New York: Reston Publishing Company.
- Cascio, W.F. (1982). Scientific, legal and operational imperatives of workable performance appraisal systems. Public Personnel Management, 11(4), 367-375.
- Clayton, K., & Gatewood, R. (1981). The development of a managerial performance appraisal system in a social service agency. Public Personnel Management Journal, 10(2), 261-269.
- Cronbach, L.J., & Furby, L. (1970). How we should measure "change" - or should we? Psychology Bulletin, 74 (1), 68-80.

- Cummings, L.L., & Schwab, D.P. (1973). Performance in Organizations: Determinants and Appraisal. New York: Scott, Foresman and Company.
- Dailey, L.A., & Madsen, A.M. (1980). How to evaluate people in Business. New York: McGraw-Hill Book Company.
- DeCotiis, T., & Petit, A. (1978). The performance appraisal process : A model and some testable propositions. Academy of Management Review, 3, 635-646.
- De Jung, J.E., & Kaplan, H. (1962). Some differential effects of race of rater and ratee on early peer ratings of combat attitude. Journal of Applied Psychology, 46, 370-374.
- Dessler, G. (1983). Improving Productivity at work. Virginia: Reston Publishing Company, Inc.
- Dosset, D.L., Latham, G.P. & Mitchell, T.R. (1979). The effects of assigned versus participatively set goals, KR, and individual differences when goal difficulty is held constant. Journal of Applied Psychology 64, 291-298.
- Dunnette, M.D. (1966). Personnel Selection and Placement. Belmont: Brooks Cole.
- Eichel, E., & Bender, H.E. (1984). Performance Appraisal : A study of current Techniques. New York: AMA Research and Information Service.
- Fletcher, C.A. (1973). Interview style and the effectiveness of appraisal. Occupational Psychology, 47(3-4), 225-230.
- Flippo, E.B. (1980). Personnel Management. New York: McGraw-Hill Book Co.
- Freedman, J.L., Sears, D.O., & Carlsmith, J.M. (1978). Social Psychology (3rd ed.). New Jersey: Prentice-Hall Inc.
- French, W. (1970). The Personnel Management Process : Human Resources Administration (2nd ed.). Boston: Houghton Mifflin Company.
- Ghiselli, E.E., Campbell, J.P. & Zedeck, S. (1981). Measurement Theory for the Behavioral Sciences. San Francisco: W.H. Freeman and Company.
- Glueck, W.F. (1979). Foundations of Personnel. Dallas: Business Publications.
- Gordon, L.V. & Medland, F.F. (1965). The cross-group stability of peer ratings of leadership potential. Personnel Psychology, 28(2), 173-177.

- Graves, P.J. (1982). Let's put appraisal back in performance appraisal: Part I. Personnel Journal, 16(11), 844-849.
- Graves, P.J. (1982). Let's put appraisal back in performance appraisal: Part II. Personnel Journal, 16(12), 918-923.
- Guion, R.M. (1965). Personnel Testing. New York: McGraw-Hill Book Company.
- Guion, R.M. (1980). On trinitarian doctrines of validity. Professional Psychology 11(3), 384-397.
- Halloran, J. (1981). Supervision - the Art of Management. New Jersey: Prentice-Hall Inc.
- Haynes, M.G. (1978). Developing an appraisal programme - Part I. Personnel Journal, 57(1), 14-19.
- Hays, W.L. (1973). Statistics for the social sciences (2nd ed.). New York: Holt, Rinehart and Winston, Inc.
- Henderson, R.I. (1984). Performance Appraisal (2nd ed.). Virginia: Reston Publishing Co., Inc.
- Heneman, H.G., (1974). Comparisons of self- and superior ratings of managerial performance. Journal of Applied Psychology, 59 (5), 638-642.
- Hyde, A.C., & Smith, M.A. (1982). Performance appraisal and training : Objectives, a model for change and a note of rebuttal. Public Personnel Management, 11(4), 358-366.
- Ivancevich, J.M. & Glueck, W.F. (1983). Foundations of Personnel/Human Resource Management (Revised ed.). Plano: Business Publications, Inc.
- Jacobs, R. & Kozlowski, S.W.J. (1985). A closer look at halo error in performance ratings. Academy of Management Journal, 28(1), 201-212.
- Jenkins, G.D., & Taber, T.A. (1977). A Monte Carlo study of factors affecting three indices of composite scale reliability. Journal of Applied Psychology, 62, 392-398.
- Johns, G. (1981). Difference Score Measures of Organizational Behavior Variables : A Critique. Organizational Behavior and Human Performance, 27 (3), 443-463.
- Kane, J.S., & Bernardin, H.J. (1982). Behavioral observation scales and the evaluation of performance appraisal effectiveness. Personnel Psychology, 35(3), 635-641.

- Kavanagh, M.J., MacKinney, A.C. & Wolins, L. (1971). Issues in managerial performance: Multitrait-multimethod analysis of ratings. Psychological Bulletin, 75(1), 34-49.
- Kellog, M.S. (1965). What to do about Performance Appraisal. New York: The Haddon Craftsmen Inc.
- Kirchner, W.K. (1965). Relationships between supervisors and subordinate ratings for technical personnel. Journal of Industrial Psychology, 3, 57-60.
- Kirkpatrick, D.L. (1984). Two ways to evaluate your performance appraisal system. Training and Development Journal, 38(8), 38-40.
- Klatt, L.A., Murdick, R.G., & Schuster, F.E. (1978). Human Resource Management - A Behavioral Systems Approach. Illinois: Richard D. Irwin, Inc.
- Korman, A.K. (1968). The prediction of managerial performance: A review. Personnel Psychology, 21(3), 295-322.
- Landy, F.J. (1985). Psychology of Work Behavior (3rd ed.). Illinois: The Dorsey Press.
- Landy, F.J., & Farr, J.L. (1983). The Measurement of Work Performance. New York: Academic Press.
- Landy, F., Zedeck, S. & Cleveland, J. (1983). Performance Measurement and Theory. London: Lawrence Erlbaum Associates.
- Latham, G.P., Mitchell, T.R., & Dosset, D.L. (1978) The importance of participative goal-setting and anticipated rewards on goal difficulty and job performance. Journal of Applied Psychology, 63, 163-171.
- Latham, G.P., & Saari, L.M. (1979). The Importance of supportive relationships in goal setting. Journal of Applied Psychology, 64, (2) 151-156.
- Latham, G.P., & Wexley, K.N. (1977). Behavioral observation scales for performance appraisal purposes. Personnel Psychology, 30(2), 355-268.
- Latham, G.P., & Wexley, K.N. (1981). Increasing Productivity through Performance Appraisal. Massachusetts: Addison-Wesley Publishing Company.
- Latham, G.P., Wexley, K.N., & Purcell, E.D. (1975). Training managers to minimize rating errors in the observation of behaviour. Journal of Applied Psychology, 60, 550-555.
- Latham, G.P., & Yukl, G.A. (1975). A review of research on the application of goal setting in organizations. Academy of Management Journal, 18, 824-845.

- Lawler, E.E., Mohrman, A.M., & Resnick, S.M. (1984). Performance appraisal revisited. Organizational Dynamics, 13(1), 20-35.
- Lemke, E., & Wiersma, W. (1976). Principles of Psychological Measurement. Boston: Houghton Mufflin Company.
- Levinson, H. (1976). Thinking ahead. Harvard Business Review, 54(4), 30-46.
- Locke, E.A. (1968). Toward a theory of task motivation and incentives. Organizational Behaviour and Human Performance, 3, 157-189.
- Loftus, G.R. & Loftus, E.F. (1982). Essence of Statistics. Monterey: Brooks/Cole Publishing Company.
- Lopez, F.M. (1975). Personnel Interviewing. New York: McGraw-Hill Book Co.
- Luthans, F. (1977). Organizational Behaviour. Tokyo: McGraw-Hill Kogakusha, Ltd.
- Mathis, R.L. & Jackson, J.H. (1982). Personnel Contemporary perspectives and applications (3rd ed.). St. Paul: West Publishing Co.
- McClelland, D.C. (1961). The Achieving Society. New Jersey: D. van Nostrand Company, Inc.
- Miles, R.E., & Snow, C.C. (1984). Designing strategic human resources systems. Organizational Dynamics, 13(1), 36-52.
- Miller, S. (1975). Experimental Design and Statistics. London: Methuen & Co. Ltd.
- Nie, N.H. & Hadlai Hull, C. (1981). SPSS-UPDATE 7-9. New Procedures and Facilities for releases 7-9. New York: McGraw-Hill Book Co.
- Odiorne, G.S. (1979). MBO II. Belmont: Fearon Pittman Publishers, Inc.
- Pagano, R.P. (1981). Understanding Statistics in the behavioral sciences. New York: West Publishing Company.
- Richardt, A.J. (1976). Performance Appraisal Systems and practices in different South African organizations. Master's Thesis, UNISA.
- Robbins, S.P. (1982). The Management of Human Resources (2nd ed.). Englewood Cliffs: Prentice Hall.
- Sausser, W.I. (1980). Evaluating employee performance needs, problems and possible solutions. Public Personnel Management, 9(1), 12-18.

- Schneier, C.E. (1977). Multiple rater groups and performance appraisal. Public Personnel Management, 6(1), 13-20.
- Schuler, R.S. (1981). Personnel and Human Resource Management. Minnesota: West Publishing Company.
- Schwab, D.P., Heneman, H.G., & DeCotiis, T.A. (1975). Behaviorally anchored rating scales: A review of the literature. Personnel Psychology, 28, 549-562.
- Siegel, J. (1980). Personnel Testing under EEO. New York: AMACOM.
- Slivinski, R.W. (1975). Performance Appraisal: What is required to make it work. Unpublished Manuscript.
- Smith, H.P., & Brouwer, P.J. (1977). Performance Appraisal and Human Development. Reading: Addison-Wesley Publishing Company.
- Spool, M.D. (1978). Training programmes for observers of behavior: A review. Personnel Psychology, 31(4), 853-888.
- Strauss, G., & Sayles, L.R. (1980). Human Behavior in Organizations. Englewood-Cliffs: Prentice Hall, Inc.
- Teel, K.S. (1980). Performance Appraisal: Current trends, persistent progress. Personnel Journal, 59(4), 296-301; 316.
- Thornton, G.C. (1968). The relationship between supervisor-and self-appraisals of executive performance, Personnel Psychology, 21 (4), 441-455.
- Thornton, G.C. (1980). Psychometric properties of self-appraisals of job performance. Personnel Psychology, 33 (1), 236-271.
- Van Wyk, J. (1983). Performance Appraisal Satisfaction Questionnaire. Unpublished manuscript.
- Wanous, J.P. (1973). Effects of realistic job previews on job acceptance, job attitudes and job survival. Journal of Applied Psychology, 58, 327-332.
- Wells, R.G. (1982). Guidelines for effective and defensible performance appraisal systems. Personnel Journal, 61(10), 776-782.
- Wexley, K.N., Sanders, R.E., & Yukl, G.A. (1973). Training interviewers to eliminate contrast effects in employment interviews. Journal of Applied Psychology, 50, 233-236.
- Wexley, K.N., Singh, J.P., & Yukl, G.A. (1973). Subordinate personality as a moderator of the effects of participation in three types of appraisal interviews. Journal of Applied Psychology, 58 (1), 54-59.

White, D.H. (1982). The effects of job stress in the South African Gold Mining Industry. Unpublished doctoral dissertation, UNISA, Pretoria.

Wikstrom, W.S. (1964). Developing managerial competence: Changing concepts, emerging practices. Journal of Applied Psychology 47(1), 18-27.

Zippo, M., & Miller, M. (1984). Performance appraisal - Current practices and techniques. Personnel, 61(3), 57-59.

APPENDIX 1

**PERFORMANCE APPRAISAL
BEOORDELING VAN WERKVERRICHTING**

Confidential/Vertroulik

1. Particulars/Besonderhede

Name/Naam:		Date/Datum:	
Job title/Posbenaming:		Department/Departement:	
Job grade/Posgraad:	Progress increment/ Vorderings inkrement:	<input type="checkbox"/>	Incentive increment/ Aansporings inkrement:
Name and position of rater/ Naam en pos van beoordelaar:			
Name and position of reviewer/ Naam en pos van nasiener:			
Total points/ Totale punte:	Present Salary/Huidige salaris:	
	Merit Increment/Meriete-inkrement:	
% increase/ % verhoging:	Adjustment/Aanpassing:	
	New Salary/Nuwe Salaris:	

2. Directions/Aanwysings

- | | |
|--|---|
| <p>2.1 Each item according to which the performance of the employee should be appraised, is defined. For each of these items, 3 categories ranging from below to above the normal requirements of the job, are also defined. Read these carefully.</p> | <p>Elke item waarvolgens die werknemer se werkverrigting beoordeel moet word, is omskryf. Vir elk van hierdie items is 3 kategorië omskryf wat strek vanaf onder tot bo die normale vereistes van die pos. Lees dit sorgvuldig.</p> |
| <p>2.2 Firstly decide in which broad category the employee falls, then decide whether he should be low, high or in the middle of this category and make a cross (X) in the appropriate block.</p> | <p>Besluit eerstens in watter breë kategorie die werknemer val. Besluit dan of hy laag, hoog of in die middel van hierdie kategorie is en maak 'n kruisie (X) in die toepaslike blokkie.</p> |
| <p>2.3 Rate on full period under review, rather than on recent specific incidents.</p> | <p>Beoordeel t.o.v. die volle tydperk watter sake is, eerder as onlangse spesiesidente.</p> |
| <p>2.4 Ratings should be done by the employee's immediate supervisor. These should then be discussed with and reviewed by the first rater's superior.</p> | <p>Beoordelings moet deur die werknemer se onmiddellike hoof gedoen word. Daarna moet dit met die eerste beoordelaar se hoof bespreek, en deur hom hersien word.</p> |
| <p>2.5 Employees with no supervisory responsibilities should only be rated on the first 4 factors.</p> | <p>Werknemers met geen toesighoudende verantwoordelikheid, moet slegs volgens die die eerste 4 faktore beoordeel word.</p> |

Performance Appraisal Scale

Skaal vir Beoordeling van Werkverrigting

1. Output

Volume van werk

Consider the volume of work (of an acceptable standard) relative to the employee's experience in the job. Also consider the employee's good sense in distributing his effort properly over the various tasks assigned to him, and in recognising priorities.

Dink aan die volume werk (van 'n aanvaarbare standaard) relatief tot die werknemer se ondervinding in die pos. Dink ook aan sy gesonde oordeel om sy kragte behoorlik te verdeel tussen take wat aan hom opgedra is, met inagnaming van prioriteite.

His work output, considered overall, does not generally maintain an acceptable level: or if some deadlines are met, this may be at the expense of other more significant tasks	1 2 3	Die volume van werk wat hy verrig, handhaaf normaalweg nie 'n aanvaarbare peil nie: of indien sommige take wel betyds afgehandel word, kan dit ten ten koste van ander, meer belangrike take wees.
Output generally acceptable, relative to the reasonable demands of the job. Normally maintains a proper balance between the various tasks that require his attention	4 5 6	Volume van werk oor die algemeen aanvaarbaar in terme van die redelike vereistes van die pos. Handhaaf gewoonlik 'n deeglike balans tussen die onderskeie take wat sy aandag vereis
Consistently maintains an output above normal job requirements. Displays a good sense of priorities	7 8 9	Handhaaf deurgaans 'n volume van werk wat hoër is as wat normaalweg in die pos vereis word. Toon 'n gesonde sin vir prioriteite

2. Quality of performance

Kwaliteit van werkverrigting

Consider the quality of his performance generally, and more specifically the skillfulness, originality, thoroughness, accuracy, neatness, reliability, and overall competence of performance. If there are other specific criteria applicable, please apply these in your assessment.

Dink aan die algemene gehalte van sy werk en meer spesifiek aan die vaardigheid, oorspronklikheid, betroubaarheid, deeglikheid, akkuraatheid en altemene bekwaamheid. Indien daar enige verdere kriteria ter sake mag wees, moet dit ook by hierdie beoordeling in ag geneem word

Quality of work is generally not of an acceptable standard. Quality of work is not predictable and tends to fall below standard when under pressure	1 2 3	Gehalte van werk is oor die algemeen nie van 'n aanvaarbare standard nie. Kwaliteit van werk is onvoorspelbaar en neem verder af wanneer onder druk.
Quality of work is usually of an acceptable standard under normal pressure.	4 5 6	Gehalte van werk is gewoonlik van 'n aanvaarbare standaard wanneer werksdruk normaal is.
Quality of work is consistently of a high standard relative to what is normally expected in the job. High quality is maintained even under pressure	7 8 9	Gehalte van werk is voortdurend van 'n hoër standaard in terme van die normale vereistes van die pos. Hoër kwaliteit word selfs onder druk gehandhaaf.

Sub total to be carried forward

Subtotaal wat oorgedra moet word

Sub total brought forward

Subtotaal oorgedra

3. Co-operation with others

Samewerking met andere

Consider the ability and willingness to work with equals and with superiors or with other units in order to reach a common goal, regardless of personal inclination.

Dink aan die vermoë en bereidwilligheid om met medewerkers, hoofde en ander eenhede saam te werk ten einde 'n gemeenskaplike doel te bereik sonder om persoonlike gevoel in ag te neem

Cannot be unconditionally relied upon to make a team effort. Alternatively may co-operate but creates friction to the detriment of the common goal	1	Kan nie onvoorwaardelik op gereken word om 'n spanpoging te ondersteun nie. Daarteen mag hy moontlik saamwerk maar ook wrywing skep wat bereiking van die gemeenskaplike doel benadeel
	2	
	3	
Co-operation with equals, superiors or with other units is usually satisfactory	4	Samewerking met medewerkers, hoofde en ander eenhede is gewoonlik bevredigend
	5	
	6	
Always goes out of his way to be co-operate regardless of personal inclination. Can be relied upon as an effective team member.	7	Doen moeite om samewerking te verleen afgesien van enige persoonlike gevoel. Kan op gereken word om 'n effektiewe spanlid te wees.
	8	
	9	

4. Personal involvement

Persoonlike Betrokkenheid

Consider the employee's attitude towards his job and its effective performance: Consider his willingness and ability to contribute something of his own to the job

Dink aan die werknemer se houding teenoor sy pos en die doeltreffende uitvoering daarvan. Dink aan sy bereidwilligheid en vermoë om uit sy eie 'n bydrae tot sy werk maak

Does a daily routine, but is disinclined to be inconvenienced by the job	1	Verrig 'n daaglike roetine maar is ongeïnteresseerd om enige ongerief as gevolg van sy werk uit te staan.
	2	
	3	
Displays a reasonable interest in the job and its effective performance	4	Openbaar 'n redelike belangstelling in sy pos en doeltreffende uitvoering daarvan
	5	
	6	
Keen and enthusiastic all the time about his job. Spontaneously thinks of ways in which the job can be done more effectively	7	Altyd entoesiasties omtrent sy werk. Dink uit sy eie aan metodes waarvolgens sy werk meer effektief verrig kan word
	8	
	9	

Sub total to be carried forward

Subtotaal wat oorgedra moet word

Sub total brought forward

Subtotaal oorgedra

5. Man Management

Bestuur van mense

Consider his effectiveness in managing the staff under his control. How effective is he in achieving the objectives of his job through the people under his control both on the long and the short term?

Dink aan sy doeltreffendheid om mense onder sy toesig te bestuur. Hoe doeltreffend is hy in die bereiking van die oogmerke van sy pos deur middel van die mense onder sy beheer, beide op die kort en langtermyn

Cannot be regarded as a particularly good manager of people. Not effective in motivating others and gaining their co-operation	1	Kan nie beskou word as 'n goeie bestuurder van mense nie. Nie suksesvol om andere te motiveer en/of hulle samewerking te verkry nie
	2	
	3	
Manages his staff well and gets good co-operation. Has no undue measure of staff problems	4	Bestuur sy mense goed en verkry goeie samewerking. Ondervind nie 'n onrealistiese mate van personeelprobleme nie.
	5	
	6	
Displays the ability to motivate all manner of people and to get them to work together as an effective team. Very successful in creating personal involvement and job orientated team spirit	7	Openbaar die vermoë om alle tipes mense te motiveer en in 'n effektiewe span te snoer. Baie suksesvol daarin om persoonlike betrokkenheid en 'n werkgeoriënteerde span te skep
	8	
	9	

9. Development of staff

Ontwikkeling van personeel

Consider his effectiveness in improving the performance of his staff through conscious training and development

Dink aan sy doeltreffendheid ten opsigte van die verbetering van sy personeel se werkverrigting deur middel van doelbewuste opleiding en ontwikkeling

Does little or nothing to improve the performance of his staff through training and development	1	Doen wynig of niks om sy personeel se werkverrigting te verbeter deur middel van opleiding of ontwikkeling nie
	2	
	3	
Is conscious of his responsibility to train and develop his staff and is reasonably effective at it.	4	Is bewus van sy verantwoordelikheid om sy personeel op te lei en te ontwikkel en is redelik doeltreffend in hierdie opsig
	5	
	6	
Looks upon staff development and training as one of his major responsibilities. Is enthusiastic about training and development and becomes actively involved in it	7	Beskou personeelontwikkeling en opleiding as een van sy belangrikste verantwoordelikhede. Is entoesiasies daaromtrent en raak aktief daarby betrokke.
	8	
	9	

Total

Totaal

If items 5 and 6 are not completed, multiply sub-total from previous page by 1,5 to obtain final score. If items 5 and 6 were completed, add totals for items 1-6 for final score.

Indien items 5 en 6 nie voltooi is nie, vermenigvuldig die subtotaal onder aan die vorige bladsy met 1,5 vir die finale telling. Indien items 5 en 6 wel voltooi is, tel die totale vir items 1-6 bymekaar vir die finale telling

FINAL SCORE

FINALE TELLING

APPENDIX 2

PERFORMANCE APPRAISAL SATISFACTION QUESTIONNAIRE

Evaluation of Performance Appraisal Practices

This questionnaire is part of a project which is being undertaken to learn more about how performance appraisals are actually done in the organization. The aim is to use the information to identify problems and to find better ways of appraising the performance of people who work here.

If this study is to be helpful it is important that you answer each question as thoughtfully and frankly as possible. This is not a test and there are no right or wrong answers.

A large number of employees have been invited to complete questionnaires. The completed questionnaires will be processed in such a way that the answers for the whole group are summarised in statistical form. This means that no individual's answers can be identified. to ensure COMPLETE CONFIDENTIALITY, please do not write your name anywhere on the questionnaire.

INSTRUCTIONS

1. Most questions can be answered by marking the number which best reflects your view. If you do not find the exact answer which fits your case, use the one that is closest to it.
2. Please answer all the questions.
3. Remember, the value of the study depends upon your being straight-forward in answering the questionnaire. You will not be identified with your answers.
4. Please note that 2 answers are required to each question. Please CIRCLE the number which describes best what you ACTUALLY EXPERIENCED. Place an "X" on the number which describes best how you WOULD HAVE LIKED IT TO BE.

THE QUESTIONNAIRE

1. Please indicate the extent to which, in your experience, performance appraisal as it is practised in the company meets the following requirements (indicate your answer with a CIRCLE).
2. Next, consider the extent to which you would prefer performance appraisal to meet each requirement and indicate your answer by placing an "X" over the number of your choice.

THE EXTENT TO WHICH:

	don't know	to a little extent	to some extent	to an adequate extent	to a large extent
1. You and your senior reached PRIOR AGREEMENT about the factors against which your performance would be measured eventually	0	1 2 3	4 5 6	7 8 9	10 11 12
2. You were given sufficient time in order to prepare yourself for the yearly performance appraisal discussion with your senior	0	1 2 3	4 5 6	7 8 9	10 11 12
3. The factors against which your performance was evaluated were of real importance in performing your job	0	1 2 3	4 5 6	7 8 9	10 11 12
4. You had the opportunity to express your own views during the annual performance appraisal discussion	0	1 2 3	4 5 6	7 8 9	10 11 12
5. Your performance was evaluated objectively	0	1 2 3	4 5 6	7 8 9	10 11 12
6. Your performance was eventually measured against those factors agreed upon previously between yourself and your senior	0	1 2 3	4 5 6	7 8 9	10 11 12

THE EXTENT TO WHICH:

	don't know	to a little extent	to some extent	to an adequate extent	to a large extent
7. The discussion was frank and open	0	1 2 3	4 5 6	7 8 9	10 11 12
8. You were informed of your senior's assessment of your performance	0	1 2 3	4 5 6	7 8 9	10 11 12
9. You and your senior reached agreement about the positive aspects of your performance	0	1 2 3	4 5 6	7 8 9	10 11 12
10. You and your senior reached agreement about aspects of your performance which could be further improved	0	1 2 3	4 5 6	7 8 9	10 11 12
11. You and your senior made plans for your further development	0	1 2 3	4 5 6	7 8 9	10 11 12
12. Plans for your further development were implemented	0	1 2 3	4 5 6	7 8 9	10 11 12
13. New objectives were set for you	0	1 2 3	4 5 6	7 8 9	10 11 12
14. The way in which performance appraisal is being handled at present had a positive influence on your performance	0	1 2 3	4 5 6	7 8 9	10 11 12
15. The discussion had a positive influence on the relationship between yourself and your senior	0	1 2 3	4 5 6	7 8 9	10 11 12
16. Performance appraisal, on the whole, is practised effectively in the organization.	0	1 2 3	4 5 6	7 8 9	10 11 12

17. How often, within the course of a year, would you prefer that your performance be reviewed in discussion with your senior? Please indicate your choice in the appropriate space below with an "X":

- (a) Monthly
 - (b) Three-monthly
 - (c) Six-monthly
 - (d) Annually
 - (e) Other (please specify briefly)
-

APPENDIX 3

List of Items compiled from the Critical Incidents

1. Ensure that all labourers are present at the start of the day.
2. Ensure the safe parking of the truck when delivering goods to a customer.
3. Liaise with customer in terms of when to start unloading, in what order and where the liquor must be unloaded.
4. Maintain good relations with his labourers.
5. Represent labourers to foreman on personal matters.
6. Handle customer complaints and explain reasons for delay in delivery.
7. Maintain good customer relations.
8. Ensure that customer signs invoice on receipt of goods.
9. Supervise the loading of empty containers and bottles and ensure that no foreign and chipped bottles are included.
10. Make daily entries on the monthly operations sheet.
11. Ensure regular cleaning and washing of truck.
12. Reconcile cash received from customer with foreman on return to depot.
13. Handle road accidents as per prescribed procedure.
14. Ensure that cargo is off-loaded according to invoice.
15. Do predriving inspection on truck and coupling of trailer.
16. Maintain time schedules and try to improve delivery times.
17. Give attention to the safety of his labourers.
18. Adhere to verbal and written instructions from transport foreman and fleet co-ordinator.
19. In case of unloaded trailer, supervise loading of trailer, report trailer number and deliver invoices to operations room.

20. Ensure the safe loading and unloading of the cargo and that the canvas and ropes holding the cargo are secure.
21. At the end of the working day, store trailer and truck at predetermined parking areas.
22. Ensure that all documentation is correct and handed in per deadline.
23. Supervise all loading and unloading of cargo.
24. Ensure that labourers are clean, wear company overalls and are presentable in general.
25. When driving, adhere to speed limits and general traffic regulations.
26. Notify operations room by radio of departure from transport bay and when arriving and departing from customer.
27. Is trustworthy as far as receiving COD payments from customers and giving credit for empty containers and bottles is concerned.
28. Ensure the presentability and quality of products loaded and delivered.
29. Liaise with customer to ensure that cash or bank guaranteed cheques are received before unloading.
30. Complete tachochart and first section of log sheet before departure.
31. Fill in driver log sheet on arrival and departure from every customer.
32. In the case of mechanical problems with the truck, get a job requisition from transport foreman and take the truck to the garage.
33. Complete receipt of amount of containers and empty bottles received from the customer and at the end of the day complete empty return slip of total amount received.
34. Hand in log sheet and tachograph to transport foreman and other documents to relevant supervisors.
35. Check cash payment from customer and calculate credit for customer in terms of breakages and empty containers and bottles correctly.
36. Overall inspection of vehicle every day, giving attention to brakes, tyre pressure, refueling, oil, et cetera.
37. On return to depot, ensure that empties are off-loaded at Dry Goods and the full containers returned to the Warehouse.

APPENDIX 4

VEHICLE CHECK LIST - DAILY

MAKE..... REGISTRATION NO. FLEET NO.

SPEEDOMETER IN ORDER FAULTY

BEFORE STARTING ENGINE CHECK:

AFTER STARTING CHECK:

Vertical column of 12 empty boxes for 'BEFORE STARTING ENGINE CHECK' items.

- ENGINE OIL LEVEL
- OIL LEAKS
- WATER LEVEL
- WATER LEAKS
- FUEL LEVEL
- FUEL LEAKS
- FAN BELTS
- BRAKE FLUID LEVEL
- FAN BELTS
- BRAKE FLUID LEVEL
- CLUTCH FLUID LEVEL

Vertical column of 15 empty boxes for 'AFTER STARTING CHECK' items.

- HANDBRAKE ACTION
- FOOTBRAKE ACTION
- CLUTCH PEDAL PLAY
- HOOTER
- WIPERS
- WINDSCREEN WASHERS
- WARNING LIGHTS
- HEAT AND CHARGING LIGHTS
- HEADLAMPS
- INDICATORS
- TAIL LAMPS
- GAUGES, FUEL, HEAT
- AIR PRESSURE
- DRAIN AIR RESERVOIR
- INSPECT BODYWORK
- TRIANGLES, PERMITS & DOCUMENTS

- TYRES - CHECK WEAR
- PRESSURE
- VALVE CAPS

Vertical column of 6 empty boxes for 'TYRES' items.

- CLEAN: WINDOWS
- MIRRORS
- REFLECTORS
- CHEVRONS

REMARKS
.....
.....
.....

Checked By

PLEASE NOTE: Any damage to the vehicle, howsoever caused, must be reported immediately.

REG. NO.

DRIVER

DATE

AS FROM THE ABOVE DATE, THE VEHICLE MENTIONED IN THE LEFT TOP CORNER WILL BE ALLOCATED TO YOU, TO BE KEPT IN A GOOD AND CLEAN CONDITION.

ALL FAULTS MUST BE REPORTED TO THE GARAGE WITHOUT DELAY.

YOU MUST SATISFY YOURSELF THAT THE DAMAGED SECTIONS ARE MARKED CORRECTLY BEFORE SIGNING THIS LETTER.

SHOULD ANOTHER DRIVER USE THE VEHICLE THAT WAS ALLOCATED TO YOU AT ANY TIME, THEN IT WILL BE YOUR RESPONSIBILITY TO CHECK AND REPORT TO A SENIOR IN CHARGE IF DAMAGE WAS CAUSED.

SIGNATURE

APPENDIX 5

Scale of Importance

The scale below is to rate the importance of every individual task in your job. When using the scale consider the time spent on the task, the frequency of performing the task and lastly, the importance of the task in terms of overall performance.

<u>RATING</u>	<u>IMPORTANCE</u>
1	Definitely not part of my position.
2	Occasionally a minor part of my position.
3	Routinely a part of my position.
4	A very important part of my position.
5	A critical part of my position, i.e. has a make or break effect on my performance.

APPENDIX 6

An Interaction Management Module: Allocation of Time

Administrator announces the interaction skill being considered and the participants read an overview of the interaction skill	5 min.
Administrator describes critical steps in handling the interaction ..	7 min.
Administrator shows a motion-picture film or video tape of a supervisor effectively handling the interaction with an employee	8 min.
Administrator and participants discuss how the critical steps were handled in the film (or tape) and discuss on-the-job situations where similar situations occur and where the critical steps could be applied	5 min.
Three participants take turns in skill practice sessions by role playing supervisors with roles provided. Their behaviour in handling interaction situations is observed by the other participants and the administrator using specially prepared Observer Guides. The use of positive reinforcement by the observers helps to build confidence and skill in the role-playing supervisors	50 min.
Participants write their own interaction situations based on job-related problems, using forms provided in the workbook	10 min.
Participants take turns in skill-practice sessions by role-playing the employee in the participant-written situations, while other participants role play supervisors using the interaction skills. The skill practice sessions are observed and discussed	55 min.
Participants read a summary of the ideas covered. Using specially designed forms, they plan on-the-job applications of the interaction skills. The administrator hands out a "Critical Steps" card for participants to utilize on the job	10 min.

(Adapted from Byham & Robinson, 1976)

APPENDIX 7

The Behavioural Observation Scale

PERFORMANCE APPRAISAL
BEOORDELING VAN WERKVERRIGTING

Particulars/Besonderhede

Name/Naam:	Job grade/ Posgraad:	Date/ Datum:
Job title/Posbenaming:	Department/Departement:	
Name and position of rater/ Naam en pos van beoordelaar:		
Total points/ Totale punte:		

TO WHAT EXTENT DOES THE EMPLOYEE PERFORM THE FOLLOWING TASKS :

1. RESPONSIBILITY

1.1 Adhere to verbal and written instructions from Transport Foreman and Fleet Co-ordinator.

Almost never |-----| Almost always
1 2 3 4 5

1.2 In case of unloaded trailer, supervise loading of trailer, report trailer number and deliver invoices to operations room.

Almost never |-----| Almost always
1 2 3 4 5

1.3 Notify operations room by radio of departure from transport bay and when arriving and departing from customer.

Almost never |-----| Almost always
1 2 3 4 5

1.4 Ensure that cargo is off-loaded according to invoice.

Almost never |-----| Almost always
1 2 3 4 5

1.5 Is trustworthy as far as receiving COD payments from customers and giving credit for empty containers and bottles is concerned.

Almost never |-----| Almost always
1 2 3 4 5

1.6 Reconcile cash received from customer with Foreman on return to depot.

Almost never |-----| Almost always
1 2 3 4 5

1.7 Hand in log sheet and tachograph to Transport Foreman and other documents to relevant supervisors.

Almost never |-----| Almost always
1 2 3 4 5

1.8 Handle road accidents as per prescribed procedure.

Almost never |-----| Almost always
1 2 3 4 5

1.9 At the end of the working day, store trailer and truck at predetermined parking areas.

Almost never |-----| Almost always
1 2 3 4 5

1.10 /

1.10 Maintain time schedules and try to improve delivery times.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

2. SAFETY

2.1 Do predriving inspection on truck and coupling of trailer.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

2.2 Give attention to the safety of labourers.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

2.3 Ensure the safe loading and unloading of the cargo and that the canvas and ropes are secure.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

2.4 When driving adhere to speed limits and general traffic regulations.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

2.5 Ensure the safe parking of truck when delivering goods to customer.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

3. SUPERVISION

3.1 Ensure that all labourers are present at the start of the day and if not, approach the leading hand for a substitute.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

3.2 Ensure that labourers are clean, wear company overalls and are presentable in general.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

3.3 Supervise all loading and unloading of cargo.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

3.4 /

3.4 Ensure the presentability and quality of products loaded and delivered.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

3.5 Supervise the loading of empty containers and bottles and ensure that no foreign and chipped bottles are included.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

3.6 On return to depot ensure that empties are off-loaded at Dry Goods and the full containers returned to the Warehouse.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

3.7 Represent labourers to Foreman on personal matters.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

3.8 Maintain good relations with his labourers.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

4. LIAISON WITH CUSTOMERS

4.1 Liaise with customer to ensure that cash or bank guaranteed cheques are received before unloading.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

4.2 Liaise with customer in terms of when to start unloading, in what order and where the liquor must be unloaded.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

4.3 Ensure that customer signs invoice on receipt of goods.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

4.4 Handle customer complaints and explain reasons for delays in delivery.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

4.5 Maintain good customer relations.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

5. MAINTENANCE OF TRUCK

5.1 Overall inspection of vehicle every day giving attention to brakes, tyre pressure, fuel, oil, etc.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

5.2 In the case of mechanical problems with the truck, get job requisition from Transport Foreman and take truck to the garage.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

5.3 Ensure the regular cleaning and washing of truck.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

6. PERFORM ADMINISTRATION

6.1 Complete tachochart and first section of log sheet before departure. P

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

6.2 Fill in driver log sheet on arrival and departure from every customer.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

6.3 Complete receipt of amount of containers and empty bottles received from the customer and at the end of the day complete empty return slip of total amount received.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

6.4 Check cash payment from customer and calculate credit for customer in terms of breakages and empty containers and bottles correctly.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

6.5 Make daily entries on the monthly operations sheet.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

6.6 Ensure that all documentation is correct and handed in per deadline.

Almost never |-----| 2 |-----| 3 |-----| 4 |-----| 5 Almost always

APPENDIX 8

Inter-Item Correlation Matrix

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	SA1	SA2	SA3	SA4	SA5	L1	L2	L3	L4	L5	L6	L7	L8	S1	S2	S3	S4	S5	M1	M2	M3	A1	A2	A3	A4	A5	A6												
R1	1,0																																																
R2	,69	1,0																																															
R3	,59	,63	1,0																																														
R4	,62	,69	,71	1,0																																													
R5	,59	,61	,69	,74	1,0																																												
R6	,59	,66	,64	,69	,73	1,0																																											
R7	,59	,52	,62	,56	,62	,76	1,0																																										
R8	,53	,64	,66	,65	,68	,78	,75	1,0																																									
R9	,53	,54	,52	,61	,58	,72	,70	,76	1,0																																								
R10	,54	,57	,56	,55	,59	,70	,70	,70	,74	1,0																																							
SA1	,51	,62	,63	,64	,59	,63	,59	,70	,64	,63	1,0																																						
SA2	,57	,58	,57	,62	,61	,59	,59	,73	,63	,59	,63	1,0																																					
SA3	,59	,63	,60	,67	,58	,66	,58	,68	,63	,59	,68	,68	1,0																																				
SA4	,47	,56	,54	,54	,59	,64	,56	,71	,58	,60	,62	,65	,75	1,0																																			
SA5	,52	,58	,56	,57	,60	,72	,64	,78	,66	,67	,68	,68	,70	,71	1,0																																		
L1	,56	,59	,50	,61	,56	,57	,42	,58	,54	,50	,64	,61	,67	,61	,63	1,0																																	
L2	,46	,51	,58	,55	,59	,59	,51	,55	,42	,53	,57	,56	,61	,58	,60	,58	1,0																																
L3	,61	,67	,60	,67	,63	,67	,57	,67	,63	,57	,71	,61	,69	,66	,68	,66	,62	1,0																															
L4	,44	,55	,59	,60	,54	,62	,57	,61	,52	,53	,59	,56	,64	,61	,61	,52	,59	,65	1,0																														
L5	,44	,55	,61	,64	,60	,67	,66	,73	,58	,59	,66	,58	,64	,66	,66	,52	,57	,65	,70	1,0																													
L6	,46	,54	,59	,64	,63	,71	,70	,70	,63	,64	,71	,62	,68	,69	,71	,55	,56	,65	,68	,71	1,0																												
L7	,49	,56	,53	,61	,54	,62	,62	,68	,63	,63	,66	,55	,63	,58	,59	,54	,58	,62	,64	,74	,72	1,0																											
L8	,55	,61	,65	,65	,66	,67	,63	,76	,69	,71	,72	,63	,67	,66	,71	,58	,58	,67	,66	,73	,74	,73	1,0																										
S1	,46	,52	,53	,55	,51	,52	,47	,51	,48	,48	,46	,49	,49	,46	,47	,54	,56	,56	,57	,55	,55	,61	,56	1,0																									
S2	,43	,58	,51	,58	,43	,47	,40	,44	,39	,41	,56	,46	,51	,44	,40	,52	,48	,59	,62	,52	,53	,48	,47	,53	1,0																								
S3	,50	,58	,56	,56	,49	,44	,42	,45	,33	,40	,49	,46	,51	,47	,43	,47	,47	,52	,53	,46	,48	,46	,54	,43	,58	1,0																							
S4	,53	,58	,57	,60	,44	,46	,34	,51	,44	,46	,50	,49	,46	,45	,41	,50	,48	,55	,54	,55	,47	,58	,54	,51	,55	,54	1,0																						
S5	,47	,55	,55	,67	,52	,53	,42	,54	,50	,46	,49	,47	,48	,48	,47	,56	,43	,59	,52	,58	,53	,55	,55	,54	,50	,53	,67	1,0																					
M1	,43	,48	,46	,47	,42	,57	,52	,59	,53	,43	,51	,46	,50	,51	,55	,41	,49	,77	,53	,59	,58	,57	,55	,49	,45	,46	,43	,48	1,0																				
M2	,42	,46	,45	,50	,46	,55	,51	,54	,59	,48	,52	,51	,52	,52	,52	,48	,43	,56	,52	,54	,58	,54	,58	,47	,41	,40	,48	,42	,49	1,0																			
M3	,48	,61	,63	,57	,50	,63	,54	,67	,53	,68	,61	,54	,59	,56	,59	,53	,53	,61	,59	,62	,62	,62	,66	,47	,50	,51	,53	,56	,52	,49	1,0																		
A1	,50	,50	,55	,55	,53	,61	,60	,61	,61	,60	,63	,52	,59	,56	,52	,48	,45	,64	,61	,61	,61	,62	,67	,51	,47	,39	,48	,45	,49	,52	,53	1,0																	
A2	,51	,58	,56	,60	,56	,66	,58	,67	,53	,61	,63	,61	,64	,60	,59	,54	,49	,65	,70	,67	,63	,64	,68	,48	,50	,48	,52	,55	,53	,53	,62	,70	1,0																
A3	,58	,59	,62	,61	,59	,61	,61	,62	,63	,56	,58	,51	,53	,51	,54	,51	,45	,58	,55	,61	,59	,63	,68	,51	,45	,43	,51	,52	,47	,52	,64	,68	,59	1,0															
A4	,53	,57	,56	,61	,51	,59	,60	,62	,61	,65	,59	,54	,53	,51	,53	,45	,43	,60	,52	,61	,62	,59	,61	,51	,49	,50	,51	,51	,50	,53	,53	,53	,58	,56	1,0														
A5	,54	,57	,61	,66	,59	,70	,67	,71	,67	,56	,64	,58	,56	,52	,60	,51	,50	,60	,62	,68	,67	,66	,71	,50	,48	,50	,54	,56	,57	,51	,62	,59	,65	,62	,57	1,0													
A6	,54	,53	,57	,62	,57	,70	,69	,71	,70	,70	,62	,63	,54	,52	,59	,50	,43	,59	,56	,63	,67	,66	,69	,46	,44	,44	,51	,55	,54	,57	,64	,60	,62	,62	,63	,71	1,0												

Correlation coefficients larger than or equal to 0,31 are significant at 0,01 level.

R = Responsibility SA = Safety L = Liaison with Customers S = Supervision M = Maintenance of Truck A = Administration N = 40