PERFORMANCE MEASUREMENT AND EVALUATION

Block 2 The Building Blocks of Performance Measurement

Unit 4 Evaluating staff performance
The Course Team

The Open University Team

Jenny Lewis, Course Team Chair
(January 1991 – June 1993)
Jacky Holloway, Course Team Chair
(from 1 July 1993)
Michael Aczel, Author
John Blunden, Author
Alan Clarke, Author
Geoff Mallory, Author
Roy Needham, Author
Sue Pearce, Author
Derek Shields, Author
Tony Stapleton, Author
Peter Stratfold, Author
Richard Wheatcroft, Author
Sheila Cameron, Critical Reader
Chris Higgins, Critical Reader
Roland Kaye, Critical Reader
Linda Harris, Course Manager
(January 1991 – July 1993)
June Payne, Course Manager
(from August 1993)
Christina Boxall, Course Secretary
Riz Sharma, Course Secretary

Production

Sally Baker, Liaison Librarian
Pam Berry, Text Processing Services
Ann Carter, Print Production
Co-ordinator
Jo Field, Project Control
Roy Lawrance, Graphic Artist
Rob Lyon, Graphic Designer
Amanda Smith, Editor
Dick Smith, Print Production Controller

BBC

Jan Bright, Producer
Francis Sealey, Producer
Julie Marshall, Production Assistant
Christine Jackson, Production Assistant

External Assessor

Professor Andrew Chambers, Director of Centre for Internal Auditing, City University Business School

External Consultants

Hillary Barnard, Author
Martin Bennett, Critical Reader
Gareth Edwards, Critical Reader
Stuart Felstead, Critical Reader
Michael Green, Critical Reader
John Holloway, Critical Reader
John Jellis, Author
David Lloyd, Critical Reader
Annie Lucas, Critical Reader
Mark Matthews, Author
Alan Neilson, Critical Reader
David Perkins, Critical Reader
Peter Perneel, Critical Reader
Roy Reece, Critical Reader
Ray Ricketts, Critical Reader
Carol Russell, Critical Reader
Elaine Thomas, Critical Reader

The Open University
Walton Hall, Milton Keynes MK7 6AA
First published 1993
Copyright © 1993 The Open University
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, without written permission from the publisher or a licence from the Copyright Licensing Agency Limited. Details of such licences (for reprographic reproduction) may be obtained from the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1P 9HE.
Edited, designed and typeset by The Open University.
Printed in the United Kingdom by Halstan and Co. Ltd, Amersham, Bucks.
ISBN 0 7492 4842 4
For further information on Open Business School short courses and the Certificate, Diploma and MBA programmes, please contact the Customer Service Department, The Open University, PO Box 481, Walton Hall, Milton Keynes MK7 6BN (tel.: 0908 653473).
1.1
Unit 4  Evaluating staff performance

Prepared for the Course Team by John Jellis and Jenny Lewis, with a contribution from Sarabajaya Kumar

Contents

Aims 4
Objectives 4
1 Measuring staff performance – an overview 5
2 Performance appraisal revisited – words and meanings 8
3 Feedback and ethics 11
  3.1 The parable of the sadhu 12
4 Accuracy in appraisal 13
  4.1 Targets 15
  4.2 Reliability and validity in human performance testing 21
  4.3 The manager as actuator: motivation and performance-related pay 23
5 Group appraisal: an underestimated method? 26
6 A systems perspective on assessment 30
  6.1 Positive feedback in human systems 32
  6.2 The importance of variety in human systems 34
7 Conclusions 36
  7.1 A story of good intentions 36
  7.2 Performance-related pay in a large company 38
  7.3 Questions about assessment 39
References 41
Acknowledgements 42
Aims

- To introduce and analyse the complex interactions involved in assessing staff performance.
- To examine a number of assumptions and attitudes which influence the behaviour of organizations and their staff during the processes of assessing staff performance.

Objectives

After studying this unit, you should be able to:
- Describe how various common assumptions can influence and distort the process of assessing staff performance.
- Explain why assessment is a form of intervention.
- Relate the evaluation of staff performance to basic management control ideas.
- Recognize that assessment techniques are never neutral but are rooted in ethical and ideological issues.
- Compare the advantages and disadvantages of various assessment techniques, in terms of their ability to measure and motivate staff performance.
1 Measuring staff performance – an overview

Humans are both unreliable instruments of measurement and difficult subjects to measure. Their behaviour is changeable and, therefore, not completely predictable, and the potential for misunderstandings between them when one judges another is enormous. The language and culture of an organization add further layers of potential opacity and confusion to any evaluative activities. But organizational performance depends on people, so it is essential that we do the best we can to measure and evaluate performance, with the intention of improving it. Recent research by Fox and McLeay (1992), which is described in Section 7, reminds us that there is a clear positive relationship between the financial performance of an organization and the degree to which its corporate strategy is integrated with its human resource management practices.

Human resource strategies, as defined in the Open Business School's MBA course of the same name (B884), attempt to align the human resources of an organization with its strategic objectives so that the necessary behaviours are generated to achieve those objectives (see Figure 1).

![Diagram](image-url)
Whether or not your own organization has an explicit human resource strategy, as a manager exerting control over many organizational sub-systems you will be required to measure and evaluate staff performance. You will also be expected to work with the results of staff evaluation done by others, and to be evaluated in some way on your own performance. Whether the main evaluation process is called an appraisal or a review or some other name is irrelevant for the purposes of this unit.

Seen in control loop terms, it is perhaps not surprising that managers find the assessment process difficult, since at various times they are expected to act as sensors, target setters, comparators and actuators, as well as being assessed themselves. The performance appraisal process is stretched to its limits by being applied to recording performance, rewarding it, disciplining poor performance, encouraging good results, assessing training and developmental needs, and judging current and future promotability.

Although selection, development and promotional decisions involve similar judgements, performance appraisals or reviews are the most obvious mechanism for such evaluations. Performance review activities fit well with theories of management control since, as pointed out by the Work Research Unit (James, 1988, p. 2) amongst other authorities:

*Performance appraisal has its roots in three well substantiated psychological principles:*

(a) adequate feedback as to how they are performing, in other words, knowledge of results;
(b) clear, attainable goals;
(c) involvement in the setting of tasks and goals.

But measuring and evaluating human work performance is never easy. In Unit 3 we saw that measuring the attitudes and opinions of customers was an imprecise science. Here, we are not looking at averages, aggregates and samples, but focusing on individuals, their achievements and their future motivation. Measuring the behaviour and potential of staff is fraught with technical, social and ethical problems. As line managers become increasingly involved in performance appraisal schemes, it becomes more important to handle the potential conflicts between various groups of staff and, indeed, between individual career objectives and those of the organization. There is no doubt that the role of assessor can be stressful.

Performance appraisal is hardly a new activity for managers: even before F.W. Taylor claimed that measuring work rates could be made scientific, the measurement of output was commonly linked to financial rewards controlled by supervisors; while psychological tests of aptitude, personality and potential became widespread during and after the Second World War. What has changed recently is the emphasis on measuring and assessing white collar work, including that of senior managers, the nature of the targets against which performance is assessed, and the revived interest in linking performance directly with pay. There seems to be a trend towards ever more complicated schemes, which are intended to integrate performance appraisal with the broader human resource strategies of the organization. The more complicated the scheme, the more assumptions about the motivating effects of feedback and incentives it will probably contain.
In the late 1980s, UK companies and consultants began to import the idea of a performance management system (PMS) from the USA. According to a national survey (IPM, 1992), a PMS is in operation when the following conditions are met by an organization:

- it communicates a vision of its objectives to all its employees;
- it sets departmental and individual performance targets which are related to wider objectives;
- it conducts a formal review of progress towards these targets;
- it uses the review process to identify training, development and reward outcomes;
- it evaluates the effectiveness of the whole process in order to improve effectiveness.

(IPM, 1992, p. 5)

Whether the implementation of a formal PMS turns out to be as beneficial as its proponents claim, its design does seem to follow the principles of management control theory, with monitoring loops to enable the organization to test whether each part of the process has been set the appropriate targets, as well as whether they are being achieved.

The idea of performance review/appraisal is clearly central to any PMS. The authors of the IPM survey concluded that:

... if it is to be used as a part of performance management, performance review is likely to contribute most effectively if it conforms to a model which emphasises clarity, openness and accuracy.

(IPM, 1992, p. 34)

In other words, clear standards must be set, appraisers must be trained to assess staff against them as accurately as possible, and the results of the process must be honestly discussed with the appraisee. Your human resource specialists should be able to provide advice on how your organization's assessment schemes were chosen. This unit aims to provide enough information for you to enter into an informed debate where necessary, by considering what difficulties may arise when managers attempt to achieve the recommended clarity, openness and accuracy, in any form of performance assessment but with most attention being paid to performance appraisal.

We will look at how language can affect assessment processes, ask you to reflect on the place of ethics in judging human performance, and analyse some of the issues surrounding accuracy in assessment. Various appraisal methods are considered, including the less commonly used but recommended group appraisal process. Section 6 presents an unusual and controversial systems perspective on performance assessment. Finally, the last section draws together various threads running through the unit and reminds you of the questions that you should ask of any assessment process that is designed for measuring and evaluating staff performance.
2 Performance appraisal revisited — words and meanings

At this point in the unit you may be thinking ‘What, again? Surely everything about assessing or appraising staff performance has been said already?’ There is, indeed, a surfeit of management textbooks which explain how to design and use appropriate methods. Here, we approach the topic of assessment with an element of scepticism. Instead of studying how to do it, we shall be more concerned with why so many of these efforts do not work as well as they might. People in all types of organization have attempted to introduce programmes for performance improvement only to find that, after a while, the excitement fades. The programmes have failed to deliver all the promised benefits.

The philosophy and practice of continuous performance improvement for the entire membership of organizations are in the formative stage. We sincerely hope that your own work experiences and learning will contribute to its development and urge you to beware of ‘accepted views.’

We often forget that words and concepts have histories. Over time, their meanings may change dramatically; they can also be lifted out of one context and used in another.

'Speak English!' said the Eaglet. 'I don't know the meaning of half those long words, and, what's more, I don't believe you do either!'

(Lewis Carroll, 1893, Chapter 3)

Beware of traps! The meaning of words — of the language used — has become a major problem in human organizations. Unlike the Eaglet, many of us believe that we know the meaning when, in fact, we do not. For example, take the word ‘counselling’ and consider how its meaning can differ significantly from one type of organization to another. In one, it may be understood to be a process for helping people to cope with personal tragedy; in another, the same word can be associated with a disciplinary process. Another familiar example is the meaning attributed to ‘appraisal’. In some organizations, it is regarded as a developmental, helpful process; in others, however, it is seen as a threatening control device intended to weed out miscreants.

In recent years, increasing attention has been given to the meanings of words commonly used in an organization. Certain words take on special meanings; observers who do not know them will have difficulty in understanding the behaviour of that particular organization.

**Activity 1**

Write down some examples of words or expressions (i.e. jargon) which you believe have special meanings in your organization.

**Comment**

You may have had difficulty in distancing yourself sufficiently to see such words. My list includes:

- she's a 9.1 manager
- a real theory X department
- a first-floor decision
- a red-carpet decision
- AOTs (decoded as ‘adults other than teachers’).
What tends to happen is that a form of shorthand evolves to speed up communication in a particular network. The language seems to bewitch the intelligence: it can help but it can also hinder. We need to beware of mental traps!

Language does not just reflect a situation, it tends to highlight certain aspects and provide certain meanings. The connections between meaning, power and language are worth teasing out. For example, it is quite common for Equal Opportunities policies in organizations to support the promotion of assertiveness training courses for women. The reader is left to infer that such training is necessary if more women are to achieve managerial status. Such policies, unfortunately, do not pay attention either to men who may also need to be trained in such skills or to the behaviour within an organization that has made it appear necessary for women to be assertive.

Vickers (1972, p. 15) draws our attention to the nature of the trap:

Lobster pots are designed to catch lobsters. A man entering a man-sized lobster pot would become suspicious of the narrowing tunnel—he would recognize the entrance as a possible exit and climb out again ...

A trap is a trap only for creatures which cannot solve the problems that it sets. The nature of the trap is a function of the nature of the trapped. To describe either is to imply the other.

...we the trapped tend to take our own state of mind for granted—which is partly why we are trapped.

Vickers' notion of the trap and the trapped can be of much help to us when we examine how certain beliefs and assumptions form mental traps: traps of our own volition! Either we solve the problems which they set or we remain prisoners of our present, and perhaps comfortable, ways of thinking about staff performance in organizations.

Before you can discuss appraisal or assessment or any other concept, it is important that you take care to define what you mean by these terms. In this unit, several words are used synonymously: performance assessment/appraisal/review; organizations/human systems. For the purposes of this section, they are regarded as having similar meanings. Thus, appraisals and reviews are forms of assessment.

Dictionaries refer to assessment as the action of setting a price or value; as the fixing of a rate or tax. The word ‘appraise’ is also described as ‘to estimate’. In this somewhat narrow meaning of fixing a price, the words ‘assessment’ and ‘appraisal’ are synonymous. Indeed, in many organizations they are interchangeable and so we find performance appraisal programmes, performance assessment, assessment centres, etc. A factor common to assessment and appraisal is that of the act of comparing behaviour with some preconceived set of measurements of ideal behaviour, or at least against an organizational standard.

While recognizing that there are different approaches to who should and what should be assessed, the following definition of assessment is proposed for our purposes in this section.

A process of intervention to help an individual to improve his or her personal effectiveness and quality of working life. Such process to be done within the context of that individual’s work group.

You may have recognized that this definition is in itself deliberately controversial. By using the idea of assessment in this wider sense, we shall be examining the nature of relationships between individuals and how they influence motivation and performance.
Motivation is an important concept which can clearly illustrate how definitions and meanings of terms can be distorted in communication. The fundamental question is ‘Do you associate the same meanings with a concept that I do?’

In the time of Julius Caesar, the ‘galley slave’ theory of motivation was popular. Speed up the drum beat, give ‘em the lash and improved performance, in this case rowing, should occur. We hope that your theories of motivation will be more enlightened. They are probably influenced by the work of Maslow, Herzberg, et al. The important lesson, however, is that we must beware of assuming that the other person knows what we mean. The importance of clarifying the definitions of terms used cannot be overemphasized if the intention is to achieve meaningful dialogue.

Attempting to understand an organization is a very difficult matter. Nevertheless, there are ways that help. Morgan (1986, p. 13) in his fascinating book *Images of Organization* reminds us that we use a metaphor whenever we attempt to understand one element of experience in terms of another. He provides a number of metaphors which can be used to help us understand an organization, its purpose and its strategies.

Organizations as machines, brains, systems or cultures are metaphors which Morgan examines in order to identify their respective benefits and disadvantages in helping us to understand. He points to the risk that we select and use a particular metaphor and, in highlighting it, force others into a subsidiary role. It is advisable, therefore, to use a variety of metaphors in your assessment of performance in a particular organization.

You will probably have heard of F.W. Taylor’s ideas on scientific management. Taylor seems to have seen organizations as machines and there is nothing intrinsically wrong with this approach. Much of early management literature was influenced or written by engineers. It is understandable, therefore, that the machine metaphor was popular.

We shall use the metaphors of organizations as systems and as sleepwalkers later in this unit. Each of them should yield different insights. The systems metaphor will be discussed in more detail because of its importance to us when examining the possible causes of the mental traps which lie in wait for us!

*If there is any common human relations mistake made by organizational superiors in their relations with subordinates, it is the mistake of assuming that the ‘real’ world is all that counts, that everyone works for the same goals, and that the facts speak for themselves.*

(Leavitt, 1978)

Many of you will have met with these comments of Leavitt’s earlier in your management studies. Their importance, however, deserves to be repeated because of the significance of the three erroneous assumptions listed in that quotation. They remind us of the possible dangers of being too sure about things and they help us to reflect on the need to use caution when assessing performance.
3 Feedback and ethics

When assessing work performance or diagnosing the possible causes of work problems, we select relevant theories and models from our 'conceptual tool kit'. The value of those theories and models is determined by their utility. If they work, we keep them in our kit; if they do not work we discard them. The assessment of an individual's or a work group's performance, however, is not just a question of theory. We need to consider also the ethics of our actions.

Tyrants may be condemned by history but it is a chilling thought that many of them would have obtained a satisfactory, and in some cases first class, rating from some performance appraisal schemes. Achieving results is one thing, but the nature of those results is a different matter. Human conduct and performance need to be examined within the context of what is believed to be right or wrong according to the values of the society in which they occur. It is crucial, therefore, that we keep in mind the fact that the techniques themselves may not always be neutral. The content of assessment documentation must be checked to see whether it reflects the right values.

Activity 2

What aspects of assessment might cause problems on ethical grounds? Spend a few minutes thinking of examples, preferably ones that relate to your own experience.

Next, we suggest that you study the Reader article 'Ethical dilemmas in performance appraisal revisited' by Clinton Longenecker and Dean Ludwig.

Commentary

We chose this article because it openly debates issues that are too often considered undiscussable in the rational world of human resource strategists. The authors emphasize that managers acting as assessors (called 'raters' in the article) are only human, and may make deliberately inaccurate judgements for a great many reasons. The purpose of the article is to demonstrate the ways in which unresolved internal conflict can cause carefully designed assessment processes to be subverted, from the organizational point of view, by the assessors.

Do not be concerned if you do not understand all the terms used in the introduction to the article, for example 'halo effect' and 'recency'. These will be explained in Section 4 of this unit. It is also not crucial to understand all the nuances of the formalist, utilitarian or bluffing perspectives, as long as you can follow the general argument.

Activity 3

Using the typology of rater motives and manipulative rating behaviour from the Reader article, make a list of any such incidents which you have experienced, either as an assessor or when being assessed. What knock-on effects might one of these inaccurate ratings have had if it had occurred several times across the organization?
3.1 The parable of the sadhu

Now please read Bowen McCoy's article 'The parable of the sadhu' in the Reader, which discusses ethical dilemmas.

**Activity 4**

Would you stop what you are doing and help the sadhu?
What is the nature of our responsibility towards others if we consider ourselves to be ethical people?

A major complication when considering ethical issues is that, in our society, they are concerned with the individual, that is, in a face-to-face relationship. However, as McCoy suggests, the individual requires, and deserves, the support of the group. The success of both individual and group would appear to depend on the quality of that relationship and a willingness to co-operate. Here again we encounter the need to assess individual performance within the wider group context if we are to identify and pursue opportunities to improve the performance of both. This will be discussed further in Section 5.
4 Accuracy in appraisal

Longenecker and Ludwig, in their Reader article considered in Section 3, argued that accuracy was not the only key to improving the effectiveness of the performance appraisal process. Nevertheless, we believe it is worthwhile, in this section, to sum up the problems of accuracy in the measurement and evaluation of staff performance.

An illuminating but not, we hope, typical discussion at a Residential School revealed that managers working in mainland Western Europe tended to feel that the performance review process was an opportunity to keep in touch with, and motivate, their staff. In contrast, many UK-based managers admitted to a dislike of the process, because of the implications of having to tell less effective staff what they thought of their performance. Clarity and openness, when suddenly introduced into the work situation in relation to poor ratings of performance, are likely to disturb everyone's equilibrium. The whole process of being seen as a judge may be found to be threatening, particularly if the underlying relationships with staff are poor, or if a manager has had any negative experience of being appraised in the past.

You may think that measuring the actual behaviour of your staff should be a little easier than measuring the opinions of a sample of consumers, as discussed in Unit 3. Measuring the outputs of production staff sounds reasonably straightforward, and is an obvious part of Ouchi's control feedback loop, as described in Unit 1. But rating performance purely by output measures may not produce effective feedback if the output measures were not entirely under the control of the person being judged. Circumstances must be taken into account to avoid a classic error of assessment: it will not be accurate to rate a production manager as a poor performer if the new machinery ordered by the previous manager turned out to be faulty, at least as long as corrective action has been taken promptly. Unit 7 on quality improvement will also emphasize the need to distinguish between individual and systemic factors.

But we are not judging just current and recent behaviour. Far greater problems arise because we quite naturally want to predict future performance, during selection for development and/or promotion. We must be aware that judging someone and recording the results always implies consequences, even if the promotion or salary reviews take place separately from performance appraisal. Appraisers will be seen as actuators because of the feedback provided by their sensing and comparing activities. Hence the continual emphasis on clarity, accuracy and openness, so that justice can be seen to be done, at least by intention. Inaccurate assessments can quickly generate feelings of injustice with potentially wide-ranging consequences of demotivation.

Just as with any decisions that we make, in acting as judges of others, even against clearly defined external standards, we are all the victims of our own cognitive limitations. Bounded rationality (discussed in Unit 2) is always with us. We cannot obtain perfect information on how the person has behaved at work throughout the time under review; we cannot necessarily fully understand the jobs of everyone we appraise; we can only deal with so much information at once and we all have our own beliefs and biases about human behaviour. Inevitably, we interpolate and extrapolate from the evidence that is available to us.
Here are a few common examples of how our imperfections can affect accuracy.

1. The halo effect – we are sure that someone is excellent at one part of the work, so we assume, without having knowledge either way, that they will have performed well in other areas that seem to us related.

2. First impressions (primacy effect) – such a strong first impression has been made on us by this person that we have not observed their performance carefully since then for evidence that might prove contradictory.

3. Recency – very good or bad performance just before the review has over-influenced our view of the whole period.

Problems caused by our own biases range from the by now familiar stereotyping, which is typical of discrimination, to the tendency some of us have to use or avoid the extremes of any rating scale (central tendency). Thus an assessor who was predisposed to central tendencies would make little discrimination between good, average and poor performers, while another, appraising the same group, might rate many people as excellent or poor and very few as average. A prejudiced person might describe a black woman manager who produced excellent results as ‘unnaturally aggressive’ when her actual behaviour was similar to that of her successful white male colleagues.

Errors associated with ‘causal attributions’ are of special interest to us because they draw attention to the way in which an assessor might explain the cause of good or poor performance. If, for example, you are explaining the causes of your own poor performance, you would usually refer to the situation and to other external attributions. However, when explaining the poor performance of other people, there is a greater tendency to see the causes as internal attributions, that is, the fixed characteristics of the person, rather than the temporary demands of the situation – ‘He is reliable, she tends to panic, I am flexible’. Causal attribution errors combined with stereotyping can produce assessments such as that shown in Table 1.

<table>
<thead>
<tr>
<th>Table 1 Assessing your manager</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Him</strong></td>
</tr>
<tr>
<td>He's young – he'll stay with us for a few years</td>
</tr>
<tr>
<td>He's having lunch with the senior manager – he must be doing well</td>
</tr>
<tr>
<td>He got an unfair deal – was he angry?</td>
</tr>
<tr>
<td>He's talking with his colleagues – they must be discussing the latest company developments</td>
</tr>
<tr>
<td>He's not at his desk – he must be in a meeting</td>
</tr>
<tr>
<td>He's getting married – he'll be more settled now</td>
</tr>
<tr>
<td>He's going on a business trip – that'll be good for his career</td>
</tr>
<tr>
<td>He's leaving to start a better job – he knows a good career move when he sees one</td>
</tr>
</tbody>
</table>
We have discussed several examples of how your attitudes and feelings considerably influence the way in which you assess a particular person's behaviour. Furthermore, your personal style of intervention, that is the method of assessment adopted, will affect your ability to obtain valid information. Through custom and habit we forget that our ideas and beliefs about the various management activities are all aspects of an underlying ideology.

When using sets or summaries of appraisal data produced by colleagues, you need to be equally aware of all these pitfalls. Appraisal training for line managers (the most usual appraisers), and the use of a third person to sit in on the appraisal or see its results, are necessary to improve the accuracy of all concerned.

4.1 Targets

In setting targets for human performance, no absolute standards exist, nor are they ever likely to. Our ideas of excellent or atrocious performance will depend on our previous experience and our beliefs about human capabilities and inclinations. To guard against all possible assessor limitations, much effort has gone into designing and clarifying performance targets for use in assessing staff. In a modern assessment members of staff themselves will probably have been involved in designing their own targets, either directly or indirectly. Senior managers are interviewed in detail about their experiences in order to build up profiles of the abilities displayed by top and average performers, while anyone may be asked to fill in a self-rating form that includes a request to suggest specific work-based targets for the next period of assessment.

Most organizations will use a complex mixture of assessment strategies, and therefore of targets, as shown in Table 2, which is based on the IPM survey referred to earlier. Changes to the process will probably also occur frequently. The IPM survey found that 58% of performance review schemes were less than three years old. Different schemes may be used in different situations, and for different categories of staff within the same organization, although many experts now feel that this can be divisive and demotivating.

Rating schemes have evolved away from the requirement to judge mysterious qualities of character against targets towards behavioural scales anchored in

<table>
<thead>
<tr>
<th>Performance measure</th>
<th>Percentage of companies using measure for all or a majority of staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objectives agreed with employees</td>
<td>49</td>
</tr>
<tr>
<td>Quality based</td>
<td>45</td>
</tr>
<tr>
<td>Managerial judgement</td>
<td>44</td>
</tr>
<tr>
<td>Time based</td>
<td>39</td>
</tr>
<tr>
<td>Objectives set by manager</td>
<td>38</td>
</tr>
<tr>
<td>Output based</td>
<td>34</td>
</tr>
<tr>
<td>Training/learning targets</td>
<td>26</td>
</tr>
<tr>
<td>Requirements set on a group basis</td>
<td>18</td>
</tr>
</tbody>
</table>

(Source: IPM, 1992, p. 28)
detailed descriptions of appropriate behaviour. More recently in the UK, this trend has favoured the definition of management competences (formerly 'competencies'), either general (related to proposed national standards) or company-specific. Self-assessment scales are now normally used for personality or aptitude tests, rather than observation and guesswork by the appraiser. Such tests belong as part of a more complex assessment process, supervised by qualified practitioners. In practice, what often happens is:

... traditional performance appraisals still tend to incorporate personality factors such as dependability and attitudes and vaguely defined standards that may not relate directly to key performance areas or even apply to the job at all ...

(James, 1988, p. 7)

One reason why personality variables will not go away is that members of staff themselves believe that they are important. In a scheme recently evaluated by a Course Team member, staff in a small company, 'XYZ', liked the new competency-based scales but questioned the absence of certain more general traits:

... there was a recognisable faction that did not see the Management Competencies as the whole story in appraisal in that they do not directly take into account such factors as knowledge, enthusiasm, effort or commitment ...

... No 'box' for attitudes, enthusiasm, application, hard work (I know this is very abstract but it is a vital plus part of some of XYZ's staff who lack, as yet, other competence).

(Stratfold, 1992)

Competences, the newest type of target, are currently the subject of much debate (and are covered in more detail in B884 Human Resource Strategies).

If you are part of the scientific management tradition, you may view competencies as the specifications for the human machinery desired to provide maximum organizational efficiency and effectiveness. If you are part of the humanistic management tradition, you may view competencies as the key that unlocks the door to individuals in realizing their maximum potential, developing ethical organizational systems, and providing maximum growth opportunities for personnel. If you are one of the people who work in organizations or one of the people who studies, thinks about and tries to help organizations utilize their human resources effectively, this model and these findings should provide a needed relief from the eclectic cynicism or parochial optimism concerning management that many of us have developed.

(Boyatzis, 1992, pp. 271–272)

At the time of writing there are two approaches to management competences. The term originated in the USA with Boyatzis and his colleagues and followers. It is used to denote characteristics underlying high performance in fairly senior management roles, such characteristics or 'behavioural repertoires' being derived principally from 'behavioural event interviews'. This type of interview consists of a 'person-centred job analysis' carried out with people who are predefined by their organizations as excellent or average performers. The result is a profile of the competences which are necessary to perform averagely in that organization ('threshold' competences) plus those found only in high performers ('differentiators'). For example, one well known study (Boyatzis, 1982) found that threshold competences included 'accurate self-assessment', 'logical thought' and 'specialized knowledge'. Another set of studies (Bethell-Fox, 1992) found 20 generic differentiators, including 'developing others' and 'organizational awareness (understanding the power relationships within organizations)'.
The alternative approach, pioneered in the UK by the MCI (Management Charter Initiative), has produced lists of detailed task-centred functional and personal competences for junior and mid-level managers, competences which are linked to actual job performance, for example as follows.

Unit I.2 Contribute to the implementation of change in services, products and systems

Copyright material removed

OpenLearn

(Source: MCI, 1990)

In this view, competence is seen as an output and must, therefore, normally be assessed at work, unlike knowledge, which is seen as an input, testable by conventional educational methods such as written examinations. The attraction of such an approach for providing external standards and the chance to gain accreditation and qualifications is obvious, for both employer and employee.

Attempts to draw up MCI standards for mid-level managers and above are more controversial and, as yet (1993), in their infancy. Some large organizations have preferred to use the USA model to derive organization-specific competences for use in their assessment programmes, particularly where there is emphasis on the search for ‘high flyers’ so that the careers of such people can be carefully developed.

The main criticisms of the two approaches to competency are as follows.

1 Neither approach has yet been comprehensively applied across all levels of management in organizations over a number of years. Their reliability and validity are still open to question.

2 The UK approach is open to accusations that it promotes a view that staff should aim for sufficient competence, rather than for excellence, and that this is not a strong motivation for improved performance.

3 Being based on existing work practice and experience, both approaches either undervalue or cannot assess the creativity and flexibility needed to respond to changes in the future. Some proponents of competences (for example Cockerill, 1989) respond that in fact core, mature, transitional and emergent competences can be identified using such techniques as behavioural event interviews.

4 Being work-based, both approaches do not help managers enough in investigating the potential of specialized employees for management.

5 Neither approach offers non-managerial women or people from minority ethnic groups a route to proving their management potential, which may have been overlooked by their employers.

Through ways (3), (4) and (5) an organization may neglect to develop some of its staff who could improve organizational performance, by focusing in on too narrow a group of managers and potential managers: those who are outwardly similar to the currently successful ones.
Whichever model is used, once the list of competences is complete, activities to test for their presence must be derived. At the simplest level, for example in a small company studied by a colleague, a list of competences is added to the performance appraisal forms. The people being appraised then assess themselves on a scale against each competency on the list, as well as being assessed by their appraiser and sometimes by a third party. At the most elaborate level, a company may develop a complete assessment system for all staff based on its own specific competences, or those required for the national system of accreditation. Assessment centres are a likely way of implementing this type of strategy. The assessment or development centre concept (more fully described in B884 Human Resource Strategies) extends the principle of trying to make the measurement of behaviour more objective. Instead of the assessor recording what a person has done or might be thought likely to do in a work-related situation, the appraisee is placed in simulations of several such situations, and panels of trained assessors assess the behaviour exhibited. All candidates can then be judged against the same criteria, for example for their readiness for promotion. Simulations and other group activities can be combined with pencil and paper tests under intensive conditions. Research on Open Business School management students by Mabey and Iles (1991) showed that managers who had experienced the process found assessment centres to be one of the fairest ways of assessing potential.

As this brief summary shows, the competences approach is far more than a way of defining appraisal targets. It has the potential to integrate the principles of an organization’s human resource strategy by specifically deciding on the competences which are necessary to fulfil organizational objectives, and using them as the base for all the constituent processes of assessment, whether for day-to-day operational requirements or for long-term succession planning.

Assessing a person’s competency remains at least as skilled a task as any other form of assessment. Consider, for example, Table 3 which shows a pair of behavioural scales provided for assessors to use while observing a candidate’s performance in six simulated situations.

**Activity 5**

Attempt to rate yourself and a colleague against the scales in Table 3. What criticisms or difficulties do you have? You could discuss this exercise with other students in a study group or tutorial.

**Comment**

These examples, even though they were taken from a reputable organization, should have illustrated the difficulties of acting as a competency assessor.

In terms of management control, targets are clearly vital, whatever their derivation but, in practice, they are sometimes ignored, because they are seen as unimportant or unrealistic. More effective target setting could greatly improve the effectiveness of performance appraisal. Staff in the small company ‘XYZ’ mentioned earlier felt that this was certainly a weak point in their system:

A comment is in the lack of action on the recommendations and targets included in the form. It should be made more clear where the responsibility for this lies …

Targets: ‘The biggest loose end of the lot’

… there was little confidence that they [targets] would be referred to before the next appraisal.
Table 3  Behavioural frameworks in a council assessment centre

(Source: Dale and Iles, 1992, pp. 138 and 155)

... there was a need for continuing opportunities to review progress throughout the year, especially with a dynamic workload.

(Stretfold, 1992)

These honest, constructive comments illustrate the importance of staff participating in the evaluation of the systems by which they are assessed.

The best of the larger, well established organizations can provide examples of good practice, for example as follows.

Case Study: Organization E

A pharmaceutical company sited in the North West, part of a large multinational, with a comprehensive and sophisticated performance management system. The system starts with strategic objectives being set at the top and communicated
downwards. These are broken down and put in measurable terms at lower levels. The team may participate in setting its own targets, and then the individual manager suggests his own targets in the light of the team targets. The manager's business role (job description, etc.) is agreed and reviewed early in the year, then targets for the year are set. These are reviewed as necessary, but ideally quarterly. Each time this review takes place, the person's development needs are also looked at, with particular reference to achieving the targets set. There is an end of year review of target achievement and this feeds into the merit pay award. No copies of the form go to the Personnel department; the forms are owned by the individual and are referred to whenever targets are reviewed during the year. Also, there are no rating scales. (This may change as the company has been doing a lot of work to identify required competencies, but if competency profiles are introduced they will be used as a development tool.) The meeting of development needs, where identified, is a line responsibility. Personnel will provide advice and self-help materials to assist.

So far as pay is concerned, a new system is being introduced with fewer grades and wider bands. There used to be merit pay and cost of living increases but the two are being merged so that in the future there will only be merit pay. Though the amount of money available is determined at the top, heads of departments have to decide how to apportion the cash available, always remembering the overall goal of improving organizational performance. All PRP [performance-related pay] is subject to peer review. Managers make recommendations about the rewards they want to give, and these reward summaries are debated with peers and the immediate senior manager thereby facilitating comparison across individuals. Consultation and comparison amongst managers on the targets set for peers is also recommended. However, the emphasis is on trying to assess people against their own standards (in terms of target achievement), and not so much in comparison with others. ...

In terms of the IPM survey, Organization E comes out relatively highly on:

- overall job satisfaction;
- overall organisational commitment;
- participation in goal setting, clarity of goals (at all levels), the feeling of knowing how an individual's contribution fits in, feedback about individual goals and the sense of how the organisation overall is doing, a challenging level of goal difficulty, and the avoidance of short-termism;
- sense of team working;
- the feeling that performance is rewarded; and
- the feeling of being well informed about what is going on.

(Source: IPM, 1992, pp. 101-102, 112)

**Case Study: ICL**

ICL pays the same attention to detail in the measurement processes for staff as for its other activities. The staff appraisal process is not very different from good practices adopted in other companies which place a high value on their staff. Certainly it is seen positively by the staff; the Service Engineers consider it to be very 'fair'. Briefly, each is entitled to a quarterly meeting lasting about 2 hours, at which objectives are set for the next period. Discussion of performance against objectives leads to a performance rating, agreement of the next set of objectives and training requirements. Pre-appraisal forms are provided to help staff prepare themselves for the discussion. The Appraisal Forms, when completed, are signed by both Manager and Engineer, the Engineers 'can also comment on what they
thought of the appraisals'. Managers were generally considered to be very will-
ing to discuss problems over achievability of targets and objectives at a later
stage.

(Source: Case Study Booklet)

4.2 Reliability and validity in human performance testing

As discussed in Unit 2, Section 4, reliability means that a test (such as a question-
aire about personality or ability) is repeatable, that it will find similar amounts
of the same qualities every time it is used on the same individual. Validity is more
complicated in that it involves a test being able to demonstrate that it does test
what it is intended to test. For occupational qualities such as competences, a
valid test would have to predict that people diagnosed as having a particular
competence did indeed perform successfully in a role requiring that compe-
tence. Here, you can see the dangers of circularity; competences are derived
from records of performance in the first place. Validity would really mean that
the artificial test of the competence, such as a simulated activity, did indeed
select people who later displayed the competence 'for real'.

It can be argued that many of the tests and inventories used in assessment do not
need to be strictly validated in the psychometric (see Box 1 overleaf) quantitative
sense. Instead, they can be used to provide plenty of behavioural data for
feedback and discussion with the person being assessed and, therefore, a richer
picture of a candidate. This would have been the argument of the famous social
psychologist Kelly who wrote that reliability is 'a measure of the extent to which
a test is insensitive to change' and that validity refers to 'the capacity of a test to
tell us what we already know' (Kelly, 1970, quoted in Bannister and Fransella,
1977, pp. 76-77).

You may have met Kelly's *repertory grid*, a non-psychometric method of examin-
ing people's constructs and ways of thinking, on other management courses. If
you have not, but find the idea intriguing, we recommend Bannister and Fransella

Whatever your views on tests, the most important consideration for a manager
considering their use in performance assessment is that the people conducting
all forms of complex tests, inventories and activities are fully trained in their
intended use, and have the skills to debrief participants where necessary.
Otherwise, demotivation of those tested is an ever-present possibility.

Validated psychometric tests, such as those of critical reasoning, personality (for
example, the Myers Briggs Type Indicator – Briggs Myers and McCauley, 1988),
interests, values, cognitive style and interpersonal or team orientation, are a
special case. Anyone administering such tests should be specifically qualified,
for example, a Chartered Psychologist, who is accredited by the British
Psychological Society. Such a person will be happy to explain the rationale
behind any given test and to provide evidence of how its properties have been
validated on a large and appropriate population.

In the UK, the widespread use of psychometric testing for occupational purposes
has tended to run in cycles: in the late 1970s it was in decline, but became more
prominent in the 1980s and began to enter the undergraduate curriculum again,
as advertisements for well paid consultants appeared in the psychology press.
Box 1 Psychometrics

Psychometrics strictly means the measurement of individual differences in mental abilities. A great many pencil and paper (or nowadays computer administered) tests have been developed that claim to be able to do this, but the following cautionary remarks by two psychologists discussing assessment in 1976 still apply.

"The psychometrics industry is a bit like the motor car industry. It has its standard models that go on seemingly for ever, and it has its nine-day wonders; it has economy products and luxury products. Some countries favour official or semi-official tests, others prefer private enterprise. Any producer, to survive, must sell, and any test must offer something to the user. Good ones must offer some kind of reliability, though it is remarkable how long certain products can survive on claims that are highly dubious. One reason for this is that in many human situations it can be better to do almost anything rather than nothing. Another reason is that in practical situations it is often extremely difficult to carry out controlled studies. Men may be selected to be salesmen, or school-leavers to be university students, and most of them do fairly well at it. But it is very hard to get good evidence that they would not have done even better at something else, or that other individuals would not have done as well as those selected. The very fact of having been selected – or rejected – may itself influence what an individual does."

(Kirby and Radford, 1976, p. 41)

In order to aspire to psychometric validity, a test must be standardized, which involves repeatedly trying out the test items on a large and supposedly representative population. The items are then refined until only those which produce reliably consistent results when subjected to a variety of statistical analyses are retained. The final version of the test then produces the distribution of scores made by the standard (that is representative) population. When you or I take the fully standardized test, our raw scores are 'normalized' (adjusted statistically) to show us where we fall against this standard population.

The standardization of tests relies on two features. First, that the mental qualities being measured are normally distributed in the general population which, as discussed in Unit 3, is a substantial assumption for some abilities. Secondly, that the population used for standardization was appropriate. Therefore, you should ask whether a test's norms refer to the general population or a managerial population or a graduate population, and whether people of your gender and culture were adequately represented.

The main criticisms of psychometric testing are that:

1. the populations on which tests are standardized under-represent women and people from minority ethnic groups
2. the questions asked may reflect behaviours and attitudes that are not as flexible and creative as they ought to be to predict future potential in a modern organization – instead, they reflect the values and beliefs of a few years ago, when they were designed
3. the characteristics of personality being measured may not be as fixed as the proponents of the tests believe – people may be able to change and develop their behaviour sufficiently to invalidate the predictive value of some tests.

All of these criticisms mean that with psychometric tests you benefit from a professional body of knowledge being applied to your assessment problems, but there is a danger of rejecting (‘deselecting’) new, innovative staff who do not fit current patterns. An organization with no innovators is ill-equipped to manage change, so if this is a concern of yours it should be raised with your human resource professionals.
4.3 The manager as actuator: motivation and performance-related pay

One of the most contentious areas for assessors is appraising specialized professionals. Line managers may be required to do this in some situations even though they are not members of the same profession; for example, in the UK, National Health Service managers are increasingly being required to make judgements about the performances of medical and nursing staff. Great clarity about targets is needed to rise to this challenge. The process is still in its infancy in many companies, where often it is not fully understood that motivations vary between professional groups.

Assessing professionals is one of the most difficult actuator roles. We have already discussed how our own cognitive limitations constrain our accuracy. Our beliefs and ideologies about what motivates others to perform better will also have an effect.

Activity 6

Examine the appraisal documents which your organization uses for junior staff and compare them with those on which you are appraised. List any different assumptions about motivation and how it varies among staff groups that you think are implied between the two types of work. If you have access to material used to appraise the most senior staff, repeat the exercise.

This activity may have been difficult without details of how the appraisal results are connected with rewards. Although it is true that being assessed as doing well against challenging targets is inherently rewarding, this is hardly the whole story of motivation. Pay, perks, bonuses, promotion and some sorts of training are all rewards. The use of these to motivate the best performers will vary greatly across different categories of staff in most organizations.

At recent conferences, and in the IPM report (1992), the question that most occupied human resource strategists was how to motivate the majority of ‘average’ workers. Special schemes to reward the brilliant can be demotivating for those who try as hard or harder and form the ‘backbone’ of the organization. This applies just as much to middle managers as it does to blue collar workers. Non-intervention in the case of very weak performers can also demotivate the rest, if they feel that management cannot distinguish solid performers from the weakest, or if their own work is suffering from the underperformance of others.

Performance-related pay (PRP) is currently (1993) fashionable in the UK. It was used in 74% of the organizations studied in the IPM survey. PRP was defined as covering profit sharing, merit pay, individual bonuses, team bonuses and share options. However, the survey found that ‘there was no correlation between the use of PRP and organizational performance – poor financial performers were as likely to have PRP as good performers’.

Designing an effective PRP scheme is a job for an expert who should consider all aspects of the organization’s human resource strategy, from job evaluation schemes to future staff career planning projections. The ramifications are extensive, requiring a whole unit in B884 Human Resource Strategies. Figure 2 and Table 4 (overleaf) illustrate some of the concepts involved.
The most obvious topics for us studying performance measurement and evaluation are the familiar challenges of connecting feedback and reward firmly to circumstances over which performers have control, and ensuring that ambiguous targets are not set. Departments which aim at targets that put another department at a disadvantage dissipate energy that should go towards organizational objectives. An organization with a training scheme that fails to impress on its assessors that not all staff can attain the top 'super excellent' category will need centralized intervention to avoid paying out more than it can afford. A further danger is that some companies, for example those in the fast-moving retail markets, need to adjust their targets frequently, and can still be caught having to reduce promised rewards if there is a sudden slump. Disappointed expectations are always demotivating, for example in the following experience of a young public sector manager:
'When I achieved my targets within eighteen months as opposed to the agreed time of three years, my manager had to recommend that I be paid in the top band, according to my performance. However, despite the fact that the money was in theory “ring-fenced”, in practice there was no money available so I was not paid the extra salary. I began to question the ethics of the organization, became very demoralized and began to look for another job.'

In summary, PRP can be an effective motivator but its current popularity may stem as much from a desire to cut costs and an ideology maintaining that money is the main motivator for most staff as from any conclusively demonstrated universal benefits. A yet more sophisticated approach involving benefits tailored to various groups may be equally suitable for forward-looking flexible organizations.
Group appraisal: an underestimated method?

In this unit we are not directly concerned with how appraisal systems should be designed but we are very much interested in assessing their contribution to the corporate objectives of the organization concerned. Therefore, we should be aware of the different means of appraisal. For example, one-to-one appraisal is the most common but may not necessarily be the most effective in all situations. While there is a growing interest in group appraisal, there is still little evidence of its use (see Table 2). Another option is to use a third party, a facilitator who is present during either one-to-one or group appraisal. The facilitator can bring 'energy' to the event and, in some cases, help to create the excitement and commitment necessary to 'wake up' the staff involved. A further option is for the person concerned to assess herself or himself before being appraised. The assessment process does not have to be the prerogative of management.

Activity 7

Write a short answer to the following questions.
Who do you think should assess an individual's performance?
What do you think should be included in the assessment?

Comment

There is, of course, no one correct answer to either question because organizations can differ considerably in their management styles, values, beliefs and attitudes.

Develop the individual members and you develop the organization' is a very common assumption. The belief that an individual can be effective, provided that the job is designed on sound principles, is widespread. Supply the right tools, clarify the objectives, ensure that the person has the requisite skills, and the rest should follow. Unfortunately, sound principles of job design are not enough. Most individuals are members of work groups and the relationships between the two are very complex: each affecting the other in various ways. This relationship is illustrated by the ways in which control loops operate. Some are confined to the individual, some to the work group, others operate across both.

There is a long history of training and development programmes which are intended to improve individual effectiveness but the crucial significance of the relationship between each individual and his or her work group has been neglected. For the purpose of this section, a work group is defined as 'a group of people who are responsible and accountable to the same person'. The members of a work group will regularly and frequently interact with each other while pursuing their individual and collective objectives.

A surprisingly large number of organizations cling tenaciously to the belief that once a sufficient number of individuals have received the same training and development, departmental performance should improve. This 'critical mass' theory underlies many existing training and development activities. Large sums of money are invested with the aim of improving, for example, a manager's competences. However, while this is necessary, it is not sufficient. An organization is much more than the sum of its individual members. It consists of work groups which influence the motivation and performance of its members and are, in turn, affected by them. Each needs the other in order to achieve overall
effectiveness. Ignoring the importance of this relationship leads to a distorted view of what should be assessed and why.

Many organizations give every encouragement to training and development programmes for individuals. Interestingly, that enthusiasm usually wanes when the need to provide similar opportunities for work groups and senior management is suggested. If the work group needs are ignored, individual members will have great difficulty when attempting to apply what they have learned on courses to their work. The work group sub-culture determines how things should be done. Attempts, therefore, to assess individual performance, without assessing the work group performance, will achieve, at best, a partial view; a sort of tunnel vision. The relationship between the work group and its individual members is so important that we cannot assess one without the other. We must examine the reciprocity of information and behaviour between group and individual if we are to understand the individual's role and performance.

We should also keep firmly in mind the fact that an organization of people is a human system. A central problem in human systems is that you cannot consider any one particular aspect of behaviour in isolation. The presence of so many variables in the system tends to obstruct our efforts to understand any one of them alone. As a sub-system, an individual might achieve all of his or her performance standards but fail to establish effective relationships with peers, senior manager or subordinates. Such failures can remain hidden unless the group is involved in appraising both the individual and itself. When this is done effectively, the group will be operating as a team.

A work group, therefore, is not just a collection of people, it develops an identity that is shared by its members. That identity helps members to give meaning to what they do. Work groups are the basic building blocks of departments, divisions, faculties, functions and entire organizations. The ways in which decisions are made and taken in groups; the manner in which conflict is avoided, suppressed or managed; all these things impinge on individual members' motivation and performance. The group's motivation is, in turn, affected by its members.

Therefore, it is intriguing that very few organizations have attempted any form of group performance appraisal – that is, appraisal of each individual member of the group by his or her colleagues, the manager of the group being included in the process.

Any assessment process is a form of intervention. In group appraisal, instead of the assessor as judge, the assessor as facilitator is the recommended metaphor. The success that is likely to be achieved from interventions by a facilitator will depend on the culture in which it occurs. A performance profile is suggested as the facilitator's first step. The profile will indicate what is feasible: the art of the possible.

The purpose of group discussion about each individual as part of the appraisal is to give advice on opportunities for improving performance and to discuss how personal development might be assisted by other colleagues. The outcome of group appraisal should be individual and group development plans, which provide the core of subsequent meetings to review progress. For meetings of this type, it is recommended that a third party is present. This person will, ideally, be trained in assessment and intervention skills.

While some managers find the idea of the group approach attractive, it is seldom taken up. It seems that organizations resist involving work groups. Consequently, many of the opportunities which can only be handled by part of or the whole work group are lost and problems remain. Are there, perhaps, some unspoken, even unconscious, forces which regard individual performance assessment as acceptable but which regard group work as threatening?
During the group appraisal process, individual and work group performance should be assessed jointly by each work group with the assistance of a 'facilitator.' (The term 'assessor' has too many connotations with the ideas of 'sitting in judgement'.) The act of assessment can be the responsibility of the work group assisted by the intervention of a facilitator who will have previously negotiated a contract of the expectations of the group, methods of working, etc. The intervention can, therefore, be the facilitator's actions and contributions that assist the group to become aware of, to 'wake up' to, and to 'break out' from ineffective behaviour. The work group and its individual members should all be further helped to identify opportunities for continuous improvement.

**Example 1**

When group appraisal was used in a large organization, a facilitator was present to help the groups with their analyses and subsequent planning. It was found that each group event was completed in half a day. The first session discussed what the key result areas (KRAs) should be for the group. Following this each member of that group prepared his or her individual KRA description which was then discussed and agreed with the group manager. Each individual was given two weeks within which to prepare the KRAs and during this period could seek advice, if necessary, from the facilitator.

The agreement of individual KRAs was part of a newly designed performance scheme at which the facilitator was also present if the appraisee and appraiser requested it. This idea of a third party proved to be helpful, especially in cases where the appraisee and appraiser had difficult relationship problems. The facilitator could accumulate a rich collection of work problems, many of which were found to be common to more than one function. Furthermore, the experience of problem solving in one group often helped to identify problem-solving ideas and methods which were made known to other groups. The ratio of facilitators to managers was 1:250, that is 1 facilitator available to 250 managers (whether in work groups or as individuals) throughout an annual cycle.

**Example 2**

A work group of eight supervisors was invited to experiment with the idea of group performance to improve collective performance and the job satisfaction of its members. KRA analyses were prepared by the group and, subsequently, by its individual members. The group work needed half a day; individuals, working alone or with the advice of a facilitator, were given two weeks to complete their KRAs. A further one-day meeting of the group took place a few days after all the KRA work had been completed and agreed. On this occasion, the characteristics of behaviour that were considered necessary for effective group working were identified and agreed. They included improved time management, decision making, feedback, delegation and inter-group working. Brief statements of desirable performance standards were agreed for each characteristic. The group then compared these with present performance standards and group improvement plans were prepared.

The group performance appraisal took place about two weeks later. Each member of the group in turn, including the group's manager, received feedback from colleagues. The feedback described the perceptions of each individual regarding present strengths and suggested areas for improvement of the appraisee. The facilitator was present to help with the process.
Example 3

We were an interagency team working with the health service, the local authority and the voluntary sector and we depended on individuals and groups working across many different sectors in order to achieve our targets. When our team started, we appeared to be in conflict and competing with a similar group with overlapping responsibility, within the same department of the same organization, yet with seemingly different goals. So we (the other team manager and myself) invited a facilitator to help us to identify and agree on solutions which would enable us (each team) to improve our performance. The individual members of each team were, throughout the course of three days, encouraged to share their ideas about their roles and how they would like to attain their individual targets. Both teams then discussed and agreed joint key result areas and the teams began to work together very effectively.

When key result areas are used as in the above examples, each KRA description, whether for a group or for an individual, will typically contain five to nine KRAs, each with a number of performance standards and how their attainment will be assessed. This assessment need not be quantitative although it is likely to be so. From a set of KRAs individuals can derive their performance improvement plan, agree it with their assessor(s) and work towards it. Of course, KRAs will vary with the nature of the work but common ones include customer relations, contribution to profits, student pass rates, patient recovery rates, research and development, productivity and employee services. For each of these, performance standards and performance indicators can be defined. The KRA method is thus a way of setting targets that are derived from the key objectives of the organization. Your organization may well use a similar process that has a different name.

The facilitator is not a judge during group appraisal but has the very important task of bringing energy and encouragement to the group. The main purpose is to help the group acquire the practice of continuous performance improvement. Sometimes the facilitator will need to criticize a group’s current and proposed work behaviour and will not be playing a purely non-directive role. Any criticism must be constructive and examples of what has been seen and heard provided as evidence for it. In a sense, the facilitator is an action-researcher helping a group to learn by doing. Learning by doing is feasible only when the people involved are clear about what they are expected to do and how well they are expected to do it. It is essential, therefore, that each individual and work group clarifies their KRAs and associated performance standards. Much depends on the present levels of trust and the skills of those involved to make constructive contributions.

Organizations which are curious enough to experiment with new approaches to assessment should not be over-ambitious at the beginning. For example, a chronic problem in the business could be declared as the area to which groups should devote their improvement efforts for the next, say, three months. The group appraisal intervention, therefore, helps groups and individuals to assess their performance against objectives which are related to overall organizational objectives and to plan for continuous improvement.
6 A systems perspective on assessment

As discussed in Sections 2 and 5, there are several false assumptions in which, as Vickers (1971) described, ... we the trapped tend to take our own state of mind for granted.

The false assumption we shall consider here is that all organizations seek to improve their effectiveness. This is not meant to be a facetious statement; indeed, it is a deadly serious one which may determine the success or otherwise of the performance assessment process as part of a human resource strategy, for example a performance management system (PMS). The complementary question is Why do organizations appear to be committed to innovation and new ideas one day and then lapse and lose interest the next?

**Activity 8**

Can you recall some examples of this sort of behaviour in your organization? Make a brief note describing the event, who started it, and how long it took for the initial impetus to fade.

**Comment**

One example we remembered was Company Z which invested a large sum of money in an 'in-company' programme of seminars and group activities designed to help it become more effective. In addition to this programme, it introduced versions of quality circles and briefing groups. Because these various initiatives had not been integrated, there was considerable confusion among participants. To make matters worse, the system of performance appraisal already in operation had no explicit connection with any of the other activities!

There is something distinctly odd about the ways in which human systems, that is organizations, seem to behave. Sometimes they say one thing but do another. Such behaviour appears to be common to all of us and it suggests that people, both as individuals and in work groups, share certain characteristics, one of which is perhaps the concept of the unconscious.

Freud's psychoanalytical writings did much to demonstrate how the concept of the unconscious can provide help in treating people who are suffering from various psychic disturbances. He observed that any thoughts now present in a person's consciousness may vanish and appear later, unchanged. This strange, indeed miraculous, phenomenon we describe as memory. It is not caused by any change in our perceptions of the world around us. We must assume that the ideas which surfaced must have been in our minds but were, in some way, latent.

This idea of a latent, sleeping state can be very helpful when we are trying to understand and explain the complex and sometimes strange behaviour of organizations and their members. The discrepancies identified between intent and action in both organizational and individual performance can be explained as forgetfulness. While people claim to seek improved performance, their subsequent behaviour so often belies it. However, if corporate excellence is desired, an organization must stay awake if it is to manage the changes that are necessary to move from the present state to a better one. Perhaps this sleepwalking tendency should be considered together with bounded rationality and unresolved internal conflict when explaining ineffective organizational performance.
There is much discussion about the accelerating pace of change and the need for change strategies. We are in a constant state of change. This turbulence will affect performance criteria and, in some instances, will eliminate, modify or replace them with others.

Any proposed changes will need to be examined to see the impact they could have in other, often unexpected, parts of the organization. While most of us seldom use the word ‘systemic’, all of us have experienced what it means: it is the effect, usually unforeseen, of an action designed to be confined to a particular part but which also affects other parts that were not intended to be involved. It is, for example, the unexpected effect of an incentive scheme leading to employee disinclination to seek promotion. Or it may be the confusion created in the minds of managers who, as a result of attending a management course and believing that certain behaviour is recommended, discover that such behaviour is anathema back at the workplace. When you start to pull a loose thread in a carpet, you can never be sure what the implications will be for the pattern woven into it. Yet, in many organizations, there are people who, in a metaphorical sense, are happily pulling threads and ignoring or forgetting the possible systemic effects of their actions. But the complications do not end there. Because so many changes are implemented in ignorance of their possible effects on other parts of the organization, they collectively can inflict appalling damage on the motivation of various individuals, their work groups and departments. Measures of effectiveness need to be regular in order to pinpoint any such damage.

**Activity 9**

How many change programmes are taking place right now in your organization? (A ‘change programme’ is an activity which has been formally approved by senior management and is intended to change one or more essential elements of the organization. Those elements include methods of managing information; the way in which the organization is structured; the design and nature of the work done; and the skills and attitudes of the people doing the work.)

**Comment**

This activity was intended to remind you of the danger that any new assessment process becomes just another change programme to be coped with by hard-pressed managers and their staff. The expressed desire to improve effectiveness and search for excellence has been the focus of much popular management literature. Mission statements generally pay special attention to it. However, as we have seen, many organizations are simply, but sadly, kidding themselves: they need to become conscious that many good intentions and policies are never implemented, and many schemes that are implemented are never evaluated.

**Example**

A company introduced a management by objectives (MBO) system, including performance appraisal, quality circles and briefing groups. For a while they all ran in parallel. Eventually, after four years, they faded out because they never properly belonged to the business; people could not see what use they were.
The distressing feature in the example was the absence of any genuine evaluation of the contribution to individual and organizational performance of any of the new schemes. This cycle of high interest followed by organizational disinterest is a common experience for many managers.

Ideas from systems thinking provide a possible explanation of why human systems differ significantly from other systems. The assessment of individual and work group performance, as we have already noted, is an extremely complex and convoluted activity. Performances of each individual and of the group are frequently interacting; attempts to isolate one from the other are unlikely to be more than partially successful. This is because some of the control loops cross the boundaries of individual and group systems.

The four organizational levels as suggested in Pugh’s Organizational Development Matrix – that is, organizational, intergroup, group and individual – remind us of the possible interactions. (Pugh’s OD Matrix was described in the Planning and Managing Change module of B880 The Competent Manager.) While control loops to achieve desired outcomes probably exist at each of the levels, they may not operate at only one level. For example, the performance and results of a sales department will depend on group as well as on individual efforts. From an observer’s perspective, the performance criteria could include product promotion and customer service, which should be ‘owned’ by the group.

At the individual level, control loops should provide the feedback necessary to take remedial action if needed. Some cynics have suggested that management consultants can make more money from diagnosing deficiencies in control loops and offering solutions to repair them than through any other activity. There is some truth in this because so many performance problems are caused by inadequate or missing feedback loops.

### 6.1 Positive feedback in human systems

We are most aware of the control loops which rely on feedback that enables us to act if there is a deviation between the desired and the actual outputs. Reducing such unwanted deviation is a subtracting process; hence it is termed ‘negative’. It returns the system to its original equilibrium. There is, however, a unique characteristic of human systems and that is how they can also use positive feedback to great advantage. Positive feedback destabilizes the system and moves it to a new state of equilibrium.

Several writers on management and organization theory have regarded human systems as ones which attempt to retain stability, to maintain order. As a consequence of this, human systems could be classified with other types of system. However, a common property of other systems is that they are subject to the Second Law of Thermodynamics which states that the world is irreversible and approaching a state of disorder or maximum entropy, that is, randomness. In this sense, the ultimate system – the universe – is a closed system. Within it the various sub-systems are assumed to be running down.

So what has this got to do with the assessment of individual performance? The answer is ‘quite a lot’, if we accept an alternative view that human systems do not inevitably degenerate into the disorder mentioned above; that they have certain unique properties which separate them from other types of system. One of these properties is that, instead of running down, they tend to drive the system to a
new state of complexity. In such transitions, the system acquires new opportunities and possibilities for action. The physiological view that human systems tend to return to a state of equilibrium after temporary disturbance is, therefore, challenged. So too are models of organizational structures which tend to minimize change and enforce stability. This capacity to move towards a new dynamic state suggests that human systems are non-equilibrium systems—sometimes referred to as ‘dissipative structures’. Human systems have a particularly high degree of energy exchange with their environment. They are richly coupled sub-systems which can absorb change. This is captured in the common saying ‘You can’t succeed in business by standing still’.

Human systems become frustrated when freedom of action is restricted, that is, when options for action are minimal. In systems terms, at such times they would have high entropy. Note that a minimal range of options will generally exist where the system has operated in much the same way for a long time and staff behaviour is much influenced by rules, regulations and procedures. An example of this is a long-term, well established bureaucracy. Such work situations are likely to demotivate staff who become bored through the dullness of routines and the absence of any excitement.

Imagine that you arrive in this situation and, as a keen educated manager, you are determined to make your mark and improve results and staff motivation. You have an MBA and you know that several options are available for you to use. You might decide to begin with quality circles or to introduce a new performance appraisal system. Soon, noticeable changes begin to occur: absenteeism decreases, performance standards agreed with staff are achieved.

In this example, you represent, in systems terms, an inflow of energy (information) and the ensuing instabilities caused by the application of your ideas help drive the system to a different state of behaviour. The fluctuations in the system have led to a different state of order; but surely this is not the first time that management initiatives such as yours have been attempted? Staying with our example, it is highly probable that, in its early years, management initiatives took place and resulted in improved effectiveness. The problem seems to be that each successive state of behaviour, in turn, develops high entropy as routine and habit become established. The effects, over a long period, might resemble the pattern shown in Figure 3.

The effects of new energy and change become institutionalized after a time and new energy is required. If the organization is lucky, it will achieve the ratchet effect shown.

---

**Figure 3** Human system’s ability to absorb change
(Source: Jellis, 1990, unpublished)
Unit 4 Evaluating staff performance

Another way to consider this sinuous behaviour is to use the 'sleepwalker' metaphor as shown in Figure 4 (Jellis, 1990). The idea of organizations as sleepwalkers can be immensely useful because it suggests that the act of awakening can be induced by design. Perhaps by the assessment process?

**Activity 10**

Can you recall an example of this suggested sinuous behaviour in your organization? Do you remember the excitement caused by the introduction of new ideas followed by the eventual dissipation as the ideas lose their original attraction and are absorbed into the organizational culture? Try drawing a diagram of this experience.

6.2 The importance of variety in human systems

Human systems thrive on variety both in the ways they do things, and in the ways they think. It is the critical ingredient in any policy for encouraging – that is, rewarding – individual and work group creativity, initiatives and innovation.

You would expect different functions and departments to have different needs. The differences need to be satisfied and there is no reason why there should not be a variety of options permitted for groups to do this, for example different key result areas. As long as the ends – the outputs – of a scheme are satisfactory, the means can be left to the discretion of the groups concerned. Variety, it is said, is the spice of life; for individual and organizational performance it is also a source of creativity and innovation.

Variety in practices can move a system from a state of routine to heightened awareness; it is the new energy shown in Figures 3 and 4. Increasing the variety in a system is a phenomenon of positive feedback. While it is desirable for certain activities, it can be destructive beyond a certain range in others. Managing the amount of variety in a system is, therefore, a very different activity from the traditional control function in which feedback is used to restore equilibrium. The latter is familiar to all managers; but the former is much less understood. Proactively managing the amount of variety which is desirable in a system is likely to become a major area of study. We want some variety but too much of it would overexcite and, possibly, destroy the purpose of the system.
Drucker (1969, p. 23) predicted the end of continuity:

Now, however, we face an Age of Discontinuity in world economy and technology. We might succeed in making it an age of great economic growth as well. But the one thing that is certain so far is that it will be a period of change ... in the knowledge needed to govern and manage ...

That knowledge includes the management of discontinuity, that is, managing sudden, large, qualitative variations in a system's behaviour when past precedents, policies and practices are no longer appropriate. Although it would seem much more desirable to stay with the controls with which we are familiar, increasingly we are faced with the dilemma of whether to manage change in a system or to be swept along by events: whether to be proactive or reactive.
7 Conclusions

In this unit you have looked beneath the familiar labels (such as appraisal, KRA, MBO, competences) to reflect on the underlying principles of assessing staff performance. Studying performance measurement and evaluation in this area and at this level is made harder by the lack of detailed published case studies. Any information that you can gain from other students will therefore be valuable in comparing the processes used.

We know that policies and practices vary immensely, even within the larger organizations let alone across industrial sectors or countries. For a small or medium sized organization, any formal assessment system will probably be welcomed as an improvement, since many still lack any form of staff appraisal. If such a system is carefully thought out and well understood by everyone who participates in it the results are likely to improve organizational effectiveness via better staff performance. Larger organizations will need more sophisticated assessment programmes, but these too will need to be flexible enough to be adjusted as business objectives change.

7.1 A story of good intentions

Originally we commissioned a case study about the implementation of a performance management system in a large UK financial organization, to compensate for the lack of relevant published material. However, halfway through the study the organization abandoned the scheme. We therefore sadly present a cautionary tale.

How not to implement a performance management system

Once upon a time a well respected UK financial institution decided that it needed a thoroughly modern performance management system (PMS) to help fulfil its new mission statement and thus improve its organizational effectiveness. Instead of implementing the PMS incrementally, it was introduced into the 200-year-old traditional, paternalistic, hierarchical organization practically overnight, in a haphazard fashion – at different times, in different phases and in different parts of the organization.

Junior white collar workers, having always been told what to do and how to do their jobs, were suddenly 'empowered' and then introduced to a complex PMS, developed for the organization initially by two different sets of consultants.

The organization was going through major changes and everyone was concerned about how to implement the PMS when they did not even know whether they were going to be employed the following week, since over a period of five months the organization had restructured and 'downsized' by 1000 people.

The following issues arose.

- Some staff perceived the scheme as very bureaucratic and costly and did not believe that its benefits would outweigh the costs, although they had no grounds for this, as the organization had not costed the scheme.
- Zero error rates when processing desk work was deemed by some to be demotivating, as it was an unattainable target. However, it was seen as motivating by others as it was aiming to be the best — 'Being right, first time, on time, every time, is the only target to aim for, as any less of a target, say for example 75%, would be indicative of a slippery slope.'
- Before the implementation of the PMS, team leaders had been asked to develop input-orientated job profiles, but the PMS was output-orientated. They were then required to develop new output-orientated job profiles which they resented, having already been through one exercise which they had to scrap within a few weeks. They were also finding it difficult to develop, as some felt that it was just a matter of semantics, there being no difference between 'input' and 'output' or, if there was a difference, they did not really understand what it was!
- Team leaders who were having to implement the PMS were doing so without clear direction from the top and no clear mechanisms for feedback on whether they were succeeding.
- There was some consternation expressed about the fact that the scheme had been developed by consultants who were no longer involved during the implementation phase so that they could not help iron out difficulties with it.
- Everyone found the terminology confusing and very complicated — job profiles, core accountabilities, current year objectives, competences, work measurement, work control management systems, inputs, outputs, functional relationships, key result areas, etc.
- Parts of the work control measurement and management system concentrated on measurable objectives, that is how much work had been processed (which was exactly the same as the 'clear desk policy' which they operated before the advent of the performance contract). Performance as linked to the new mission in terms of quality — that is, completion of tasks with a zero error rate — was not measured!
- A rise in the overall percentage of work cleared would be shown as an improvement in performance of the team whereas there might in fact have been no improvement, as only completion of tasks is measured, assuming a zero error rate. Conversely, a drop in the overall percentage would indicate a deterioration in performance and productivity, whereas in fact the opposite might have been true — there may be a marked improvement as, although the activities completed may be fewer, there might actually have been a zero error rate and therefore a higher quality of service delivered to the customer.
- The language used to describe activities meant different things to different people and there were no commonly accepted definitions. So, for example, one individual could have ascribed the work he or she had just done showing a new recruit how to undertake a certain task to the section headed 'training', which was not seen as 'time productively spent' but dead time and would bring down the overall performance of the team in terms of productivity; whereas another individual in another team could have ascribed the same activity to another heading such as 'one-off/miscellaneous' which would have been seen as 'productive' and would add to the overall performance of the team.
- The information gleaned from the work control measurement was seen to be too ambiguous to be of any use in measuring performance and was only meaningful if seen in relative terms and interpreted by the team leader.
- It was seen as unethical if other managers had access to reports of teams other than their own team, as they would not know the team members and would have to judge their performance and draw conclusions in a vacuum, without knowing the individuals' capacities and capabilities.

Open University MBA Programme 37
In this case, the business environment moved too quickly for the substantial ‘teething troubles’ to be dealt with so that, despite the use of experienced consultants and the efforts of the implementing departments, the scheme failed. It was far too complex for an organization whose human resource management functions were not as well integrated as they might have been, and whose staff felt threatened by restructuring. The scheme as presented to those who had to use it lacked clarity, openness and accuracy. To have had more chance of success, implementation should have been stage-by-stage, with plenty of time for the communication of its workings and purpose to all participants.

7.2 Performance-related pay in a large company

Next look at the Reader article ‘Performance-based management incentives at the Fiat Group: a field study’ by Kenneth Merchant and Angelo Riccaboni. This article, although written primarily for a purely academic/accountancy audience, is valuable for its wealth of detail. We recommend that you spend at least 45 minutes on it.

There is of course no need to learn every facet of the scheme described. It has not been selected as an example of particularly good or bad practice, but as a complex real life example for you to analyse in terms of the subjects discussed in this unit. It is also interesting because it makes public the details of a managerial as opposed to a blue collar incentive scheme.

Activity 11

List five strengths and five weaknesses of the scheme described. They might for example be concerned with problems of accuracy, such as when adjustments are made in the company’s favour; or ethics, such as the chances of deliberately inaccurate assessments, the range of targets on which managers are assessed, and assumptions about managerial motivation.

Comment

The article’s authors speculate that the Fiat scheme will need to be adjusted as the strategy of the group changes. This tension between long and short term objectives will be considered in more detail in Block 4.

Although the Fiat article concentrates on the effects of financial incentives, it tells us a great deal about the company’s views of its senior managers. The managers themselves are said to believe that significant performance-based incentives were an important and beneficial part of the restructuring process. As mentioned in Section 1 of this unit, researchers are starting to find good evidence that the widely held belief that ‘people matter’, and that therefore they should be rewarded and developed, does improve organizational performance when put into practice systematically.

Fox and McLeay (1992) and their team from the Universities of Lancaster and Wales studied 49 companies mainly in the engineering and electronics sectors. Annual financial performance indicators covering a period of over 10 years were used, rather than the broader range recommended by Eccles (Unit 1). Many senior managers were interviewed and the policies and practices they reported were independently assessed. The research team was especially careful to dis-
tistinguish between intent (the systems that existed, as found in policy documents, etc.) and practice (what actually happened from day to day in the company). The coherence of all the activities that were carried out as part of human resource management was judged, as well as the presence or absence of effective practice in the five sub-areas defined. These were (1) management education, training and development, (2) recruitment and selection, (3) performance appraisal, (4) rewards and remuneration, and (5) company-level career planning.

The most striking results were as follows.

- There was a clear positive relationship between financial performance and the degree of integration between the corporate strategy of an organization and its human resource management practices. For example, the ROCE (return on capital employed) of a firm with a high degree of integration between its human resource management function and its corporate strategy was likely to be substantially above the average within its sector.
- A policy of systematically rewarding good staff performance, when this was a main focus of the human resource management functions in practice, was significantly and positively correlated with ROCE.

### 7.3 Questions about assessment

By now you should realize that improving all forms of assessment of staff performance that contribute to human resource strategy is a worthwhile although daunting task. We have tried to be honest about just how difficult it is to do well, because many texts tend to gloss over the problems. It is always tempting to measure and assess anything that is simple, rather than those elements of performance that are relevant but more challenging to capture. In the interests of accuracy and ethics, we have to rise to this challenge, or the motivational aspects of assessment will work against instead of towards improved organizational effectiveness. Assessment is a powerful management tool that should never be taken for granted.

When thinking about better ways of assessment, you will realize that many contradictory ideas have contributed to management practice and that you cannot safely accept the word of any one expert who claims to offer the ideal performance management scheme. We suggest that you consider any scheme against the following two models, from each of which a set of questions can be derived. You will no doubt by now be able to suggest improvements and amendments to our questions.

### Model 1: The control loop

This by now familiar model can be used at any level in the organization to check how effectively feedback is being provided.

- Does the feedback from the assessment relate to issues over which the person or group being assessed has control?
- How accurately could anyone judge that type of performance against that type of target?
- Do the targets relate ultimately to the objectives of the organization?
- Is the feedback clear and constructive?
- Can the assessor be an effective actuator in this situation?
Does the assessor understand enough about the motivation of the assessed?
Do the people being assessed perceive the process to be clear, open and accurate?
Do the assessors value clarity, openness and accuracy?

**Model 2: Variety in systems**

The following questions concentrate on what a new scheme will add.

- Will it provide the organization with any valuable new information about staff performance?
- Will this scheme add to the motivation of all the staff involved?
- Who is going to provide the energy to get this scheme up and running?
- Does this scheme help to compensate for the limitations that bounded rationality imposes on human achievement?
- Will this scheme help to reduce unresolved internal conflict between work groups?
- Will it wake up the organization, or become lost because there are too many non-integrated change schemes happening already?
- Will the language used be understood, and will it trigger appropriately challenging mental models and metaphors?
References


Acknowledgements

Grateful acknowledgement is made to the following sources for permission to reproduce material in this unit.

Cover

Photo: Tony Stone Worldwide.

Text

*The Management Charter Initiative and The Employment Department* (1990), Management Charter Initiative.

Figures


Table