Chapter 13
The Dilemma of Performance Appraisal

Peter Prowse and Julie Prowse

Abstract This paper deals with the dilemma of managing performance using performance appraisal. The authors will evaluate the historical development of appraisals and argue that the critical area of line management development that was identified as a critical success factor in appraisals has been ignored in the later literature evaluating the effectiveness of performance through appraisals. This paper will evaluate the aims and methods of appraisal, the difficulties encountered in the appraisal process. It also re-evaluates the lack of theoretical development in appraisal and moves from the psychological approaches of analysis to a more critical realisation of approaches before re-evaluating the challenge to remove subjectivity and bias in judgement of appraisal.

13.1 Introduction

This paper will define and outline performance management and appraisal. It will start by evaluating what form of performance is evaluated, then develop links to the development of different performance traditions (Psychological tradition, Management by Objectives, Motivation and Development). It will outline the historical development of performance management then evaluate high performance strategies using performance appraisal. It will evaluate the continuing issue of subjectivity and ethical dilemmas regarding measurement and assessment of performance. The paper will then examine how organisations measure performance before evaluation of research on some recent trends in performance appraisal.

This chapter will evaluate the historical development of performance appraisal from management by objectives (MBO) literature before evaluating the debates between linkages between performance management and appraisal. It will outline
the development of individual performance before linking to performance management in organizations. The outcomes of techniques to increase organizational commitment, increase job satisfaction will be critically evaluated. It will further examine the transatlantic debates between literature on efficiency and effectiveness in the North American and the United Kingdom) evidence to evaluate the HRM development and contribution of performance appraisal to individual and organizational performance.

13.2 What is Performance Management?

The first issue to discuss is the difficulty of definition of Performance Management. Armstrong and Barron (1998:8) define performance management as:

A strategic and integrated approach to delivering sustained success to organisations by improving performance of people who work in them by developing the capabilities of teams and individual performance.

13.2.1 Performance Appraisal

Appraisal potentially is a key tool in making the most of an organisation’s human resources. The use of appraisal is widespread estimated that 80–90% of organizations in the USA and UK were using appraisal and an increase from 69 to 87% of organisations between 1998 and 2004 reported a formal performance management system (Armstrong and Baron, 1998:200). There has been little evidence of the evaluation of the effectiveness of appraisal but more on the development in its use. Between 1998 and 2004 a sample from the Chartered Institute of Personnel and Development (CIPD, 2007) of 562 firms found 506 were using performance appraisal in UK. What is also vital to emphasise is the rising use of performance

<table>
<thead>
<tr>
<th>Workplace size</th>
<th>Any non-managerial employees (%)</th>
<th>60% or more non-managerial employees (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10–24 employees</td>
<td>73</td>
<td>64</td>
</tr>
<tr>
<td>25–49 employees</td>
<td>65</td>
<td>60</td>
</tr>
<tr>
<td>50–99 employees</td>
<td>81</td>
<td>72</td>
</tr>
<tr>
<td>100–199 employees</td>
<td>82</td>
<td>75</td>
</tr>
<tr>
<td>200–499 employees</td>
<td>89</td>
<td>74</td>
</tr>
<tr>
<td>500 or more employees</td>
<td>91</td>
<td>81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sector ownership</th>
<th>Any non-managerial employees (%)</th>
<th>60% or more non-managerial employees (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private sector</td>
<td>68</td>
<td>62</td>
</tr>
<tr>
<td>Public sector</td>
<td>86</td>
<td>77</td>
</tr>
</tbody>
</table>

Source: Adapted from Table 4.5 Kersley et al. (2006) page 83.
appraisal feedback beyond performance for professionals and managers to nearly 95% of workplaces in the 2004 WERS survey (see Table 13.1). Clearly the use of Appraisals has been the development and extension of appraisals to cover a large proportion of the UK workforce and the coverage of non managerial occupations and the extended use in private and public sectors.

### 13.2.2 The Purpose of Appraisals

The critical issue is what is the purpose of appraisals and how effective is it researched and used in practice throughout organizations?

The purpose of appraisals needs to be clearly identified. Firstly their purpose. Randell (1994) states they are a systematic evaluation of individual performance linked to workplace behaviour and/or specific criteria. Appraisals often take the form of an appraisal interview, usually annual, supported by standardised forms/paperwork. The key objective of appraisal is to provide feedback for performance is provided by the line manager. The three key questions for quality of feedback:

1. What and how are observations on performance made?
2. Why and how are they discussed?
3. What determines the level of performance in the job?

It has been argued by one school of thought that these process cannot be performed effectively unless the line manager of person providing feedback has the interpersonal interviewing skills to provide that feedback to people being appraised. This has been defined as the “Bradford Approach” which places a high priority on appraisal skills development (Randell, 1994). This approach is outlined in Fig. 13.1 which identifies the linkages between involving, developing, rewarding and valuing people at work.

### 13.2.3 Historical Development of Appraisal

The historical development of performance feedback has developed from a range of approaches. Formal observation of individual work performance was reported in Robert Owens’s Scottish factory in New Lanark in the early 1800s (Cole, 1925). Owen hung over machines a piece of coloured wood over machines to indicate the superintendent’s assessment of the previous day’s conduct (white for excellent, yellow, blue and then black for poor performance). The twentieth century led to F.W. Taylor and his measured performance and the scientific management movement (Taylor, 1964). The 1930s Traits Approaches identified personality and performance and used feedback using graphic rating scales, a mixed standard of performance scales noting behaviour in likert scale ratings. This was used to recruit and identify
management potential in the field of selection. Later developments to prevent a middle scale from 5 scales then developed into a forced-choice scale which forced the judgement to avoid central ratings. The evaluation also included narrative statements and comments to support the ratings (Mair, 1958).

In the 1940s Behavioural Methods were developed. These included Behavioural Anchored Rating Scales (BARS); Behavioural Observation Scales (BOS); Behavioural Evaluation Scales (BES); critical incident; job simulation. All these judgements were used to determine the specific levels of performance criteria to specific issues such as customer service and rated in factors such as excellent, average or needs to improve or poor. These ratings are assigned numerical values and added to a statement or narrative comment by the assessor. It would also lead to identify any potential need for training and more importantly to identify talent for careers in line management supervision and future managerial potential.

Post 1945 developed into the Results-oriented approaches and led to the development of management by objectives (MBO). This provided aims and specific targets to be achieved and within time frames such a specific sales, profitability, and deadlines with feedback on previous performance (Wherry, 1957). The deadlines may have required alteration and led to specific performance rankings of staff. It also provided a forced distribution of rankings of comparative performance and paired comparison ranking of performance and setting and achieving objectives.

In the 1960s the development of Self-appraisal by discussion led to specific time and opportunity for the appraisee to reflectively evaluate their performance in the discussion and the interview developed into a conversation on a range of topics that the appraisee needed to discuss in the interview. Until this period the success of the
The dilemma of performance appraisal was dependent on skill of interviewer. In the 1990s the development of 360-degree appraisal developed where information was sought from a wider range of sources and the feedback was no longer dependent on the manager-subordinate power relationship but included groups appraising the performance of line managers and peer feedback from peer groups on individual performance (Redman and Snape, 1992). The final development of appraisal interviews developed in the 1990s with the emphasis on the linking performance with financial reward which will be discussed later in the paper.

13.2.4 Measures of Performance

The dilemma of appraisal has always to develop performance measures and the use of appraisal is the key part of this process. Quantitative measure of performance communicated as standards in the business and industry level standards translated to individual performance. The introduction of techniques such as the balanced scorecard developed by Kaplan and Norton (1992). Performance measures and evaluation included financial, customer evaluation, feedback on internal processes and Learning and Growth. Performance standards also included qualitative measures which argue that there is an over emphasis on metrics of quantitative approach above the definitions of quality services and total quality management. In terms of performance measures there has been a transformation in literature and a move in the 1990s to the financial rewards linked to the level of performance. The debates will be discussed later in the paper.

13.3 Criticism of Appraisals

Critiques of appraisal have continued as appraisals have increased in use and scope across sectors and occupations. The dominant critique is the management framework using appraisal as an orthodox technique that seeks to remedy the weakness and propose of appraisals as a system to develop performance.

This “orthodox” approach argues there are conflicting purposes of appraisal (Strebler et al, 2001). Appraisal can motivate staff by clarifying objectives and setting clear future objectives with provision for training and development needs to establish the performance objective. These conflicts with assessing past performance and distribution of rewards based on past performance (Bach, 2005:301). Employees are reluctant to confide any limitations and concerns on their current performance as this could impact on their merit related reward or promotion opportunities (Newton and Findley, 1996:43). This conflicts with performance as a continuum as appraisers are challenged with differing roles as both monitors and judges of performance but an understanding counsellor which Randell (1994) argues few managers have not received the raining to perform. Appraisal Manager’s reluctance to criticise also stems from classic evidence from McGregor that managers are
reluctant to make a negative judgement on an individual’s performance as it could be
demotivating, lead to accusations of their own support and contribution to individual
poor performance and to also avoid interpersonal conflict (McGregor, 1957).

One consequence of this avoidance of conflict is to rate all criterion as central
and avoid any conflict known as the central tendency. In a study of senior managers
by Longnecker et al. (1987), they found organisational politics influenced ratings of
60 senior executives. The findings were that politics involved deliberate attempts by
individuals to enhance or protect self-interests when conflicting courses of action are
possible and that ratings and decisions were affected by potential sources of bias or
inaccuracy in their appraisal ratings (Longenecker et al., 1987). There are methods
of further bias beyond Longenecker’s evidence. The political judgements and they
have been distorted further by over rating some clear competencies in performance
rather than being critical across all rated competencies known as the halo effect and
if some competencies are lower they may prejudice the judgment across the positive
reviews known as the horns effect (ACAS, 1996).

Some ratings may only include recent events and these are known as the recency
effects. In this case only recent events are noted compared to managers gathering
and using data throughout the appraisal period. A particular concern is the equity of
appraisal for ratings which may be distorted by gender, ethnicity and the ratings of
appraisers themselves. A range of studies in both the US and UK have highlighted
subjectivity in terms of gender (Alimo-Metcalf, 1991; White, 1999) and ethnicity of
the appraise and appraiser (Geddes and Konrad, 2003). Suggestions and solutions
on resolving bias will be reviewed later.

The second analysis is the radical critique of appraisal. This is the more critical
management literature that argues that appraisal and performance management are
about management control (Newton and Findley, 1996; Townley, 1993). It argues
that tighter management control over employee behaviour can be achieved by the
extension of appraisal to manual workers, professional as means to control. This
develops the literature of Foucault using power and surveillance. This literature
uses cases in examples of public service control on professionals such a teachers
(Healy, 1997) and University professionals (Townley, 1990). This evidence argues
the increased control of public services using appraisal as a method of control
and that the outcome of managerial objectives ignores the developmental role of
appraisal and ratings are awarded for people who accept and embrace the culture
and organisational values. However, this literature ignores the employee resistance
and the use of professional unions to challenge the attempts to exert control over
professionals and staff in the appraisal process (Bach, 2005:306).

One of the different issues of removing bias was the use of the test metaphor
(Folger et al., 1992). This was based on the assumption that appraisal ratings were a
technical question of assessing “true” performance and there needed to be increased
reliability and validity of appraisal as an instrument to develop motivation and
performance. The sources of rater bias and errors can be resolved by improved
organisational justice and increasing reliability of appraiser’s judgement.

However there were problems such as an assumption that you can state job
requirements clearly and the organisation is “rational” with objectives that reflect
values and that the judgment by appraisers’ are value free from political agendas and personal objectives. Secondly there is the second issue of subjectivity if appraisal ratings where decisions on appraisal are rated by a “political metaphor” (Hartle, 1995). This “political view” argues that appraisal is often done badly because there is a lack of training for appraisers and appraisers may see the appraisal as a waste of time. This becomes a process which managers have to perform and not as a potential to improve employee performance. Organisations in this context are “political” and the appraisers seek to maintain performance from subordinates and view appraisals as internal customers to satisfy. This means managers use appraisal to avoid interpersonal conflict and develop strategies for their own personal advancement and seek a quiet life by avoiding censure from higher managers. This perception means managers also see appraisee seeks good rating and genuine feedback and career development by seeking evidence of combining employee promotion and pay rise. This means appraisal ratings become political judgements and seek to avoid interpersonal conflicts. The approaches of the “test” and “political” metaphors of appraisal are inaccurate and lack objectivity and judgement of employee performance is inaccurate and accuracy is avoided. The issue is how can organisations resolve this lack of objectivity?

### 13.3.1 Solutions to Lack of Objectivity of Appraisal

Grint (1993) argues that the solutions to objectivity lies in part with McGregor’s (1957) classic critique by retraining and removal of “top down” ratings by managers and replacement with multiple rater evaluation which removes bias and the objectivity by upward performance appraisal. The validity of upward appraisal means the removal of subjective appraisal ratings. This approach is also suggested to remove gender bias in appraisal ratings against women in appraisals (Fletcher, 1999). The solution of multiple reporting (internal colleagues, customers and recipients of services) will reduce subjectivity and inequity of appraisal ratings. This argument develops further by the rise in the need to evaluate project teams and increasing levels of teamwork to include peer assessment. The solutions also in theory mean increased closer contact with individual manager and appraises and increasing services linked to customer facing evaluations.

However, negative feedback still demotivates and plenty of feedback and explanation by manager who collates feedback rather than judges performance and fail to summarise evaluations. There are however still problems with accuracy of appraisal objectivity as Walker and Smither (1999) 5 year study of 252 managers over 5 year period still identified issues with subjective ratings in 360 degree appraisals. There are still issues on the subjectivity of appraisals beyond the areas of lack of training.

The contribution of appraisal is strongly related to employee attitudes and strong relationships with job satisfaction (Fletcher and Williams, 1996). The evidence on appraisal still remains positive in terms of reinvigorating social relationships at work (Townley, 1993) and the widespread adoption in large public services in the UK such as the national health Service (NHS) is the valuable contribution to line managers...
discussion with staff on their past performance, discussing personal development plans and training and development as positive issues. One further concern is the openness of appraisal related to employee reward which we now discuss.

### 13.3.2 Linking Appraisals with Reward Management

Appraisal and performance management have been inextricably linked to employee reward since the development of strategic human resource management in the 1980s. The early literature on appraisal linked appraisal with employee control (Randell, 1994; Grint, 1993; Townley, 1993, 1999) and discussed the use of performance related reward to appraisals. However the recent literature has substituted the chapter titles employee “appraisal” with “performance management” (Bach, 2005; Storey, 2007) and moved the focus on performance and performance pay and the limits of employee appraisal. The appraisal and performance pay link has developed into debates to three key issues: The first issue is has performance pay related to appraisal grown in use? The second issue is what type of performance do we reward? and the final issue is who judges management standards?

The first discussion on influences of growth of performance pay schemes is the assumption that increasing linkage between individual effort and financial reward increases performance levels. This linkage between effort and financial reward increasing levels of performance has proved an increasing trend in the public and private sector (Bevan and Thompson, 1992; Armstrong and Baron, 1998). The drive to increase public sector performance effort and setting of targets may even be inconsistent in the experiences of some organizational settings aimed at achieving long-term targets (Kessler and Purcell, 1992; Marsden, 2007). The development of merit based pay based on performance assessed by a manager is rising in the UK Marsden (2007) reported that the:

Use of performance appraisals as a basis for merit pay are used in 65 per cent of public sector and 69 per cent of the private sector employees where appraisal covered all non-managerial staff (p.109).

Merit pay has also grown in use as in 1998 20% of workplaces used performance related schemes compared to 32% in the same organizations 2004 (Kersley et al., 2006:191). The achievements of satisfactory ratings or above satisfactory performance averages were used as evidence to reward individual performance ratings in the UK Civil Service (Marsden, 2007). Table 13.2 outlines the extent of merit pay in 2004.

The second issue is what forms of performance is rewarded. The use of past appraisal ratings as evidence of achieving merit-related payments linked to achieving higher performance was the predominant factor developed in the public services. The evidence on Setting performance targets have been as Kessler (2000:280) reported “inconsistent within organizations and problematic for certain professional or less skilled occupations where goals have not been easily formulated”. There has been inconclusive evidence from organizations on the impact of performance
Table 13.2 Establishments reporting use of “merit pay” for some employees (% of establishments in sector declaring use of merit pay)

<table>
<thead>
<tr>
<th>Weights</th>
<th>Private % merit pay</th>
<th>Public % merit pay</th>
<th>Private % merit pay</th>
<th>Public % merit pay</th>
<th>Private N(all establish)</th>
<th>Public N(all establish)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>13</td>
<td>–</td>
<td>27</td>
<td>–</td>
<td>210</td>
<td>–</td>
</tr>
<tr>
<td>Electricity, gas and water</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Construction</td>
<td>11</td>
<td>–</td>
<td>18</td>
<td>–</td>
<td>92</td>
<td>–</td>
</tr>
<tr>
<td>Wholesale and retail</td>
<td>13</td>
<td>–</td>
<td>22</td>
<td>–</td>
<td>461</td>
<td>–</td>
</tr>
<tr>
<td>Hotels and restaurants</td>
<td>17</td>
<td>–</td>
<td>20</td>
<td>–</td>
<td>161</td>
<td>–</td>
</tr>
<tr>
<td>Transport and communications</td>
<td>9</td>
<td>29</td>
<td>25</td>
<td>21</td>
<td>70</td>
<td>50</td>
</tr>
<tr>
<td>Financial services</td>
<td>29</td>
<td>–</td>
<td>43</td>
<td>–</td>
<td>95</td>
<td>–</td>
</tr>
<tr>
<td>Other business services</td>
<td>19</td>
<td>0</td>
<td>33</td>
<td>0</td>
<td>274</td>
<td>17</td>
</tr>
<tr>
<td>Public administration</td>
<td>–</td>
<td>6</td>
<td>–</td>
<td>28</td>
<td>–</td>
<td>97</td>
</tr>
<tr>
<td>Education</td>
<td>1</td>
<td>10</td>
<td>8</td>
<td>19</td>
<td>15</td>
<td>179</td>
</tr>
<tr>
<td>Health</td>
<td>11</td>
<td>11</td>
<td>9</td>
<td>22</td>
<td>105</td>
<td>156</td>
</tr>
<tr>
<td>Other community services</td>
<td>31</td>
<td>3</td>
<td>20</td>
<td>12</td>
<td>72</td>
<td>70</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>10</td>
<td>26</td>
<td>21</td>
<td>1557</td>
<td>589</td>
</tr>
</tbody>
</table>
pay and its effectiveness in improving performance. Evidence from a number of individual performance pay schemes report organizations suspending or reviewing them on the grounds that individual performance reward has produced no effect in performance or even demotivates staff (Kessler, 2000:281). More in-depth studies setting performance goals followed by appraisal on how well they were resulted in loss of motivation whilst maintaining productivity and achieved managers using imposing increased performance standards (Marsden and Richardson, 1994). As Randell (1994) had highlighted earlier, the potential objectivity and self-criticism in appraisal reviews become areas that appraisees refuse to acknowledge as weaknesses with appraisers if this leads to a reduction in their merit pay. Objectivity and self reflection for development becomes a weakness that appraises fail to acknowledge as a developmental issue if it reduces their chances of a reduced evaluation that will reduce their merit reward. The review of civil service merit pay (Makinson, 2000) reported from 4 major UK Civil Service Agencies and the National Health Service concluded that existing forms of performance pay and performance management had failed to motivate many staff.

The conclusions were that that employees found individual performance pay divisive and led to reduced willingness to co-operate with management, citing managerial favorites and manipulation of appraisal scores to lower ratings to save paying rewards to staff (Marsden and French, 1998). This has clear implications on the relationship between line managers and appraises and the demotivational consequences and reduced commitment provide clear evidence of the danger to linking individual performance appraisal to reward in the public services. Employees focus on the issues that gain key performance focus by focusing on specific objectives related to key performance indicators rather than all personal objectives. A study of banking performance pay by Lewis (1998) highlighted imposed targets which were unattainable with a range of 20 performance targets with narrow short term financial orientated goals. The narrow focus on key targets and neglect of other performance aspects leads to tasks not being delivered.

This final issue of judging management standards has already highlighted issues of inequity and bias based on gender (Beyer, 1990; Chen and DiTomasio, 1996; Fletcher, 1999). The suggested solutions to resolve discrimination have been proposed as enhanced interpersonal skills training are increased equitable use of 360 degree appraisal as a method to evaluate feedback from colleagues as this reduces the use of the “political metaphor” (Randell, 1994; Fletcher, 1999).

On measures linking performance to improvement require a wider approach to enhanced work design and motivation to develop and enhance employee job satisfaction and the design of linkages between effort and performance are significant in the private sector and feedback and awareness in the public sector (Fletcher and Williams, 1996:176). Where rises be in pay were determined by achieving critical rated appraisal objectives, employees are less self critical and open to any developmental needs in a performance review.
13.4 Conclusion

As performance appraisal provides a major potential for employee feedback that could link strongly to increasing motivation, and an opportunity to clarify goals and achieve long term individual performance and career development why does it still suffer from what Randell describes as a muddle and confusion which still surrounds the theory and practice?

There are key issues that require resolution and a great deal depends on the extent to which you have a good relationship with your line manager. Barlow (1989) argued `if you get off badly with your first two managers, you may just as well forget it (p. 515). The evidence on the continued practice of appraisals is that they are still institutionally elaborated systems of management appraisal and development is significant rhetoric in the apparatus of bureaucratic control by managers (Barlow, 1989). In reality the companies create, review, change and even abolish appraisals if they fail to develop and enhance organisational performance (Kessler, 2000). Despite all the criticism and evidence the critics have failed to suggest an alternative for a process that can provide feedback, develop motivation, identify training and potential and evidence that can justify potential career development and justify reward (Hartle, 1997).

References


