

The Crown Prosecution Service

Equality and Diversity Unit &
Human Resources Directorate

PERFORMANCE

APPRAISAL

RATINGS:

A Review and Impact Assessment

Actions taken since the impact assessment

The CPS has drawn upon the recommendations of the impact assessment of performance appraisal ratings to inform the more recent Invest programme and the performance development and review system. This has involved moving away from a box marking system to training on performance development and review for managers and staff and involvement of HR Business Partners in proactively supporting performance management.

1. INTRODUCTION

- 1.1. In line with many Civil Service departments, the *CPS Annual Equalities in Employment Report 2002–2003* highlighted the fact that some groups of employees were receiving less favourable performance appraisal review (PAR) box markings than others.
- 1.2. A review of the literature indicated that the question of differential PAR ratings in relation to Black and Minority Ethnic (BME) personnel, disabled staff, and women employees had previously been researched in a number of Civil Service and local government organisations. Previous research has invariably found that BME and disabled employees received less favourable PAR ratings. This research (in the Civil Service) also found that men received much lower PAR ratings than women did.
- 1.3. To date the Crown Prosecution Service (CPS) has not carried out an impact assessment of PAR ratings of key demographic groups within the Service. The current study, therefore, seeks to encourage enduring improvements in equality in employment by examining PAR ratings across a diverse range of employee groups.
- 1.4. This report represents the findings of the research, which has been carried out by independent equality and diversity consultants, Sankofa Exchange Limited.
- 1.5. We are heartened that this review found no significant evidence of discrimination in the application of the PAR process. We are however far from complacent. We are alert to the differences in perception of a range of employees, and to the differences in experiences of disabled employees. We will ensure that these issues are taken forward and addressed in 2006-07 and beyond as we both prepare our response to the disability equality duty and review our PAR process.

- 1.6. We intend to build on this review with future reviews of HR processes and practices in our efforts to continuously improve performance on equality and employment.

Séamus Taylor
Director of Equality and Diversity
The Crown Prosecution Service
15 March 2005

Angela O'Connor
Human Resources Director
The Crown Prosecution Service

2. EXECUTIVE SUMMARY

Introduction

- 2.1. This report sets out the findings of the research into PAR box markings in the CPS. The research forms part of the Service's overall programme on equality and diversity.

EXECUTIVE SUMMARY

- 2.2. In the first phase of the research, PAR box markings were analysed to determine whether any significant differences in PAR box markings existed in the CPS. The *Annual Equalities in Employment Report 2002–2003* suggested that some differences existed with regard to ethnicity and disability status. However, it was impossible to determine from the information available whether these differences were consistent enough to have a potentially adverse impact on the employment experiences of particular demographic groups.
- 2.3. Irrespective of whether the findings actually did demonstrate differences in PAR box markings that could not be explained by chance, the Service is committed to improving equality in employment for all its employees.
- 2.4. The first stage of this impact assessment was a quantitative analysis of PAR reports. The first part of the quantitative analysis was informed by a detailed examination of a random sample of PAR reports. The output of this analysis was summarised and further analysed using a range of statistical techniques. This examination of paper PAR reports was supported by detailed high-level statistical analyses of the whole collection of PAR reports for the Crown Prosecution Service for the years ending 2003 and 2004. In total, 5,030 records were analysed. The data were analysed to determine whether any differences existed with regard to gender, ethnicity, or disability status.
- 2.5. The second phase of the impact assessment consisted of 20 depth interviews, focus groups and a snapshot survey of 131 employees. Given the self-

selecting nature of the samples and the number of respondents obtained in phase two, less reliance can be placed on the findings of that stage. Some of the samples in the second stage of the review were not fully representative and so may reflect the views of those who had the most negative (or positive) perceptions about the PAR system.

- 2.6. The Service is committed to change and will act on the findings of the impact assessment to make real improvements in employment practices.

Main findings of the CPS PAR ratings impact assessment

- 2.7. **The review found no cogent evidence of discrimination. The overall key finding is that although the impact assessment found no evidence of a statistically significant difference in PAR box mark ratings between different staff groups, it did identify a concern about how the current system PAR system is perceived amongst all groups. However, almost half of the respondents felt their PAR report was an accurate reflection of their performance. A difference in PAR ratings on the basis of disability was found. This may be explained by other factors - such as working patterns, a greater number of small objectives being agreed by this group, lower levels of job satisfaction – and would benefit from further investigation.**
- 2.8. **The other key findings were:**
- (a) no statistically significant differences in PAR ratings on the basis of gender and ethnicity were found. This contrasts with the Civil Service/public sector trend on PARs;**
 - (b) almost half of respondents felt their PAR report reflects their performance;**
 - (c) a majority of respondents did not feel the need to ask for support in achieving their objectives;**

- (d) **a majority of respondents intended to stay with CPS for the next 12 months;**
- (e) **respondents placed significant weight on the potential contribution of the new HR Business Partners in helping to improve performance.**

2.9. *Key areas for further consideration, review and/or improvement:*

- (a) the limited value added by the Countersigning Officer (CSO) role;
- (b) the limited value of having a box rating system - with gravitation towards average ratings;
- (c) the perception of unfairness in the present system.

Comparisons with other Civil Service departments:

2.10. The current impact assessment should be viewed in the context of the Civil Service study of PAR ratings carried out by the Institute of Employment Studies in 2001 (IES, 2001). The main findings of the Civil Service study were as follows:

- (a) men, BME, and disabled employees all tended to receive relatively unfavourable PAR box markings;
- (b) unfavourable markings were not driven by ability, but were determined by other factors;
- (c) evidence of cultural bias against particular groups was found;
- (d) operation of an in-group/out-group culture existed in the Civil Service;
- (e) both managers and staff were dissatisfied with the lack of objectivity in the PAR system.

2.11. The findings were a useful benchmark for the CPS to use when carrying out its own impact assessment of its PAR process. The detailed findings of the CPS impact assessment are set out below.

Highlights of the findings of phase one: quantitative impact assessment

Analysis of PAR files

- 2.12. The analysis of PAR files compared, by demographic group (gender, disability and ethnicity), the PAR box markings for post-holders at the same pay level.
- 2.13. Each employee's PAR report was compared with regard to the number of objectives set, the number of objectives achieved, and whether feedback intended to improve performance was offered. Finally, the PAR reports were also compared to determine whether the diversity competence was consistently assessed by the appraiser.
- 2.14. The analysis of the files found that disabled employees had a large number of small objectives in comparison to their non-disabled counterparts who had a few major objectives. The main disadvantage associated with this higher number of objectives was that they are more difficult to achieve. There was a negative relationship between the number of objectives set and PAR box marking. The higher the number of objectives the higher the PAR box marking. A high box marking (5 – 4) indicates substandard performance and a low box marking (1 – 2) is indicative of above average or excellent performance.
- 2.15. The finding that disabled employees agree more objectives than their non-disabled counterparts does not necessarily signal discrimination.
- 2.16. Phase two of the impact assessment examined these differences in the number of objectives in more detail.
- 2.17. Although there were some variations in the number of targets set for women (versus men) and BME (versus White) employees, no statistically significant pattern was found.

- 2.18. To determine whether or not the variations found could be explained by chance and to examine these differences in more detail, the most salient variations identified in the reports were subjected to statistical analysis.
- 2.19. Statistical analysis of the reports confirmed that no statistically significant differences were found in respect of the number of objectives set or achieved by ethnic group or gender.
- 2.20. Statistically significant differences were found in respect of the number of objectives set and achieved by disability status.
- 2.21. The results of the statistical analysis of PAR files suggest that the more objectives that are set, the more difficult it is to achieve these objectives.

Quantitative impact assessment of PAR rating data

- 2.22. A complete census of all electronic PAR report summaries was carried out. The electronic summaries of PAR reports did not contain the detailed summaries that were found in the paper PAR reports.
- 2.23. The census of all PAR records held by the CPS found no statistically significant differences in respect of the PAR ratings of BME and White employees when all employee PAR records were examined.
- 2.24. A detailed review of the PAR records of lawyers was undertaken. No statistically significant differences in respect of PAR ratings of BME and White lawyers were found.
- 2.25. The impact assessment did not find any statistically significant differences between the PAR ratings of men and women. However, disabled employees received less favourable PAR ratings than their non-disabled counterparts.

- 2.26. A review of the performance appraisal literature suggested that factors such as staff satisfaction may impact on performance appraisal ratings. This is because staff who are least satisfied with their jobs are least likely to invest sufficient effort in achieving their objectives.
- 2.27. The data available from the CPS 'Your Voice' staff survey 2004 allowed a limited analysis of the relevant employee satisfaction data. Exploration of the results found that disabled employees were most negative about the PAR process, the amount of variety in their jobs and did not feel that their jobs gave them a sense of accomplishment. Disabled employees reported a lower level of motivation towards their jobs than other demographic groups. Finally, disabled employees were most negative about job security and were more neutral than other employee groups about their intention to stay in the CPS. **Further research is required in order to understand these differential levels of job satisfaction and the less favourable PAR ratings of disabled employees. HR and the EDU should prioritise this further work in 2006–2007 in anticipation of the new disability equality duty.**
- 2.28. **When the overall results of this impact assessment are benchmarked against the published literature on performance appraisal ratings, the CPS performed much better than its counterparts in the public and private sectors. There was no statistically significant adverse impact on BME employees, women or men. Although improvements to the PAR system have been identified, phase one of this impact assessment suggests that the Service is making progress in achieving equality in employment.**

Highlights of the findings of phase two: qualitative impact assessment

- 2.29. Phase two of this report was a detailed qualitative analysis of the evidence as part of our equality and diversity impact assessment of the PAR process.

Performance Appraisal Ratings: a Review and Impact Assessment

- 2.30. The Service is committed to providing effective feedback for employee development and fairly managing the performance of all employee groups.
- 2.31. The Service has a performance appraisal and review system and is committed to identifying any potential for discrimination in the system and ensuring that the process is seen and experienced as fair to all and driving through change processes which are developmental and objective.
- 2.32. In spite of several criticisms of the current PAR system, the findings of the phase two survey (n=131) showed almost half (44%) felt that their last PAR rating accurately reflected their performance and most respondents intended to stay within the organisation (just 26% expressed any intention to leave).
- 2.33. The overwhelming finding of the second part of the study is that the vast majority of employees (88%) considered the PAR rating method to be an ineffective technique.
- 2.34. Further, some 87% of employees felt that the PAR process was not an effective way of recognising individuals who perform well. A survey using a mainly quantitative questionnaire found that some 84% of employees considered that the Service had not been successful in introducing a performance culture. The survey results showed that some 77% of employees considered the PAR rating method to be unfair – a message that was a consistent theme of the associated focus groups and depth interviews.
- 2.35. The Service could do more to ensure that there is adequate feedback about job performance throughout the year, as some 65% of employees felt that they had inadequate feedback.
- 2.36. The first phase of the research identified a disparity in the number of objectives agreed and achieved by disabled employees. That study found that disabled employees agreed a higher number of objectives than their non-disabled counterparts. It is interesting to note that 69% of respondents felt that it was

better to have a few major objectives rather than many minor objectives. However, this finding needs to be considered in light of the clear benefits articulated by some participants in having a larger number of small manageable objectives.

- 2.37. Some 62% of employees felt that the appraisal system could be made more objective; a number of suggestions for doing this were made - such as more systematic collection of evidence to support progress against work objectives and better use of any rating scales.
- 2.38. A content analysis of participant responses in the free response section of the questionnaire revealed that:
- (a) staff were mainly dissatisfied with the removal of rewards such as pay and promotion from the PAR system;
 - (b) there was a very strong feeling about the failure of the CPS to effectively manage poor performance;
 - (c) many of the employee concerns about the performance appraisal system can be traced to their dissatisfaction with the box marking system. In particular employees believed the system to be ineffectively operated as the overwhelming majority of employees obtained box 3 performance appraisal rating.
- 2.39. Overall employees at pay level A and B were more satisfied with the appeals mechanisms and the effectiveness of the PAR rating method in recognising performance than employees at levels C to E.
- 2.40. Employees who described their sexuality as heterosexual were more likely to regard the PAR rating method as accurately reflecting their performance than employees who did not disclose their sexuality.

Performance Appraisal Ratings: a Review and Impact Assessment

- 2.41. It is also interesting to note that Countersigning Officers were more likely than non-Countersigning Officers to prefer a few major objectives and were more likely to say that the CPS had been successful in introducing a performance culture.
- 2.42. A statistically significant finding of the survey was that older people were more likely to report discomfort with providing written evidence to support their PAR review. The depth interview and focus group results also indicated that the African Caribbean male participant and the disabled woman expressed discomfort in providing evidence to support their PAR review.
- 2.43. New recruits were more likely to express a desire to leave the Service in 12 months than those employees who had been employed by the Service for longer than 12 months.
- 2.44. Whilst most employees considered that their PAR rating accurately reflected their performance, two of the African Caribbean participants felt that their PAR rating under rated their performance. Finally, the disabled woman and the woman engaged in on a part-time contract expressed most dissatisfaction about the number of objectives they agreed and the ability to renegotiate or obtain support in achieving these objectives.
- 2.45. There were a number of examples of excellent practice where managers were engaged in coaching ineffective employees, motivating employees by awarding high box markings and quality assuring PAR reports across whole Areas. There was also a demand from managers and willingness from the HR Directorate for the HR Business Partners to effectively challenge poor performance at every level of the organisation.

Conclusion

- 2.46. **We conclude that there is no evidence of a significant adverse impact against any particular demographic group arising out of the operation of the current PAR system. However, the evidence that disabled employees**

were more likely to receive lower PAR ratings and have more objectives should be explored further as a priority and steps taken by EDU and HR to ensure fair treatment in the PAR process.

3. CURRENT RESEARCH ON EQUALITY IN PAR RATINGS

Introduction

- 3.1. This section considers some of the key research literature on equality in performance appraisal ratings. The literature reviewed indicates that some demographic groups tend to receive less favourable performance appraisal ratings than other demographic groups. This unfavourable rating effect is prevalent in a number of organisations, as confirmed by numerous studies in the public and private sectors. The processes leading to less favourable PAR ratings are considered. Benchmarks for determining adverse impact are also discussed.

THE CIVIL SERVICE 2001 STUDY OF EQUALITY IN PAR RATINGS

- 3.2. The main findings of the Civil Service study on performance appraisal box markings provides a very useful overview of the problems experienced with the PAR system in a number of government departments (the Cabinet Office, DETR, DSS, MAFF, Inland Revenue and the Department of Health). The study also provided a useful framework in which to analyse the PAR system in the CPS and to compare outcomes.

Main findings of the IES Civil Service study

- 3.3. In 2001, the Institute of Employment Studies carried out a review of performance appraisal box markings in a number of key Civil Service departments and agencies.
- 3.4. The IES review of PAR ratings involved an analysis of more than 180,000 records, performance appraisal files, and in-depth interviews with some 470 employees. The sample included disabled staff, women, White men, managers, and BME employees. The IES (2001) study found statistically significant differences in PAR box markings in relation to ethnicity, gender and disability.

- 3.5. The IES (2001) research concluded that in some cases negative attitudes may adversely influence both performance and performance ratings. IES also found that some managers held negative attitudes towards BME and disabled employees. These perceptions adversely affected their people management skills, by inadvertently allowing their negative attitudes to influence their assessment of BME and disabled employees' performance.
- 3.6. A further factor affecting performance ratings was the lower level of job satisfaction and organisational commitment exhibited by disabled and BME employees, which may have affected their performance.
- 3.7. The IES (2001) study explored the application of the performance appraisal process through focus groups. The key themes that arose from discussions in focus groups made up of BME staff, women, disabled employees, White men, and managers were:
- (i) inconsistent application of performance measures;
 - (ii) lack of countersigning officer credibility;
 - (iii) overly complex PAR process;
 - (iv) unhelpful box marking classifications;
 - (v) unhelpful guidance and support;
 - (vi) use of a perceived quota system.
- 3.8. The IES study of Civil Service departments found that men, BME and disabled employees all received unfavourable PAR box markings when compared with White women. Unfavourable markings were not driven by ability, but were determined by other factors. There was evidence of cultural bias against particular groups. Furthermore, an in-group/out group culture existed. Both managers and staff were dissatisfied with the lack of objectivity in the PAR system.

A SUMMARY OF THE LESS FAVOURABLE PAR RATINGS LITERATURE

3.9. The process by which particular demographic groups obtain low PAR ratings appears to be particularly complex. However, it was possible to construct a theoretical model from the literature and discussions with the CPS project advisory group. The model comprises several key elements: under-or over-management, lack of training, stress linked to discrimination, manager's perception, and stress linked to cultural restrictions.

Under- or over-management

3.10. Over-management can sometimes occur in organisations when managers more actively manage a member of staff belonging to particular demographic group. The reason for this over-management may stem from mistrust or anxiety about the employee's abilities. Under-management is often related to a fear of effectively managing the employee, in the belief that if the employee's performance is questioned, a complaint of discrimination could be lodged.

3.11. A culture of taking the line of least resistance in an organisation will also result in under-management of employees. Challenging unacceptable performance may be unattractive to managers, as considerably more work is required to justify a box 4 or 5 marking than is required to justify a performance box score of 3 or above. Unfavourable performance appraisal box markings are more likely to be appealed against by employees.

Lack of training

3.12. Lack of effective training for both managers and employees has been identified by a number of studies (see, for example, Hatrup, Rock, & Scalia, 1997; Hedge & Kavanagh, 1988; IES, 2001). This lack of training relates mainly to the interpersonal skills needed to effectively carry out the performance appraisal review and also the provision of training for employees in the areas of development identified in their performance appraisal reviews.

Stress linked to discrimination

- 3.13. There is now a considerable body of literature which clearly demonstrates that discrimination may result in occupational stress (Cooper & Melhuish, 1984; Davison, 1997; Davison & Cooper, 1992, 1983). The particular stressors which have been isolated in this research are stress from:
- the strains of coping with prejudice and stereotyping;
 - overt and indirect discrimination from fellow employees;
 - other employees, the organisational structure and climate;
 - lack of role models and feelings of isolation;
 - the burden of coping with the role of being seen as a token.
- 3.14. Dewberry (2001) examined the differential pass rates of BME and White bar vocational course (BVC) students. After controlling for several variables which could have confounded the results, Dewberry erroneously concluded that BME bar final students performed less well than their White counterparts because BME bar final students lacked the ability to perform better.
- 3.15. While noting the research which shows that BME employees tend to consistently receive lower performance ratings than their White counterparts, (Ford, Kraiger, & Schechtman, 1986; Landau, 1995; Waldman & Avolio, 1991), Dewberry (2001) argues that this difference in bar final examination marks is due to a lack of ability of BME student lawyers.
- 3.16. Dewberry's (2001) research does have a number of limitations. One important factor not taken into account was the perceptions about fairness of the BME bar final students. If the perception of these students was that their examiners would discriminate against them, this could have adversely affected their performance (Greenberg, 1986). Furthermore, negative perceptions about one's ability can result in a self-fulfilling prophecy, where students or employees do not perform to the best of their ability because they do not believe that excellent performance is expected of them.

3.17. Fear of racist testers rather than lack of ability can contribute to the poorer performance of BME students and employees. Mishra (1980) found that performance on tests was influenced by the ethnicity of those who administered the tests. This finding was also supported by Terrell, Terrell, and Taylor (1981), who examined the relationship between fear of racism and subsequent scores on the Wechsler Adult Intelligence Scale. Terrell *et al.* (1981) found that when individuals doubted whether they would be treated fairly because of their ethnicity, their performance on the intelligence test suffered. These findings suggest that factors other than the ability of BVC students or of employees may explain differential performance ratings.

Manager's perception

3.18. The literature reviewed suggests that manager's perception manifests itself in the form of in-group and out-group attitudes and negative stereotypes about the performance of particular demographic groups. The studies of Ilgen and Youtz (1986), Andrew (1996) and Dansereau, Graen, and Haga (1975) all support this view.

3.19. Tackey, Tamkin, and Sheppard's (2001) review of the available literature on BME employees and performance appraisal ratings is based on the premise that BME groups experience greater disadvantage in employment in the labour market. Within the broad parameters of such disadvantage, they have drawn on other available research evidence (Jones, 1993; Madood *et al.*, 1997; TUC, 1999, 2000). Tackey *et al.* (2001) observed that the widespread and persistent nature of the problem, despite much variety in performance management, would tend to suggest that there are other processes at work other than the inherent deficiencies in the PAR systems used.

3.20. Studies based on nationally representative samples (Baldi & McBrier, 1997) conclude that there is clear rater bias on the basis of ethnicity; BME employees were significantly less likely to be promoted than White employees with similar levels of education and work experience. This is often reinforced by institutional

racism, whereby BME managers also rate White employees above their BME counterparts. Cianni and Romberger (1995) came to similar conclusions and identified a perceived denial of development opportunities by supervisors and managers on the basis of ethnicity and gender as the cause of the difference in PAR ratings. Baldi and McBrier (1997) have identified this as a form of discrimination which has replaced direct discrimination. Pulakos and Wexley (1983) examined gender and performance ratings and found that managers give more favourable ratings to subordinates who they perceive as being similar to themselves. Kraiger and Ford (1985) also found a rater bias based on ethnicity.

(i) *Cognitive structures*

- 3.21. Managers may look for information and evidence to reinforce broader stereotypes based on race and gender (Ilgen & Youtz, 1986).
- 3.22. The second form of (attribution) bias results from what are perceived as the reasons for good performance. Good performance by BME employees is often attributed to luck or extraordinary effort (Greenhaus & Parasuraman, 1993) rather than abilities and talents.
- 3.23. Rick, Tamkin, Pollard, Tackey, and Heron (1999) found that White raters and BME raters tend to use distinct and different types of competencies when describing the competence of individuals:
- White raters use individual management skills, their knowledge of the organisation, as well as competencies from a cognitive skill cluster;
 - BME raters, by contrast, are more likely to describe performance in terms of organisational skills, time management, team working, and a self-management cluster of competencies.

(ii) In-groups and out-groups

3.24. Two types of groups within organisations are identified by Alderfer (1986): identity groups and organisation groups. Organisation groups are identified by task, function or hierarchy, whilst identity groups are characterised by other characteristics such as age, sexuality, gender and ethnicity.

3.25. Work by Dansereau *et al.* (1975) helps to explain the impact of 'in' and 'out' groups. Dansereau developed a model which suggests that managers assign subordinates to an 'in-group' or an 'out-group' and behave differently towards these two groups. In-groups are treated more favourably, whilst out-groups are managed in an authoritarian style and through contract compliance.

Stress linked to cultural restrictions

3.26. There is a growing recognition that organisational culture can contribute to exclusion and discrimination, despite the existence of equal opportunities and procedures (Tackey *et al.*, 2001).

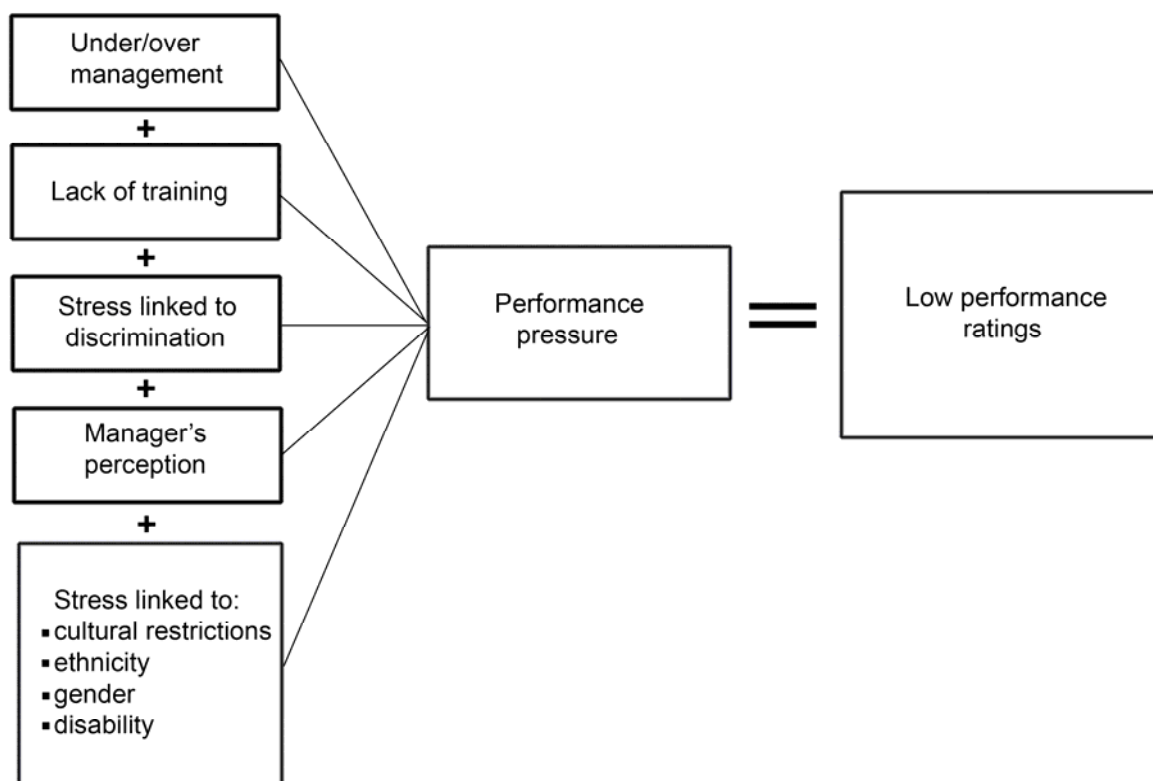
3.27. Connor, La Valle, Tackey, and Perryman (1996) have defined organisational culture in terms of shared symbols, language, practices (how we do things here), and deeply embedded beliefs and values. They argue that organisational culture can represent an impediment to change because culture is difficult to detect or pin down and is therefore difficult to challenge.

3.28. Two issues arise from this:

- Whether the employer's policies address statistical under-representation and the quality of candidates they approach;
- Whether the organisation's culture is such that women, disabled and BME employees react negatively in rebellion against the lack of opportunity, of which underperformance is an outward manifestation.

- 3.29. The final major area impacting on performance appraisal ratings relates to the impact of the potential lack of acceptance of cultural diversity within the workplace. In such situations it is contended that employees who appear different from the in-group are likely to suffer additional occupational stress caused by the lack of acceptance of cultural diversity. Whilst this is a new and innovative area of work, there is now a growing body of research which confirms this effect.
- 3.30. Leong (2001) examined the effect of acculturation on job satisfaction, occupational stress and performance appraisal ratings in a sample of Asian workers in the USA. Leong found that job satisfaction, occupational stress and job performance ratings varied according to how strongly Asian workers held onto their traditional cultural values. 'Traditional' Asian-Americans were more likely to experience higher levels of occupational stress and lower levels of job satisfaction compared with 'assimilationist' (more Westernised) and 'integrationalist' Asian-Americans.
- 3.31. Leong's (2001) research also examined the relationship between acculturation and performance appraisal ratings. Leong found that employees with high levels of acculturation (more Westernised) gained higher performance appraisal ratings than the employees with low levels of acculturation.
- 3.32. Leong's (2001) work was extended by Horsford's (2003) work in the UK. Horsford examined the relationship between career progression and occupational stress linked to cultural restrictions. His study examined BME and White managers across all industry sectors and found that for some BME managers, high levels of occupational stress were related to maintaining a strong cultural identity. Low levels of occupational stress were linked to maintaining a Western cultural identity and in-group attitude. As can be seen from Figure 3.1, it is suggested that a number of key factors lead to performance pressure and consequently to low performance appraisal ratings.

Fig 3.1: Causes of low performance ratings.



THE IMPACT OF COMPETENCIES ON PAR BOX MARKINGS

- 3.33. A competence framework may adversely affect the performance appraisal box markings when it can be established that the relevant competencies disadvantage a particular demographic group.

Competencies and performance assessment

- 3.34. Eichinger (2000) defines competencies as individual performance behaviours that are observable, measurable, and critical to successful individual or corporate performance. Eichinger identifies three prerequisites for successful implementation of competencies in organisations for HR purposes.

3.35. These three prerequisites for successful implementation are that the competencies:

- (i) can be seen;
- (ii) can be measured;
- (iii) are connected to what is deemed to be important.

3.36. There are a number of problems with competence-based performance appraisal systems:

- (i) The language used to describe competencies may be inaccessible. Competencies may be described in masculine terms or in a way that appears biased towards or against a particular ethnic group.
- (ii) Competencies are based on what good performers have done in the past; this approach does not allow the organisation to respond to rapid change. This approach may produce clones rather than a team with a mixture of skills who complement each other.
- (iii) Because of earlier discrimination against certain groups in society, models used for developing competencies tend to exclude the competencies particular to women, disabled people, or Black and Minority Ethnic employees. These groups are consequently disadvantaged by the institutional discrimination in mainstream competence frameworks.
- (iv) Some organisational competencies may also fail to effectively recognise more diverse leadership skills and in particular the transformational leadership styles used by many women. A number of competencies may measure what are essentially masculine or in-group traits. This means that members of the out-group who do not necessarily possess these traits are at a significant disadvantage.

ADVERSE IMPACT AND STATISTICAL SIGNIFICANCE

3.37. Two benchmarks for identifying potential discrimination in the PAR process were identified. The first benchmark was the use of statistical significance. Statistical significance is based on the probability that the results are not due to chance or random error. The second benchmark applied is the adverse impact model; this model was developed in the USA and is essentially based on differences between the actual proportions and an absolute threshold. The threshold in this model is where the success rate of one particular group is 80% or more, while the success rate of the comparator group is 20% or less.

Use of statistical significance to monitor discrimination

3.38. In statistics, 'significant' means 'probably true' (i.e. not arising due to chance). A research finding may be true without being important. A result which is 'highly significant' means that it is 'very probably true'.

3.39. Significance levels indicate how likely a result is due to chance. The benchmark used to mean that something is good enough to be believed is .95. That is, the finding has a 95% chance of being true. Put another way, the finding, therefore, has a 5% (.05) chance of not being true, the converse of a 95% chance of being true. (This normally appears in reports as, $p < .05$.)

3.40. The current study uses the following accepted benchmarks to refer to results and to assess how confident one can be in accepting the results. Not significant (*ns*) means that $p > .05$. Significant is defined as $p < .01$ to $p < .05$; a result that is very significant is shown as $p < .001$ to $p < .01$; and a finding that is extremely significant will be shown as $p < .001$ or lower.

The use of the adverse impact model to monitor discrimination

- 3.41. The US standard for determining adverse impact is drawn from guidelines issued in the USA – *The Uniform Guidelines for Employee Selection* (1978). The *Uniform Guidelines* include a basic statistical formula, described as the ‘four-fifths rule’. Adverse impact analysis is a very effective way of looking at general monitoring data, especially in large organisations.
- 3.42. Where a particular demographic group has less than four-fifths (or 80%) of the success rate (in comparison with) the group with the highest success rate, this will generally be regarded as evidence of adverse impact.
- 3.43. For example, an employer may have processed 200 job applications from 100 BME and 100 White applicants. Out of these 200 applicants, 100 were hired (80 Whites and 20 of BME origin). Thus the percentage of Whites hired is 80% and the percentage of BME applicants hired is 20%. The acceptance rate for BME candidates is only one-quarter of that of Whites (20 versus 80) and this would thus constitute an adverse impact.
- 3.44. A finding of adverse impact does not equate to a finding that discrimination has necessarily taken place. It provides a basis for examining particular practices, processes or actions.

Key points

- 3.45. In considering the benchmarks to be used in the review, the following points should be noted:
- (a) two models of impact-assessing the allocation of PAR box markings were available to us;
 - (b) where possible, both impact assessment models were applied;
 - (c) in most instances the statistical significance model was applied, as this was the most effective measure to use;

- (d) **whilst some findings may not reach statistical significance or reveal an adverse impact, these findings may nevertheless prove to be important areas for further exploration;**
- (e) particular demographic groups are more likely to experience stress linked to discrimination, and this stress may adversely affect their performance;
- (f) perceptions in terms of stereotyping and in-group attitudes may partly explain differences in performance ratings;
- (g) cultural restrictions based on gender, ethnicity, sexuality, age or disability may result in increased stress and lower performance
- (h) employee perceptions of unfairness may result in lower performance through a process of self-filling prophesy.

4. FRAMEWORK FOR ANALYSIS OF PAR REPORTS

Introduction

- 4.1. This section sets out the framework for evaluation of the paper PAR reports. The PAR reports considered in this impact assessment were drawn from a random sample of available PAR reports. Incomplete and illegible reports were removed for the analysis. In analysing the PAR reports, consideration was given to the findings identified from the Institute of Employment Studies (2001) review of Civil Service PAR box markings. Where possible, reports were matched on the key demographic variables of interest.

FRAMEWORK FOR ANALYSIS OF PAR REPORTS

- 4.2. Completed performance appraisal reports were evaluated using the CPS performance appraisal system guidelines as a framework for evaluation. In addition to this, a research technique known as frame analysis was used; this involved applying the key themes identified in the IES (2001) study to the impact assessment of the CPS performance appraisal process.
- 4.3. Countersigning officers' (CSO) comments were examined. In particular the study examined whether the CSO had commented in any meaningful way in the reports. This measure was particularly relevant when it is obvious from the face of the report that the CSO should have made detailed comments on the report.
- 4.4. The reports were reviewed to examine whether employees from one particular demographic group needed to achieve relatively more objectives in order to obtain the same PAR box rating than an employee in another demographic group.

- 4.5. The PAR reports were assessed for any obvious cultural bias. In particular, the reports were analysed to see whether there were comments which suggested cultural bias against a particular group.
- 4.6. The reports were also assessed to identify whether the standards set were consistent across all demographic groups.
- 4.7. In reviewing the PAR reports it was noted that the structure of the PAR report makes it difficult for managers to give effective feedback about performance. This is because the report asks managers to set out the employee's performance under each competence heading. The form has insufficient space to clearly advise employees about where their performance is below standard. While some of the areas for development may be addressed in the employee's forward plan, it would be more useful if the main areas for improvement were included in the PAR report itself.

5. ANALYSIS OF PAR FILES

Introduction

- 5.1. This section sets out the details of the qualitative analysis of the PAR reports. The analysis compares, with respect to demographic group (gender, disability and ethnicity), PAR box markings when the pay level of the post-holder is held constant. The reports are compared by the number of objectives set, the number of objectives achieved, whether feedback to improve performance was offered, and whether the diversity competence was assessed.

RESULTS OF QUALITATIVE ANALYSIS OF PAR FILES

- 5.2. A total of 49 PAR reports were available, for analysis. The sample consisted of eight PAR reports from BME men and eight PAR reports from BME women. Of the PAR reports available for review, 15 were from White women and 18 were from White men.
- 5.3. The comparisons of PAR reports were made by box markings. Box 1 is the highest box marking and means the employee's performance is classified as outstanding. Box mark 2 represents performance significantly above the standard required. A box mark 3 represents performance which fully meets the standard required. Whilst a box mark 4 represents performance which does not meet the standard required. A box mark 5 is awarded for an unacceptable standard of performance.
- 5.4. Within each demographic group PAR reports were also compared by pay band. Level A represents the lowest pay band consisting of two pay levels A1 and A2. Level B includes pay levels B1, B2 and B3. B3 is the highest pay level within this pay band. Level C consists of pay levels C1 and C2 and is used exclusively for lawyers. Level D is followed by level E, these pay levels are generally used for non-lawyer staff. Chief Crown Prosecutors and Directors have their own Performance Appraisal Scheme and a different pay structure. Their reports were not considered in this impact assessment.

Disability

Box 1, B1 PAR reports of disabled and non-disabled women compared

- 5.5. There were no differences in the number of objectives set and achieved by the disabled and non-disabled women. Neither of the participants had any detailed feedback about how their performance could be improved. However, the non-disabled woman's report referred to performance against the equality and diversity competence.

Box 2, A2 PAR reports of disabled and non-disabled women

- 5.6. Two disabled women at pay level A2 received a box 2 PAR rating. Their PAR ratings were compared with that of a non-disabled woman who gained a box 2 PAR rating. Seven objectives were set for the disabled employees. This compares to six objectives for the non-disabled employee. All objectives were achieved by all the employees. Achievements in relation to the diversity competence were discussed with one of the disabled employees; however, there is no evidence for any corresponding discussion with the non-disabled employee. None of the employees received feedback on the areas where their performance needed to improve.

Box 2, B1 PAR reports of disabled and non-disabled men compared

- 5.7. The PAR reports of two male B1 employees were examined for any differences. Ten objectives were set for the disabled employee – eight of which were achieved. This was higher than his non-disabled comparator (who had seven objectives set, all of which were achieved). No specific feedback was apparent from the disabled employee's report; however, some feedback on how to improve performance was noted on the form of the non-disabled employee. The diversity competence was mentioned in performance appraisal reports of both employees.

Box 3, B1 PAR reports of disabled and non-disabled men with compared

- 5.8. There were some differences in the number of objectives set for the disabled man and his non-disabled counterpart. (The non-disabled counterpart was matched by gender, grade and box marking). The PAR report of the disabled employee revealed that 14 objectives had been set during that reporting year and eight of these objectives had been achieved.
- 5.9. The non-disabled comparator had a total of six objectives, all of which had been achieved. In contrast, the non-disabled employee received detailed feedback on how to improve his performance. Reference was made to the diversity competence when assessing the performance of the disabled employee; however, no reference was made to the diversity competence in the non-disabled employee's report.

Box 3, C2 PAR reports of White disabled and non-disabled men compared

- 5.10. A review of the performance appraisal forms revealed that the disabled employee had five objectives to meet (all of which were met) compared with five objectives for the non-disabled employee (four of which were achieved). There was no significant feedback in the performance appraisal report, which advised employees how to improve their performance in future years. Reference to diversity competence appeared in both reports.

Ethnicity

- 5.11. A total of 16 PAR reports relating to BME employees formed part of the initial analysis. Four of the reports were eventually excluded from the analysis, as no comparators were available. Further details relating to the distribution of PAR box marks are given in Table 5.1. Table 5.1 summarises the profile of the sample, including details of the PAR box mark awarded. Slightly fewer PAR reports were actually subjected to qualitative analysis than appear in the overall

pool, as it was not possible to match all PAR reports by the key demographic variables.

Table 5.1. Summary of PAR files analysed

PAR marking	Male		Female		Total
	BME	White	BME	White	
1	–	1	2	–	3
2	3	7	5	8	23
3	5	10	1	7	23
Total	8	18	8	15	49

Box 2, A2 PAR reports of BME and White women compared

- 5.12. No differences were found in any of the key factors (objectives, feedback and diversity) in the analysis of these PAR reports.

Box 2, B1 PAR reports of BME and White women compared

- 5.13. Both of the PAR reports relating to BME women indicated that they were set slightly more objectives (six, all of which were achieved) than their White counterparts. Reference to equality and diversity was made in all three PAR reports. One of the PAR reports, of one of the White women, gave guidance on how performance could be improved. However, this focused on improving confidence and assertiveness. These are areas which are often adversely affected by discrimination and stereotyping.

Box 2, B1 PAR reports for BME and White men compared

- 5.14. Comparison of PAR reports of BME and White men showed that there were fewer objectives set in the BME PAR report. Both individuals achieved their objectives; however, only the PAR report belonging to the White employee

demonstrated evidence of feedback about how performance should be improved.

Box 2, B3 PAR reports of BME and White men compared

- 5.15. Comparative analysis of the PAR reports demonstrated that there were more objectives set and achieved (six) in the BME employee's PAR report than in the PAR report of his White counterpart (four). Neither employee received feedback on how their performance could be improved. Both PAR reports made reference to the diversity competence.

Box 2, level D PAR reports of BME and White women compared

- 5.16. When the PAR reports of the BME and White women were compared, the BME PAR report recorded more objectives (eight) than the report of her White counterpart (six). Both employees achieved the same proportion of objectives. There was no difference in the assessment of the diversity competence or the amount of feedback on how performance should be improved.

Box 2, C2 PAR reports of BME and White women compared

- 5.17. No difference in the number of objectives set and achieved was found when the PAR reports of a BME and a White C2 woman were compared. Neither of the reports indicated that the diversity competence was tested. However, there was clear guidance about how performance should be improved in the BME woman's report.

Box 3, A2 PAR reports of BME and White men compared

- 5.18. When the PAR files of a BME man at level A2 was compared with the PAR reports of the two White men, no difference in the number of objectives set was found.

5.19. The BME employee did not receive feedback on how to improve his performance. However, feedback on how to improve performance was evident from the PAR reports of the White employees. The diversity competence was referred to in the performance appraisal reports of the White employees.

Box 3, C2 PAR reports of BME and White men compared

5.20. The PAR report belonging to the BME employee recorded four objectives, all of which were achieved. This compared to five objectives for his White counterpart (four of which were achieved). The diversity competence was mentioned in all reports. None of the reports reviewed demonstrated that detailed guidance had been provided on how to improve performance.

Gender

5.21. A total of 23 PAR reports that related to women were available for comparison. The PAR reports most clearly matched for box marking, grade, ethnicity and disability status were compared.

Box 2, A2 PAR reports of a BME man and woman compared

5.22. More objectives (six) were observed in the PAR report of the woman when compared with the PAR report of the man (five). All objectives were achieved in both cases. There was no variation in the amount of guidance on improving performance. Nor was there any evidence of an assessment of the diversity competence.

Box 2, B1 PAR reports of a White man and woman compared

5.23. When PAR reports at box mark 2 were analysed for a White man and woman, the White man was found to have agreed more objectives (seven) than the White woman (five). The reports indicate that both employees achieved all the

objectives set and were both assessed on the diversity competence. However, only the White man received specific written feedback on how to improve his performance as part of the PAR report.

Box 2, B3 PAR reports of a White woman and man compared

- 5.24. A comparison of relevant box 3 PAR reports demonstrated that the woman was set and achieved more targets (six) than her male counterpart (four). There was no difference in the proportion of objectives achieved nor was there any difference in the amount of feedback to employees on improving performance.

Box 2, C2 PAR reports of a BME man and woman compared

- 5.25. When the box markings of BME men and women were compared, the BME woman was set and achieved more objectives (seven) than her male counterpart (four). Both employees received no feedback on how their performance should be improved in the PAR report. However the PAR report for the BME man made reference to the diversity competence.

Box 3, A2 PAR reports of White men and women compared

- 5.26. When the number of objectives set for the women and men at level A2 were compared, it was observed that women had more objectives than men. The two women participants had six objectives (one employee managed to achieve all six, whilst the other employee achieved four of her objectives). The male comparators had four or five objectives. One employee achieved two of the five objectives set, while the other achieved all four objectives.

Box 3, B1 PAR reports of White men and women compared

- 5.27. No difference in the number of objectives set or achieved was observed. Both employees achieved all their objectives. No differences in the amount of

feedback offered were found. However, there was reference to the diversity competence in the performance appraisal report of the woman.

Box 3, B2 PAR reports for disabled White men and women compared

- 5.28. No difference in the number of objectives set or achieved was noted. Nor was there any difference in the feedback, or in the assessment of the diversity competence noted.

Box 3, level E PAR report of a BME man and woman compared

- 5.29. When the box markings of a BME man and woman were compared, the woman was found to have more objectives (eight of which seven were achieved). While some feedback on how to improve performance appeared on the man's PAR report, no feedback was apparent from the woman's PAR report. Reference to diversity was made in both PAR reports.

Key points

- 5.30. *Disability:* More targets were found to be set for disabled employees, leading to an increased likelihood that they would not be able to achieve the targets set as a result. A review of the reports also suggested that disabled employees have to achieve more objectives to obtain the same PAR box mark as their non-disabled counterparts. It is impossible to determine from the reports how these higher-than-recommended number of targets were agreed, that is, whether they were driven by the wishes of the employee or the wishes of the employee's manager.
- 5.31. *Gender:* Although there are some differences in the number of objectives set, the amount of feedback provided and the competencies assessed, the variations are too small to be statistically significant. They are not so acute as to amount to an adverse impact against either demographic group.

5.32. *Ethnicity*: The results suggest that whilst there are inconsistencies in the number of objectives set, these inconsistencies are unlikely to be statistically significant or to result in an adverse impact. There were also inconsistencies in the assessment of the diversity competence.

6. STATISTICAL ANALYSIS OF PAR FILES

Introduction

- 6.1. The qualitative analysis yielded useful information about potential variances in the assessment of performance by demographic groups. However, it was necessary to determine whether or not these variations could be explained by chance. In order to examine these differences in more detail, the most salient variations identified in the reports were subjected to statistical analysis.

STATISTICAL ANALYSIS OF PAR REPORTS

- 6.2. Quantitative data derived from PAR document analyses were entered into SPSS¹ and analysed to check for subgroup differences in objectives set, objectives achieved, and shortfalls in objectives.

The distributions of the subgroups of interest are shown in Table 6.1.

Table 6.1. Number of participants per subgroup

		N
Ethnic Group	BME	15
	White	34
Gender	Female	23
	Male	26
Disability	No	38
	Yes	11

- 6.3. Chi-squared (χ^2) tests showed that the distribution of the genders within BME and White groups was exactly as would be expected in this sample and the distribution of disability status across the genders was also as would be expected. However, there were no disabled participants in the BME group, although the characteristics of the sample led to an expectation of approximately three disabled people in the BME group.

¹ Statistical Package for Social Sciences

- 6.4. Subgroup comparisons on number of objectives set showed no statistically significant differences by gender or ethnicity. However, there was a statistically significant difference in the number of objectives set according to disability status, with the mean number of objectives set for disabled people being significantly higher than the number set for non-disabled people. Details of this analysis are summarised in Table 6.2.

Table 6.2. Comparison by *t*-test of objectives set according to disability status

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
Disabled	11	8.00	3.34	2.27	10.87	.044
Non-disabled	35	5.66	1.23			

- 6.5. As shown by the standard deviations (*SD*) in Table 6.2, the spread of scores for the group of disabled staff was considerably greater than that of the non-disabled group. There was also quite a large difference in the number of people in each group (11 compared to 35). The level of significance was borderline (.044) and it must be considered that the unequal spread of scores and the unequal sample sizes may have influenced this result.
- 6.6. Further *t*-tests were performed across the three subgroups of interest to test for differences in the number of objectives met. As with the number set, there were no statistically significant differences in the number of objectives met by gender or ethnicity. Yet again, a statistically significant difference was found in the number of objectives met according to disability status. Disabled people achieved a greater number of objectives than their non-disabled counterparts, as shown in Table 6.3.

Table 6.3. Comparison by *t*-test of objectives achieved according to disability status

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
Disabled	11	7.00	2.14	3.01	44	.004
Non-disabled	35	5.34	1.39			

- 6.7. There were no statistically significant differences between objectives set and objectives achieved according to gender, ethnicity or disability status.
- 6.8. In order to examine the relationship between PAR box markings and the number of objectives set and achieved, a Pearson's correlation was performed. No statistically significant correlation between PAR mark and number of objectives set was found. Nor was there any correlation between the number of objectives achieved and PAR box mark. No correlation was observed between PAR box marks and pay band.
- 6.9. As expected, the number of objectives set and the number of objectives achieved were highly correlated ($r = .865$, $p < .000$). However, in practice, the more objectives that are set the more difficult it is to achieve these objectives. Therefore, individuals who have been given a large number of objectives are likely to be placed at a disadvantage.

Key points

- 6.10. In summary, the following key findings emerge for the statistical analysis of the paper PAR reports:
- (a) no statistically significant differences were found in respect of the number of objectives set or achieved by ethnic group or gender;
 - (b) a reliable statistically significant difference was found in respect of the number of objectives set and achieved by disability status where disabled people had more objectives;
 - (c) not surprisingly, the greater the number of objectives set, the greater was the number of unmet objectives, indicating, perhaps, that it is difficult to achieve a high number of objectives;
 - (d) no relationship was found between the number of objectives set and achieved and PAR box marking, although the small sample size must be acknowledged.

7. QUANTITATIVE ANALYSIS OF ARCHIVE REVIEW OF PERFORMANCE RATING DATA (2004)

Introduction

- 7.1. A complete census of all PAR reports held by the CPS was analysed. This analysis helped to determine how confident we could be with the findings generated from the analysis of paper PAR files. The much larger sample size available in the archive data also enabled more sensitive and more complex statistical analyses to be performed. However, as the archive data is held on an electronic database, it was not possible to examine some of the key variables (such as the number of objectives set) identified in the review of paper PAR files.

QUANTITATIVE ANALYSIS OF ARCHIVE REVIEW RATING DATA (2004)

Background to data preparation and analysis

- 7.2. Quantitative analyses of performance ratings were conducted via SPSS. The primary aim of this set of analyses was to investigate any statistically significant differences in ratings between various employee subgroups. The comparison groups were selected to include groups traditionally found to be more commonly disadvantaged in the workplace, thus:
- BME employees were compared with White employees;
 - disabled employees were compared with employees without stated disabilities;
 - men were compared with women.
- 7.3. The data-sets originally provided included several missing data cells and had to be confined to cases with data in all cells used for analysis. This resulted in a usable cleaned data file.

Performance Appraisal Ratings: a Review and Impact Assessment

- 7.4. The original data-set for 2003 contained a total of 7,563 cases. However, we were concerned about the large amount of missing data – a total of 1,721 cases lacked ethnicity information (235 objectors; 1,486 ‘no reply’), and 1,723 cases lacked information regarding disability status. Gender details were omitted from seven cases.
- 7.5. The original data-set for 2004 contained a total of 8,099 cases. However, a total of 1,625 cases lacked ethnicity information (224 objectors; 1,401 ‘no reply’), and 29 cases lacked information regarding disability status. Gender details were omitted from six cases.
- 7.6. The missing data for different variables did not always belong to the same cases, hence the relatively high number of discarded cases.
- 7.7. In addition to problems of missing data, the subgroups for comparison were of quite disparate sizes. For example, for the year 2004 the proportions for the key groups of interest were as presented in Tables 7.1–7.3.

Table 7.1. Proportions of BME and White employees

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Not Ethnic Minority (White)	4392	87.3	87.3	87.3
Ethnic Minority	638	12.7	12.7	100.0
Total	5030	100.0	100.0	

Table 7.2. Proportion of disabled and non-disabled people

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Non disabled	4760	94.6	94.6	94.6
	Stated disability	270	5.4	5.4	100.0
	Total	5030	100.0	100.0	

Table 7.3. Proportion of men and women

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	1650	32.8	32.8	32.8
	Female	3380	67.2	67.2	100.0
	Total	5030	100.0	100.0	

Creation of subsample data-sets

- 7.8. In order to minimize potential problems of different group size, various data-sets were drawn from the main data-set to provide random samples of majority groups that would be approximately the same size as the minority group with which they would be compared. To minimize potential problems of drawing a typical random sample for comparison, three random samples were drawn from each majority group of interest.
- 7.9. The first step in the creation of subsample data-sets was the isolation of the minority group of interest as a separate minority group data-set. For example, all cases identified as having disability status were taken from the main data-set and saved as a 'disability status' data file.
- 7.10. The second stage entailed the separation of a random sample of cases (approximately equal in size to the associated minority group) from the

remaining majority group. In this stage, the selected random sample was saved as random selection 1 (non-disability status). The process of drawing a random sample of cases from the majority group of interest was repeated twice to provide random selection 2 and random selection 3 data files.

7.11. The third stage entailed the separate merging of each random selection file with the minority data-set file (e.g. disability status with non-disability status random selection 1; disability status with non-disability status random selection 2; disability status with non-disability status random selection 3). This provided three new data files that could be used to investigate group differences within groups of approximately equal sizes.

7.12. The same principle was applied to the derivation of three sets of comparable data files for investigation of ethnicity and gender differences. In each case, comparisons were made between the minority group of interest and all three of the randomly drawn comparison samples. In addition to the three similar sized groups compared for each group of interest, comparisons were made within the entire sample set.

Statistical analyses

Initial analyses

7.13. The first set of analyses included cross-tabulations to assess the extent to which the distribution of personnel matched what would be expected by chance, if they were in accordance with the characteristics of the population represented by the data-set. These analyses were conducted within the entire cleaned data-set relating to 2004.

7.14. The comparisons included distributions of:

Disability by	Gender	(<i>ns</i>)
	Ethnicity	(<.01)
	Lawyer status	(<i>ns</i>)
Ethnicity by	Gender	(<.01)
	Lawyer status	(<.05)
Lawyer status by	Gender	(<.001)

Full details of this cross-tabulation analysis are given in Appendix 1.

7.15. The distribution of disabled people was not significantly different between the genders, or between lawyers and non-lawyers. However, there were significantly fewer disabled people within BME groups than would have been expected by chance alone according to the characteristics of the data-set population.

7.16. There were significantly more women and fewer men within the BME groups than would be expected by chance distributions and significantly fewer lawyers within the BME group.

7.17. The strongest statistically significant result revealed a disproportionately low level of female lawyers within the data-set analysed.

Overall comparisons

7.18. On completion of data cleaning and exploratory analyses, statistical tests were focused on investigating associations between PAR ratings and demographic characteristics of review ratings, and demographic characteristics by ethnicity, disability and gender.

7.19. This stage of the analysis began with an initial exploration of the entire cleaned data-set (2004), using factorial ANOVA (analysis of variance) to test for main

effects of ethnicity, disability status and gender; and to test for interaction effects across these independent variables.

7.20. The ANOVA (see Tables 7.4 and 7.5) showed only one statistically significant relationship – that of a main effect relationship between PAR rating and disability status. In other words, there was something that was causing the disabled employees to be rated less favourably. Examination of group means showed that non-disabled people tended to achieve more favourable PAR ratings than their disabled colleagues (mean rating scores of 2.5 and 2.7, respectively), indicating that disabled people were, on aggregate, rated less favourably than their non-disabled colleagues. This could amount to direct or indirect discrimination. More details of this ANOVA are given in Appendix 2.

7.21. The analysis showed no statistically significant interaction effects, and disabled people achieved slightly higher (less favourable) review ratings than people without stated disabilities in both genders and in all ethnic groups.

Table 7.4. Numbers per group in factorial ANOVA to investigate main effect and interaction relationships between review ratings and minority group

Between-subjects factors			
		Value label	N
Disability status	0	Non-disabled	4,760
	1	Disability	270
Sex	0	Male	1,650
	1	Female	3,380
01 numerical version of collapsed ethnic group	.00	White	4,392
	1.00	Ethnic Minority	638

Table 7.5. Results of factorial ANOVA to investigate main effect and interaction relationships between review ratings and ethnicity

Source	Type III sum of squares	df	Mean squared	F	Significance
(V21) Disability status	2.296	1	2.296	8.294	.004
(V23) Gender	0.144	1	0.144	0.521	.470
(V26) Ethnicity	0.859	1	0.859	3.103	.078
Disability × Gender	0.341	1	0.341	1.233	.267
Disability × Ethnicity	0.459	1	0.459	1.659	.198
Gender × Ethnicity	0.637	1	0.637	2.302	.129
Disability × Gender × Ethnicity	0.517	1	0.517	1.869	.172
Error	1,389.968	1	0.277		
Total	34,336.000				

Comparison of review ratings between BME and White employees

7.22. Overall, four *t*-test comparisons by ethnic group were performed. The first comparison (Table 7.6) examined the entire data-set; whereas the following three sets of comparisons (Tables 7.7–7.9) were undertaken on the three separate derived data-sets as described above.

Table 7.6. Comparison by *t*-test of review ratings by ethnicity – entire data-set

Group	N	Mean	SD	t	df	Significance (2-tailed)
BME	638	2.58	0.54	1.313	5,028	.19
White	4,392	2.56	0.52			

Table 7.7. Comparison of *t*-test for review ratings by ethnicity – subsample 1

Group	N	Mean	SD	t	df	Significance (2-tailed)
BME	638	2.58	0.54	0.71	1,271	.48
White	635	2.56	0.51			

Table 7.8. Comparison of *t*-test for review ratings by ethnicity – subsample 2

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
BME	638	2.58	0.54	0.63	1,321	.530
White	685	2.57	0.52			

Table 7.9. Comparison of *t*-test for review ratings by ethnicity – subsample 3

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
BME	638	2.58	0.54	0.72	1,319	.473
White	683	2.56	0.52			

Clearly, there is no evidence of differences in PAR ratings within the 2004 data-set.

Comparison of review ratings between disabled and non-disabled employees

7.23. This analysis also involved the use of four separate data-sets and, as expected from the initial ANOVA results, all four yielded statistically significant results. In all cases the mean PAR rating of the ‘disabled’ group was higher than for the non-disabled group. The *t*-tests are summarised in the tables, with Table 7.10 showing results for the entire data-set and Tables 7.11–7.13 showing the results for the three subsample tests.

Table 7.10. Comparison by *t*-test of review ratings by disability status – entire data-set

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i> ²	Significance (2-tailed)
Disabled	270	2.67	0.49	3.71	305.93	.001
Non-disabled	4,760	2.55	0.53			

Table 7.11. Comparison by *t*-test of review ratings by disability status – subsample 1

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
Disabled	270	2.67	0.49	2.892	567.7	.042
Non-disabled	301	2.54	0.52			

Table 7.12. Comparison by *t*-test of review ratings by disability status – subsample 2

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
Disabled	270	2.67	0.49	3.33	550.4	.001
Non-disabled	284	2.52	0.54			

² Where degrees of freedom are given with decimal points, homogeneity of variance does not apply and *t* is calculated with the appropriate formula for this situation.

Table 7.13. Comparison by *t*-test of review ratings by disability status – subsample 3

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
Disabled	270	2.67	0.49	2.34	534.02	.020
Non-disabled	272	2.56	0.55			

7.24. These results indicate that there is a statistically significant difference between the PAR ratings of disabled and non-disabled people. The consistently higher means shown for disabled employees indicates that this group, on aggregate, tends to be regarded less favourably (in terms of PAR ratings) than their non-disabled colleagues. Obviously, the reason for this difference is not explicable from these data.

Comparison of review ratings between men and women

7.25. Yet again, separate analyses were performed on the entire data-set and each of the three derived data-sets. The results of these tests were not as straightforward as the previous comparisons, as two of the three tests yielded statistically significant results and one yielded a non-significant result. The two results that showed statistical significance both had more equal group sizes (Male = 1,659, Female = 1,668; $t = 2.05$, $p = .04$: Male = 1,659, Female = 1,659; $t = -2.46$, $p = .01$) than the data-set that produced the non-significant result (M = 1,659, Female = 1,751; $t = 1.56$, $p = .12$). Therefore, another data-set of 1,659 females was randomly drawn and tested against the male sample. This test was non-significant ($t = 1.88$ $p = .06$). Whatever the statistical significance of the test outcome, the difference between group means was consistently very small, with females having the slightly higher mean (Male = 2.54, Female = 2.57–2.58), indicating a slightly less favourable rating.

7.26. Summaries of the results of the *t*-tests between the sexes are shown in Tables 7.14–7.16.

Performance Appraisal Ratings: a Review and Impact Assessment

Table 7.14. Comparison by *t*-test of review ratings by gender – entire data-set

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
Male	1,650	2.54	0.53	-2.08	3,220.81	.038
Female	3,380	2.57	0.52			

Table 7.15. Comparison by *t*-test of review ratings by gender – subsample 1

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
Male	1,650	2.54	0.53	-1.81	3,330	.070
Female	1,682	2.57	0.53			

Table 7.16. Comparison by *t*-test of review ratings by gender – subsample 2

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
Male	1,650	2.54	0.53	-1.96	3,300	.050
Female	1,652	2.57	0.53			

Table 7.17. Comparison by *t*-test of review ratings by gender – subsample 3

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
Male	1,650	2.54	0.53	-2.56	3,292.95	.01
Female	1,650	2.58	0.51			

7.27. Perhaps the most important attribute of tests between the genders was that of sample size. The minority group in this case (males) was substantially larger ($N = 1,659$) than the other minority groups: BME group ($N = 638$); disabled staff group ($N = 270$). Therefore, the derived sample size for gender comparisons was approximately 3,000, whilst the derived sample sizes for ethnicity were approximately 1,300, and for disability approximately 600. As test power increases with sample size, the likelihood of identifying even very small differences between groups increases in line with sample size. Thus, the relatively large sample size of the gender differences test could be almost as influential as real group differences in the production of statistically significant results in this case. Whatever the reason, examination of the actual means shows that there is very little material difference between men and women in this case.

Comparison of review ratings between lawyers in BME and White groups

7.28. One objective of the research was to test for ethnicity effects within the population of lawyers in this data-set. To test for differences in review ratings according to ethnicity, a *t*-test was undertaken within the entire (2004) data-set of lawyers. The results showed that there was no significant difference in review ratings between BME (*N* = 172) and White (*N* = 1,358) groups (*t* = 0.142, *p* > .89). The mean scores of the groups were remarkably close (5.52 and 5.53, respectively) as were the standard deviations (0.61 and 0.53, respectively). Three further tests using derived data-sets with equal-sized groups (approximately *N* = 170), also yielded three non-significant results. These results are summarised in Tables 7.19–7.21.

Table 7.18. Comparison by *t*-test of review ratings of lawyers by ethnicity – entire data-set

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
BME	172	2.52	0.61	1.28	205.73	.887
White	1,358	2.53	0.53			

Table 7.19. Comparison by *t*-test of review ratings of lawyers by ethnicity – subsample 1

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
BME	172	2.52	0.61	-0.10	340	.923
White	170	2.53	0.57			

Table 7.20. Comparison by *t*-test of review ratings of lawyers by ethnicity – subsample 2

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
BME	172	2.52	0.61	0.48	342	.635
White	172	2.49	0.52			

Table 7.21. Comparison by *t*-test of review ratings of lawyers by ethnicity – subsample 3

Group	<i>N</i>	Mean	<i>SD</i>	<i>t</i>	<i>df</i>	Significance (2-tailed)
BME	172	2.52	0.61	0.33	345	.744
White	175	2.5	0.56			

7.29. These results indicate no differences at all in the PAR ratings of lawyers from Black and Minority Ethnic groups and those from White groups.

Key points

- 7.30. The results of the analysis of all PAR reports for the reporting period ending 2004 tended to confirm the initial analysis in that:
- (a) no statistically significant differences in respect of the PAR ratings of BME and White employees were found;
 - (b) no statistically significant differences in respect of the PAR ratings of BME and White lawyers was found;
 - (c) no statistically significant differences in respect of the PAR ratings of men and women were found;
 - (d) disabled employees were more likely to receive less favourable PAR ratings.

8. ANALYSIS OF 'YOUR VOICE 2004' STAFF SURVEY RESULTS

Introduction

- 8.1. We have included a further statistical analysis of the staff survey questionnaire because these data provide some insight into the variations in job satisfaction across the various demographic groups. Staff satisfaction may impact on PAR ratings, as dissatisfied staff are less likely to expend the same amount of effort on their objectives as staff who are satisfied with their job.
- 8.2. An analysis of the 'Your Voice' staff survey questionnaire results was carried out. The analysis examined factors which are known to be related to less favourable performance appraisal ratings, such as satisfaction with the current performance appraisal scheme and job satisfaction. Whilst the previous analyses were conducted on raw data, the data available for the staff survey was only at the collapsed percentage level, and therefore was only really appropriate for chi-squared analyses of favourable, neutral and unfavourable response groups.

ANALYSIS OF 'YOUR VOICE 2004' STAFF SURVEY RESULTS

Analysis of replies to the statement "I believe there is an effective system for recognising people who perform effectively" by demographic group

- 8.3. The analysis revealed a significant difference between White employees and other ethnic groups in relation to the belief that there is an effective system of recognising staff performance. Some BME employees were more positive in their belief that there is an effective system for managing performance than others.
- 8.4. When the small sample of Chinese respondents were removed from the analysis of replies to this question, no statistically significant differences were

found in the response to the statement “I believe there is an effective system for recognising people who perform effectively”.

- 8.5. The analysis of results also revealed that employees under the age of 20 and those in the 60+ category were more positive about there being an effective system for recognising people who perform effectively than employees in other age categories ($\chi^2 = 27.72$, $df = 10$; $p < .01$).
- 8.6. In order to examine whether sexuality was related to the belief that there was an effective system for recognising people who perform effectively, a chi-squared analysis was carried out on the responses. No significant differences were found.
- 8.7. An analysis of questionnaire responses was also carried out by gender. No significant difference was found between men and women on the question that there is an effective system for recognising performance.
- 8.8. The analysis of the responses of disabled and non-disabled employees revealed that disabled employees were significantly more negative than their non-disabled counterparts about the question of whether or not there is an effective system for recognising people who perform effectively ($\chi^2 = 8.48$, $df = 2$; $p < .01$).

Analysis of satisfaction with the amount of responsibility in current job by demographic group

- 8.9. The χ^2 test showed that there were no statistically significant differences by ethnic group, sexuality or gender in relation to satisfaction with the amount of responsibility given in the employee's current role. However, there was a highly significant difference in relation to age; employees in the 20–29 age category were the most negative about the amount of responsibility given to them in their current role.

8.10. A statistically significant difference was also found in relation to disability. Disabled employees were more negative about the amount of responsibility given to them in their current roles ($\chi^2 = 11.90$, $df = 2$; $p < .001$). It is difficult to ascertain whether this response means that they felt there was too little or too much responsibility given to them in their current roles. Certainly, the analysis of PAR reports shows that disabled employees have proportionally more objectives set for them than for their non-disabled counterparts.

Analysis of responses to satisfaction with variety in job by demographic group

8.11. A statistical analysis of demographic variable by variety in job showed that there were no statistically significant differences by ethnic group, sexuality or gender in relation to satisfaction with variety with current job.

8.12. Statistically significant differences were observed by age; employees in the 20–29 age category were least satisfied with the amount of variety in their jobs ($\chi^2 = 96.39$, $df = 10$; $p < .001$). There was also a statistical difference found when the responses of the employees with and without disabilities were analysed. Disabled employees were more likely to respond negatively to the question that they felt satisfied with the amount of variety in their job ($\chi^2 = 11.69$, $df = 2$; $p < .01$).

Analysis of replies to the statement “my job gives me a personal sense of accomplishment” by demographic group

8.13. Statistical analysis of responses by demographic group to the statement “my job gives me a sense of personal accomplishment” was carried out using the chi-squared procedure. No statistically significant differences were observed when responses by sexuality were analysed. There were also no statistically significant differences by gender.

8.14. An analysis of the results by age found that employees in the 20–29 age category were most negative about the statement that their job gives them a

personal sense of accomplishment. This finding was highly significant ($\chi^2 = 62.61$, $df = 10$; $p < .001$).

- 8.15. When the responses to the statement “my job gives me a personal sense of accomplishment” was analysed by disability, the results revealed that employees with a disability were more negative about this statement ($\chi^2 = 14.88$, $df = 2$; $p < .001$). This result is highly significant.

Analysis of responses to the statement “I feel motivated to do the best I can in my job” by demographic group

- 8.16. No statistically significant differences were found by ethnic group, age or gender in relation to this question.
- 8.17. A statistically significant difference was found in the responses to this question by sexuality. Employees who preferred not to disclose their sexuality were least likely to say that they felt motivated to do the best they can in their job ($\chi^2 = 21.93$, $df = 8$; $p < .01$).
- 8.18. There was also a highly significant difference between disabled and non-disabled employees in response to the statement “I feel motivated to do the best I can in my job”. Disabled employees were far more likely to respond negatively to this statement ($\chi^2 = 41.02$, $df = 2$; $p < .001$).

Analysis of responses to the statement “I feel satisfied with the level of job security” by demographic group

- 8.19. No statistically significant differences were found by ethnicity, sexuality, age or gender in relation to the statement “I feel satisfied with the level of job security”.
- 8.20. A statistically significant difference was observed when the results of disabled and non-disabled employees were analysed in relation to this question. Employees with disabilities were more negative about their level of job security ($\chi^2 = 20.85$, $df = 2$; $p < .01$).

8.21. Responses were re-analysed with the Chinese respondents removed from the main sample, because of the small number of Chinese respondents in the sample. When Chinese respondents were removed from the sample, the result changed from non-significant to significant ($\chi^2 = 13.41$, $df = 8$; $p < .05$). Asian respondents were more negative about their job security.

Analysis of responses to the statement “I intend to be working for the CPS in 12 months’ time” by demographic group

8.22. No statistically significant difference was found in relation to responses to the statement “I intend to be working for the CPS in 12 months’ time” by ethnic group. Nor were any statistically significant differences found by gender. However, a statistically significant difference in respect of sexuality was found. Employees who preferred not to disclose their sexuality were more likely to respond neutrally to the question that they would be working in the CPS in 12 months’ time ($\chi^2 = 36.06$, $df = 8$; $p < .001$).

8.23. Disabled employees were also more neutral about whether they would be working for the CPS in 12 months’ time than other groups ($\chi^2 = 7.44$, $df = 2$; $p < .05$).

Key points

8.24. The following significant differences emerged from the analysis of the ‘Your Voice’ staff survey 2004 questionnaire data:

- (a) employees in the 20–29 and 60+ age categories were the most positive about there being an effective system for recognising people who perform effectively;
- (b) disabled employees were the most negative about there being an effective system for recognising people who perform effectively;
- (c) disabled employees and employee in the 20–29 age category were most negative about the amount of variety in their jobs;

- (d) disabled employees and employees in the 20–29 age category were most negative about their jobs providing them with a personal sense of accomplishment;
- (e) disabled employees and employees who declined to disclose their sexual orientation were least likely to state that they feel motivated in their jobs;
- (f) disabled employees and Asian employees were most negative about their level of job security.
- (g) disabled employees and employees who declined to disclose their sexuality were most neutral in their responses to the statement that they would be working for the CPS in 12 months' time.

9. QUALITATIVE IMPACT ASSESSMENT CONCLUSIONS

- 9.1. At the beginning of this study we indicated that our research objectives were to determine whether or not PAR box markings are adversely affected by the membership of a particular demographic group. The literature review strongly suggested that performance appraisal ratings would be adversely affected by the membership of a particular demographic group irrespective of real job performance. This compelling conclusion, which followed from the previous research carried out in the UK and more particularly in the Civil Service, was that statistically significant differences would be found between BME and White PAR ratings reports. Whilst the findings from the research in the Crown Prosecution Service do not necessarily undermine the considerable amount of studies which did find evidence of discrimination or adverse impact, no statistical evidence for discrimination on the basis of ethnicity or gender was found in the CPS.
- 9.2. The only real subgroup differences to emerge from this study were that, when the performance ratings were analysed, disabled employees tended to achieve lower performance appraisal ratings than non-disabled employees . These findings did reach statistical significance.
- 9.3. Many of the findings in this study are unusual, insofar as the literature previously surveyed (Greenhaus & Parasuraman, 1993; IES, 2001; Tackey *et al.*, 2001) indicates that there will be a statistically significant difference in the PAR box markings of BME employees and their White counterparts. Further, it was expected that this difference would show that BME employees would obtain significantly less favourable PAR box markings than their White counterparts. The review was unable to confirm this finding in the CPS. If the results of this study are compared with the research across other major Civil Service departments (IES, 2001) it would appear that the BME employees in the CPS are attaining better PAR scores than BME people employed in the other government departments.

- 9.4. There is, of course, the possibility that the findings of the current review represent a statistical anomaly. However, the study also found no evidence of discrimination within the review of paper PARs. Furthermore, the large number of statistical tests carried out with different subsamples and at different times means that the possibility that the current findings are incorrect is extremely remote. It is acknowledged that in order to be absolutely certain that the results were not due to chance – and also to monitor ongoing performance – the analysis carried out within this review will need to be repeated in future years.
- 9.5. The study also demonstrates that when the PAR box markings of disabled employees are looked at as whole, disabled employees tend to obtain comparatively less favourable PAR box markings.
- 9.6. Clearly a further analysis of the PAR statistics of disabled employees in future years will serve as a useful tool in monitoring the effectiveness of equality and diversity strategies within the CPS.
- 9.7. The findings of the current study are remarkable insofar as they clearly demonstrate that BME employees are able to obtain good PAR box mark ratings when managed in a workplace that has relatively high regard for equality and diversity. The findings do not negate the perceptions of members of various key demographic groups that they feel disadvantaged by the PAR process. Phase two of this research explored why these perceptions persist.

10. ANALYSIS OF FOCUS GROUP AND DEPTH INTERVIEWS

Introduction

10.1. Phase two of the impact assessment was a qualitative assessment of the evidence supported by a snapshot survey. Themes for the focus groups, depth interviews and mini-survey were derived from our analysis of the issues arising from phase one. All participants in phase two were self-selecting.

ANALYSIS OF FOCUS GROUP AND DEPTH INTERVIEWS

10.2. Focus groups were convened and in addition to the focus groups, 20 depth interviews were carried out. These depth interviews allowed for greater analysis of the issues affecting staff.

10.3. In total five employees participated in the focus group. Women, part time employees, BME disabled and older workers were well represented amongst respondents. There was also good representation from individuals at every level of the organisational hierarchy. The sample also included individuals from the lesbian, gay and bisexual communities.

10.4. On average each interview lasted for approximately one hour. Many interviews exceeded this timeframe by far. All interviews and focus group discussions were recorded in order to enable accurate and detailed analysis of the discussions.

10.5. Whilst it was not possible to include all the transcripts of interviews and focus groups some of the key messages to emerge from the qualitative study are reproduced below. Key demographic information has been removed from the quotations to maintain confidentiality.

(a) Core questions

Researcher: *“Do you think that your PAR rating accurately reflected your performance? If not, in what ways did it under-and/or over-rate your performance?”*

Summary

- 10.6. Most employees considered that their PAR rating accurately reflected their performance however there were two cases the first case an African Caribbean man and the second case an African Caribbean woman, who considered their PAR rating did not accurately reflect their performance. Both employees felt that the box marking did not accurately reflect their effort.

Example of responses

- 10.7. BI Caseworker: *“Yes but previously, I’d say most of the years I’ve been quite happy with my report, I mean I’ve felt that I’ve got the box markings that I’ve truly deserved.”*
- 10.8. HR manager: *“No. I’ve never had problems”.*
- 10.9. Crown Prosecutor: *“No but probably I’ve been more accepted because of being a gay man it can now work in your favour, not intentionally but because of the way things are. As diversity is one of the strongest things now in the CPS and probably the rest of the Civil Service as well”.*

Researcher: *“Which competencies do you find most relevant and most irrelevant to your job?”*

Summary

- 10.10. The competencies which employees considered to be most relevant to their job differed according to the nature of the post which the participant occupied. More senior employees cited leadership as the most relevant job competence whereas employees further down the hierarchy referred to achieving results, planning and organising. There was some concern about the fact that the latest generation of job descriptions referred to competencies which are inconsistent with the current core competence framework. This is likely to lead to some confusion as to which competence set will apply when the performance of the job holder is reviewed.

Example of response

- 10.11. Trials Unit Business Manager: *“I’ve been concerned about and it’s pretty much where the core competence sits with the new recruitment process. For example the job profiles now are drafted very much in isolation and if you look at a few, whilst we’ve got the generic ones any other strange ones they’re all done independently, and yet they have to be analysed before they’re advertised but they no longer sit neatly within the core competence framework because the bullet points that are now being used are different – in fact people are creating their own.*
- 10.12. *Consequently if you for example advertise a post and if the description under the competencies no longer match the core competence framework how can you then set the standards for the job in accordance with the generic standards that have been advertised when the two no longer reflect each other? Consequently if the recruitment process has changed such that now the job will or the job profile gives you a better idea and you can recruit the right person for the post, how can you then sit that together with a generic standard when you’re doing your PAR when the two no longer mirror each other?*
- 10.13. *For the most part I think the core competence framework is now a bit of a dead duck because if I advertise a job on the 1st of April for example and fill it in by the 1st of May, not that it could happen that fast, but if you did and then you’re looking at writing a report in 12 months time – what standard are you measuring them against? Is it the job profile you go back to or do you go to the core competence framework because they don’t fit together anymore”.*
- 10.14. Case Progression Officer: *“I think the most relevant would be to achieve results in planning and organising in my present role. No I just think it’s the competencies have been brought in over the last few years obviously and I just think it’s more a political thing - everyone friendly, race friendly, sex friendly”.*
- 10.15. Area Business Manager: *“I suppose in my job I suppose leadership. Achieving results, communication and negotiation, planning and organising”.*

Researcher: *Did you find any of the competencies easier or more difficult to meet than others?*

Summary

10.16. Many of the participants indicated that the most difficult competence for them to meet was that of continuous self development, this is either because there was not the time or resources to ensure that continuous self development took place. Some employees also considered the diversity and leadership competencies difficult to meet.

Examples of responses

10.17. Part-time BI Caseworker: *“Continuous self-development again is an issue with the continuous self-development for me personally – I’m part-time and I’m restricted to certain hours because of children in school. It’s not something that is very well embraced here in this office. I might point out at the beginning that I don’t think that they’re achievable – but it’s a national standard so it has to be set and then I spend probably my time in my quarterly review explaining to her why I didn’t meet the national standard because I only work part-time and I’m only in the office three days a week. When somebody has set the national target, of, for example, when somebody is given the target to send a brief out to Counsel within fourteen days of the case being committed, well that might be ten working days for a full time member of staff but for me there is no sort of compensation for me being in the office only three days a week as opposed to five and I am now penalised and having to do my briefs within six working days as opposed to ten”.*

10.18. Case Progression Officer: *“But I do find leadership quite hard because obviously everyone is different and I’ve got my team and that’s various people, their standards of work are different, their personality is different, they even make people think ‘yeah our leadership you can just delegate’. I’m not a person to just delegate, I look at the person, learn and see where they come from, what they can do, what they can achieve, what I can push them to do. So it is quite a hard one really to achieve”.*

10.19. Area Business Manager: *“I think the one competence which is the most difficult to meet – again, I dealt with it this morning – giving some people some lawyers’ feedback*

Performance Appraisal Ratings: a Review and Impact Assessment

on a recent selection exercise is diversity awareness. Its not as easy to cling onto – not as tangible and I can think of examples when people say can you give me an example to put in form I can think where someone has done a really good job as a witness whose disabled or dealt with a member of a victim who doesn't speak English as a first language there are all sorts of things. Then I can think of things I've done but it is frequently the most difficult to hang your hat on. Because we got off to a bad start as you will remember when we got criticised about five years ago for our attitude to diversity and it was then brought into the Department in a slightly heavy handed way, you know, you will all go on this course and you will be lectured upon this. It was rather than selling the idea and it was sold in the wrong way and I think it's taken some time to recover from that. We're much more professional about it now. But it is a more difficult one to demonstrate that you have met"

Summary

10.20. A large number of employees indicated that they did not get adequate feedback about their job. In addition to this it became evident that employees were more likely to receive negative feedback or criticism as opposed to positive feedback.

Researcher: *“Please list the standards of behaviour you would expect from an appraiser and an appraisee?”*

Summary

10.21. Honesty and the ability to provide constructive feedback through dialogue were considered to be the key attributes required by those leading the performance appraisal process.

Example of response

10.22. **Chief Crown Prosecutor:** *“Well at every level, a combination of pleasant approach, supportive approach from the manager, honesty mixed with a degree of common sense, you avoid being brutally frank, that would be harmful, the chance to improve the performance of the Officer, a chance to praise for good performance during the year and an attempt to identify whether performance is improving are they becoming more confident with the job, better at their job and so on. And from the Officer, I mean the*

interviews that work well is a contribution coming back from the Officer recognising what the Reporting Officer is saying – be it good or bad, being encouraged obviously in making that point constructively with the praise, so that both parties finish up having been fairly honest, frank discussion about the past year, development needs and then plan what’s going to happen the following year so that I see people get something out of it as opposed to complete very formal process, very cautious on what’s said based on generalised recollections rather than specifics and each party is doing this sort of dance and avoiding any problems and no one really gets anything out of it very satisfactory, those would be the contrasts”.

Researcher: *“Can the objectivity of the PAR system be improved? If so, how?”*

Summary

10.23. Participants made a number of suggestions for improving the objectivity of the performance appraisal process, one of the most pertinent was to ensure that both parties collect a suitable amount of data throughout the reporting year so that the PAR interview can be carried out effectively in the light of the available evidence. Assisting managers in framing work objectives in a SMART way would also assist tremendously in improving the objectivity of the PAR process. In addition to this, employees also indicated that if a box marking system is used there must be a process to more accurately differentiate people, so that all employees are not given a box mark 3.

Example of responses

10.24. Chief Crown Prosecutor: *“Well if... you first of all do need the Reporting Officer to have a good knowledge of his staff, they need to have regular contact with the staff, sometimes the span of control is too large and people don’t have regular contact I think it must be very difficult to objectively assess these staff. If I can just illustrate obviously the core of our work is prosecuting, if the Reporting Officer doesn’t get to court to see his staff in court how does he know how good they are and accurate, how well prepared are they for court? How good are they? The reality would be for many people they never get to see who they’re reporting on and yet it’s at least half of their work so it’s ... people might say oh I can tell from talking to somebody, you can tell a good advocate, you can tell and things like that. But it’s not really true so I think for a lot of*

Performance Appraisal Ratings: a Review and Impact Assessment

the process there just isn't the adequate information to be objective and so I think people finish off with the sort of subjective assessment of those people who seem to be articulate, forward, cheerful, outgoing you know and so on. So I think for many Lawyers they resent this process because it's based on very thin, objective data and as I've mentioned before I think an awful lot of Line Managers don't jot down during the year the good things that happen so we get people writing in saying they were very pleased with the Lawyer who prosecuted their case and he's very pleasant and he introduced himself, all that sort of stuff should go straight onto somebody's file.

10.25. *It also ought to be brought up with other comments so that when you come to do that report it's relatively easy because you've got quite a few pieces of data which are specific and actually justify the assessment you've probably got by then.*

10.26. *I think the staff don't think the process is objective I think they think that an element of favouritism; certain managers like certain people more therefore they get a better box marking”.*

10.27. BI Admin Manager: *“Well it sort of relates to number one then in a way whereby you've got a good worker that just gets a basic three and he has to exceed the hard worker. Yeah I just think we need to identify the box marking three so that we got a, b and c within it. So A, a person that's nearly getting towards the box marking 2 and the way that they've done the box marking 2 whereby they expect you to be acting up, I don't think that should be the case because not everybody likes to act up and there's not always the opportunity either”.*

10.28. Trials Unit Business Manager: *“In an ideal world you would have a good Line Manager and an excellent Counter Signing Officer and there wouldn't be an issue with improving quality of stuff that was recorded. Some of the most difficult issues that I've found, talking to people that I manage and other managers, is framing objectives in a way that makes them SMART and can be very difficult especially when the job is ever changing, new initiatives every day, there are new things that often we're trying to capture in a forward job plan, initiatives that we're not sure in our own mind what they're going to be yet and trying to capture them in terms that will make them achievable and understandable can often be very difficult but that's just the day to day*

Performance Appraisal Ratings: a Review and Impact Assessment

stuff, some of the stuff that confuses people, or certainly confuse people when it first came in, it's less of an issue now I think, is in and around the diversity, an objective on diversity because initially there was a wide misunderstanding. I think, that diversity meant race awareness; gradually I think we're getting there".

Researcher: *"Are you satisfied with the current mechanisms for appealing against PAR reports? (a) If yes, please indicate why; (b) if no, please state why?"*

Summary

- 10.29. The current mechanisms for appealing against PAR reports could be improved. This point is quite effectively made by the Chief Crown Prosecutor who participated in this study.

Example of response

- 10.30. **Chief Crown Prosecutor:** *"Well ... to the extent that the person appealed to is really only looking at the mechanics of the process, it doesn't seem a very good system and almost instinctively Managers support other Managers, it's usually based, unfortunately to, it does break down into fairly subjective assessments but provided the mechanics have been done the recollection of the last one I was involved in was, well, that's it. You know it's not very satisfactory, on the other hand without the data from the Officer or the Reporting Officer how does one go behind the marking?"*

Researcher: *"Is there any additional advice that you would like from the HR department to help you participate better in PAR reviews?"*

Summary

- 10.31. Participants were quite positive about the input HR could make to improving participation in PAR reviews. One respondent suggested that the HR Directorate provide new 'How do I' guides on PAR and managing poor performance. There were also suggestions that emails should be sent just before the reporting cycle with a link on it to the relevant section of the PMM.
- 10.32. Other participants felt that it would be useful for the HR Department to re-emphasise the importance for all individuals to keep notes or records

demonstrating how they have performed against each job competence. Participants felt that it would be useful if HR remind individuals to set objectives which are significant and add value to the business. Examples of good PAR reports for different grades of staff were felt to be useful and it was noted that the new system would have to be effectively marketed to make it a success.

Example of responses

- 10.33. Trials Unit Business Manager: *“On the intranet I find it difficult to work around. I actually keep a paper copy because I find it simpler, you know the small booklets that we’ve had issued to us on work life balance, flexible working, I think that we should have new ones on PAR and managing poor performance.*
- 10.34. *Right, it’s the kind of thing that if it only happens once a year people will forget where the guidance is and they’ll forget what they need to know and forget how to do it and so if somebody says on the intranet in March when we all started to think about PARs, to remind people of the process, you know this is.. or a hot link through to the section of the PMM on the intranet so that people can get to it quickly and so they will know – ‘oh yes I’ve just seen that, I’m aware of that’, and they can click through to it to check what the process should be, there’s no reason why we shouldn’t use technology available and put it in its place, that might be quite useful”.*
- 10.35. Crown Prosecutor: *“Yeah. It would I suppose be helpful if HR would re-emphasise to all job holders that it’s in their interest to be familiar with the guidance and you should also be keeping information about what you have been doing during the course of the reporting period and that’s all I can come up with otherwise I think all the information is there”.*
- 10.36. *I think what might be helpful is for different grades of staff, different types of staff, to have three or four examples of what you mean, so people could see what a good report, a sensible report would actually look like. You never see each others reports; you never see other Area reports. I’ve never been shown an example in my life of what a good PAR looks like”.*

Researcher: “*Is there anything that discourages you from writing down information to demonstrate how you have performed against each job competence?*”

Summary

10.37. Although most participants did not consider that there was any reason why they would be unable to write down information to demonstrate how they had performed against each job competence, it is important to note that the disabled participant and the African Caribbean man indicated that they would experience difficulties in collating the relevant information. This could be overcome by appropriate training and support.

Example of responses

10.38. Designated Case worker: “*A lot of the time people could write exactly what they wanted. And sometimes it would be very difficult to disprove or prove. I don’t like blowing my own trumpet to be honest with you. I find it very hard to do. I don’t see that I’d do any better or worse than the average person in my group. I’m certainly not as good as some and I’m not as bad as others. I’m just on a line and I find it very difficult to I can identify my defects better than I can identify my qualities*”.

10.39. AI Support worker: “*I’m not a writer and I need to brush up on my writing skills because unfortunately when people work in the office they write most things down. Unfortunately I don’t really write things down, when you put things down in writing it’s a powerful weapon. So I probably need to brush up on my letter writing skills*”.

Researcher: “*Do you think that it is better to have a large number of small objectives or a few major objectives? Why do you prefer this option? What do you see as the advantages and disadvantages of each approach?*”

Summary

10.40. There was some variety, in respect to whether or not employees preferred a few major objectives or a large number of small objectives, the disabled woman participant felt that it was not appropriate to have any objectives at all. It was felt that objectives were set in addition to work objectives which naturally flowed from the employee’s job description.

Example of responses

- 10.41. BI Admin Manager: *“I think I would prefer few major objectives because sometimes when you get the..., say you get 10 small objectives, if they are silly objectives you think ‘I can’t be bothered you know how to do them basically because they’re just a waste of time because they’re not thought through’. Whereas if we’re giving them major ones you know then they’re thought through and you can relate to them and you will spend time doing them”.*
- 10.42. Case Progression Officer: *“I think a large number of small objectives are a far easier; you can do them in a week. I know one time one of my objectives was to act up and this comes back to there being available opportunities, I had to make that objective happen. If I only had a few major ones and I didn’t have the opportunity, then my PAR would be basic even though I’ve been doing all my work to do the standard of whatever... So if I’ve got a large number of small ones you can still get the box marking you deserve and that’s how I think”.*

Researcher: *“Do you feel able to negotiate the number of objectives that you agree each year?”*

Summary

- 10.43. Most employees felt able to negotiate the number of objectives in their PAR report. However, the female part-time case worker felt that her needs and circumstances were not taken into consideration when she tried to negotiate objectives.

Example of response

- 10.44. Part-time BI Caseworker: *“ I’ve said to you earlier, you know, I try to negotiate my objectives, in particular one of them where there is a national standard and I’m getting nowhere and I just feel terribly frustrated that at the end of the day it’s a pointless exercise – my manager’s trying to set me this objective, we’re discussing it, I’m putting it down on paper and at the end of the day it won’t be achieved, not because I don’t want to achieve it, you know, I’m not being obstructive, I’m simply pointing out that it is unrealistic to set me these same objectives as you are setting somebody who is here in the office more who will have more time to achieve it”.*

Researcher: “Does your current role give you a personal sense of accomplishment? If it does, why – what is the best bit about it? If it does not, how could this be improved?”

Summary

10.45. Despite the difficulties highlighted by the participants, respondents overwhelmingly reported that their current role gave them a sense of personal accomplishment.

Example of responses

- 10.46. Area Business Manager: “Yeah. The nature of the job (and most BM’s will tell you the same) is extremely demanding. I go home absolutely drained and quite often I don’t know what I’ve done, I haven’t really done anything. I’m fighting fire and keeping plates spinning and dealing with difficulties whether they are discipline difficulties or like this morning I spent an hour giving two lawyers feedback which is constructive”.
- 10.47. Part-time BI Caseworker: “The best thing about my role, I like what I do, I like the job. I’ve been here at the CPS for sixteen years and I’m not really likely to go anywhere else, I’m not looking to go anywhere else. I like the work; I like the people I work with. I’m very experienced in what I do so I feel confident about what I’m doing and I can do things and feel that I have achieved at the end of the day and perhaps assisted other members of staff who aren’t so experienced, so that gives me a sense of achievement, what I don’t like is unnecessary pressure through other people who aren’t performing as they should do and they’re not being sort of put back on track
- 10.48. Up until I had a family, and I’ve got a young family, I would have said that wholly the only reason I was here was that I loved the work and I liked the people I worked with so I would have no need to go anywhere else. I suppose, as well, the financial gain at the end of the month is not that bad. I don’t have a problem with my pay; I think we’re paid very well for what we do. Now I have a young family, my stance on work has changed to another dimension inasmuch as I have to consider the flexibility that this job offers and I would not find it elsewhere. You know, so that is extremely important to me and probably has now become my number one priority. And that is why I will not be going anywhere else”.

Researcher: *“Would you like to see a change in the amount of variety in your current role? If so how? If not, why?”*

Summary

10.49. Respondents found their roles varied and interesting. Consequently none of the respondents interviewed reported that they would like to see the amount of variety in their roles increased.

Example of response

10.50. Chief Crown Prosecutor: *“No. My job in a sense is almost a Jack of all trades. I mentioned that in the comments before. I used to be a Lawyer, I used to be an Advocate, I used to be in Court, and I’m like the teacher who has become headmaster. Now you know, my job range is far, far wider, so I’ve got more than enough variety”.*

Researcher: *“Is there any reason why you might not feel motivated to do the best job that you can?”*

Summary

10.51. The most significant de-motivators highlighted by the respondents were instances where their performance had been criticised by their managers. In one case this resulted in inefficiency papers being served shortly after the employee in question had achieved a box mark 2 PAR rating.

Example of responses

10.52. Designated Case worker: *“I had a situation where last year I got a box 2 rating and then a fortnight later I was served with inefficiency papers because I’d had, I was singled out, basically what had happened – I’d gone to Court and I’d collapsed in Court with a suspected heart attack, I thought I’d got like a viral infection because I’d got the flu and hadn’t taken time off and gone to work. I was told to have as much time as I needed because of it, had three weeks off, came back to work and had a migraine. Well I do suffer from migraines and because I took that day off because of the migraine I then came back, was told I was inefficient and served with an inefficiency paper. I thought it was nonsense, I mean, a few weeks before being told I was a box 2, so I think my case can be classed as a disability because you just cannot work with them. I’m not talking about headache; I’m talking about the full vomiting, not being able to sleep,*

your head and your mouth not really being able to co-ordinate, so yes I think it is a disability”.

Researcher: *“Are you satisfied with your level of job security?”*

Summary

10.53. Employees were generally satisfied with their level of job security; however this satisfaction decreased in lower level posts - in particular posts at pay level A and B. The driver for this lack of security was the constant change in the public sector. Employees were anxious that their post would be deleted as part of broader efficiency savings as part of cuts across the whole of the public sector.

Researcher: *“Do you intend to find another job within the next 12 months – if so, why; if not, why? What do you like best and dislike most about working here?”*

Summary

10.54. Those employees who felt their jobs were at risk as a result of restructuring, were more likely to indicate that they intend to find another job within the next 12 months. It was also notable that the A1 Black male indicated that he intended to find another job as a result of the difficulties that he had experienced.

Example of response

10.55. A1 Support Assistant: *“I wouldn’t be here today I would be somewhere else”.*

Researcher: *“Have you ever needed to ask for support in achieving your objectives? If yes, what was the response?”*

Summary

10.56. Most employees felt that they could ask for support in achieving their objectives. Their responses to this question were broadly positive. However, the disabled woman who participated in the study indicated that she did not receive adequate support in achieving her objectives.

Example of response

10.57. *Designated Case worker*: “Yeah. I’ll go to my line manager and say ‘I really don’t understand’. Keeping up with the legislation or whatever, ‘don’t understand this’ there’s somebody who’s going to have to explain it to me like to a two year old. I’ve got no problems with going to management and asking for extra training. But sometimes the answer is ‘it’s in there, read it’. I’ve read it and I still don’t understand it. Because at that time my line manager was probably up to it with her work. But what they should be doing is saying ‘OK’ because as far as I’m aware we haven’t got a Training Officer – we have to have a Training Officer here and local training whereby you know they can do a training matrix, identify problems from the PAR”.

Example of response

(b) *Additional questions for the Countersigning Officer focus group*

Researcher: “Do you experience any difficulties in effectively acting as a Countersigning Officer?”

Summary

10.58. The Countersigning Officers participating in the study found it difficult to effectively discharge their roles as they knew very little about the actual work of the employees for whom they acted as Countersigning Officer (CSO). Consequently, they felt that their role as CSO does not really add much value.

Example of response

10.59. *Chief Crown Prosecutor*: “I finish at counter signing for people who support the level unit head. Well if I speak to that person four or five times a year that will be it – I know her, she’s been there for some years and totally reliant on his assessment and although no-one’s ever objected to what was said in counter sign I wonder what the value of it is really because the person will know I know little about their work, will know I’m reliant really on the Reporting Officer, it’s not been the case, if the Officer had objected about that Reporting Officer’s report and was looking to me to bring some objectivity or additional data to it in reality I rarely can and so although I do it – I’ve been doing the

job a long time – it doesn't take me very long to do but I don't think there's any added value in what I am doing".

Researcher: *“What do you consider the advantages and disadvantages of using the full range of assessment criteria – for example rating employees from 1 to 5 on the form?”*

Summary

10.60. Respondents considered that there are clear advantages in using the full range of assessment criteria. However, there is very little evidence of this happening in practice. Nevertheless, at least one participant was able to cite a situation where she nearly assessed an employee as box 4.

Example of response

10.61. *Trials Unit Business Manager: “Disappointed if I felt that we were in a box 4 situation with any member of staff, having said that if there were performance issues with an individual, I would have no hesitation in supporting the manager in dealing with that and then and if it was sufficiently bad, you know calling someone out of cycle to have a full report and issue a box 4 milestones ready to seek improvement. I mean we got very close to going down that road with a member of staff, what eighteen months ago, but in the event, I mean then... we'd started with issues during the probationary period, she managed to scrape through probation and we were performances slipped back and we were actually looking then at doing an out of cycle box 4 but in the event she got herself another job and left so we never quite got there but I would have no hesitation in using what was available to us if we needed to get someone to improve or to decide that their career lay elsewhere.*

Researcher: *“Can you think of any improvements to your role in the appeals process? Do you think that CSOs should be given the role of hearing the appeal?”*

Summary

10.62. The main factor that determined whether or not a manager should act as a CSO was whether or not, in reality they had any knowledge or information which would allow them to make an accurate assessment of the employee's performance. Participants also felt that there would be some value in taking the

CSO role outside of the immediate management chain of command, as this would add real or perceived objectivity to the process.

Example of response

- 10.63. Trials Unit Business Manager: *“Yeah, I mean for our unit and for my part I would like to think that most of the staff I countersign would be more than happy for me to hear an appeal because I think they would... well, I believe that they would trust me to uphold a decision if it was right, having said that if you take it out of the CSO’s hand you’ve got a more independent view on the matter and depending on the nature of the appeal there would be an occasion where you would need to have it escalated beyond the line management chain and I don’t think you should ever take that away from staff because I think it’s important to know that if you need to go beyond your immediate managers you can do”.*
- 10.64. Chief Crown Prosecutor: *“Well, we’ve got to identify really what the appeal process should be about. If it’s just the mechanics then anybody could deal with it, it would just be a question of looking at the evidence wouldn’t it, and looking at what’s happened when and has it been done properly. If it’s an attempt to look at the actual assessment marking, I suppose logically an outsider should be able to do that as well because a report should be based on the comparative between the objectives and targets that have been set and the data in the report saying how well you’ve gone that year in achieving that, so in one sense it should be possible to have somebody from the outside, a Reporting Officer route, it might bring a little more confidence to the process, I think without wanting to echo all my previous points I think that if you started with the Officer doing their own assessment of what they’ve done and then the Reporting Officer commenting on, yes that’s a good accurate assessment of the performance this year or you think that it is over in certain areas. That report then almost ought to be agreed-based on the performance that year from which I think you or I could probably make an assessment on a scale as to markings”.*

Researcher: “*What actions can be taken to more effectively support people who are under-performing (potential box 4 and 5 candidates)?*”

Summary

10.65. Respondents felt the newly appointed HR Business Partners were to be key allies in helping them more effectively support people who are under performing. The general approach favoured by participants was developmental in the first instance, in that a poor performer would be given training and the opportunity to improve. Following that developmental opportunity other avenues for improving performance would be considered.

Example of response

10.66. *It is very time consuming. People have got a feeling over a number of years that you have to keep unbelievable amounts of records to show what you're doing, and so there's a certain tension in supporting a poor performer, you really ought to be focusing really on fully supporting that person but I think being lawyers as well – certainly in a lawyer environment – people tend to keep one eye on the possible outcome of this and they should obviously keep some records of what's going on, although the main focus should be on trying to improve performance. I think people lack confidence in dealing with the improvement of performance, how do you go about it? Can you do it on your own? Are there others that can help you? Do you want some trainers involved? Do you need to set out a clear programme of what's involved? I think people struggle with this, I think there's very few people, I don't know, in the Service getting box 4's.*

10.67. *It would be very helpful if these new Human Resource Business Partners could get involved at a very early stage, identifying a poor performer. I think maybe it would be helpful to have an outside objective person coming in and saying what's the problem? What have you done? Right, this is the best way forward, and then maybe be a sort of contact point on the phone asking if it is improved, is it going well? I think we might benefit from that. I'm not sure that we've got that within the Areas at the moment.*

Researcher: “*What changes to the role of the Countersigning Officer would you find most helpful?*”

Summary

10.68. The main change suggested to the role of the CSO was whether or not the role should exist at all. This echoes earlier comments in that where the team of reports is too large to effectively assess performance, the role of the CSO was considered to be of obscure value.

Example of response

10.69. ***Area Business Manager:*** “*I would – there is an argument, a lot of people think you ditch it. I’m not convinced you should ditch it. I think you should ditch it if it’s remote and meaningless. Sometimes you have people countersigning where I did not see you or your work. In those cases I think there is an argument for doing away with it. But generally speaking I am in favour of it. If we’re having that sort of process I’m in favour of it because – and I think it is important for the individual to have the countersigning officer involved*”.

Researcher: “*How can individual PAR objectives be more effectively linked to Area/Directorate performance indicators and CPS targets?*”

Summary

10.70. There was an excellent suggestion for better linking PAR objectives to Area and Directorate performance indicators. This involved using images to show how individual objectives related to key CPS and Area targets.

Example of response

10.71. ***Area Business Manager:*** “*In South Wales they did a very good thing – one of their schemes did – what they did was if it was a national Area target on witness care they put a smiley face next to it. And if it was to do with resources and saving money it was a Euro sign or a pound sign and then two or three hieroglyphics so then when they cascaded it down in individuals, so you got an individual as an A2 in the office doing something in the Witness Care Unit you got a smiley face so that that was the link to what was in the business plan. I mean it’s not my idea, I can’t take credit for that*”.

Key Points

- (a) Existing PAR system was considered to not adequately support a performance culture;
- (b) Most employees considered the PAR rating reflected their performance. However, two African Caribbean participants felt it under rated their performance;
- (c) There is a need to ensure the competence framework is consistent with competencies in current job descriptions;
- (d) Continuous self development and diversity were the most difficult competencies to meet;
- (e) Participants felt that the guidelines could be improved to demonstrate how to obtain box 2 and above markings;
- (f) Employees felt that they did not obtain enough feedback throughout the year;
- (g) The ability to provide constructive feedback was seen as a key competence to be developed;
- (h) More effective collection of data to support the PAR review and use of the full range of the rating scale were seen key to improving the objectivity of the PAR process;
- (i) The appeals process could be improved;
- (j) Participants would like to see HR provide more guidance information to support the PAR process.
- (k) The disabled and BME respondents reported most difficulty in being able to document their performance for the purpose of the PAR review;
- (l) Some respondents believed that the PAR ratings were determined by factors other than performance;
- (m) The disabled respondent felt that it was better to have no objectives as these were regarded as additional to the duties in her job description;
- (n) All employees reported that their roles gave them a sense of personal accomplishment were adequately varied.
- (o) The disabled respondent felt she had been provided with inadequate support in achieving her objectives from the line manager;

- (p) Some respondents believed the Service had more distance to travel to fully embed a performance culture;
- (q) Countersigning Officers did not consider that their role added much value to the PAR process;
- (r) There was limited evidence of managers using the full range of box markings;
- (s) The CSO role could be abolished, as it adds little value;
- (t) The role of HR Business Partners was welcomed - respondents felt that they had real potential to challenge underperformance;
- (u) There was evidence of real innovation in linking individual performance targets to corporate objectives.

11. ANALYSIS OF PAR QUESTIONNAIRE SURVEY

Introduction

- 11.1. The primary aims of this set of analyses were to investigate general feelings regarding key aspects of the PAR process and to explore any relationships between socio-demographic characteristics and questionnaire responses.

ANALYSIS OF PAR QUESTIONNAIRE SURVEY

BACKGROUND TO DATA PREPARATION AND ANALYSIS

- 11.2. Opinions about the PAR process were elicited using a questionnaire developed from the key themes verified in stage one and the initial analysis of depth interview and focus group data. Each of the 16 opinion items on the questionnaire was presented together with a 7 point answer scale with bi-polar anchor endpoints (see a copy of the questionnaire in appendix 3). Most of the scales (questions 3-9; 11-12; 15 and 16) had anchor endpoints of 'Not At All' (1) and 'Very Much So' (7). Question 1 had end points of 'Very Unfair' (1) to 'Very Fair' (7); question 2 had endpoints of 'Very Ineffective' (1) to 'Very Effective' (7); question 10 had endpoints of 'Lots of Minor Objectives' (1) to 'A Few Major Objectives' (7) and question 14 had endpoints of 'Much Less' (1) to 'Much More' (7). So, generally a high score represented a favourable opinion.
- 11.3. For ease of overview interpretation, response ranges on the opinion items were collapsed into three general categories of negative, (response options 1-3), neutral (response option 4), and positive (response options 5-7).
- 11.4. In addition, due to low numbers of respondents within job categories A, D & E responses to this variable were also collapsed to provide two groups (A&B) and (C,D & E) each with sufficient numbers to facilitate sub-group comparisons.
- 11.5. Key analyses performed included frequency counts, chi squared tests and independent t-tests.

STATISTICAL ANALYSES

Sample Profile

11.6. The questionnaire was intended to be a snapshot of the organisation; the timescales did not allow extensive distribution of the questionnaire. Nevertheless, the sample of 131 questionnaires was large enough to allow an 85% confidence that the sample results are fairly representative of the organisation as a whole. It is based in self selection and therefore more likely to include those who feel aggrieved or hold strong views about the PAR process. Initial frequency counts of the variables are presented in summary tables 11.1 to 11.8 below.

Table 11.1 Distribution of Participant Job Roles

		job role			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	A	11	8.4	8.7	8.7
	B	51	38.9	40.2	48.8
	C	51	38.9	40.2	89.0
	D	3	2.3	2.4	91.3
	E	11	8.4	8.7	100.0
	Total	127	96.9	100.0	
Missing	declined	3	2.3		
	System	1	.8		
	Total	4	3.1		
Total		131	100.0		

Table 11.2 Proportion of Counter Signing Officers among Participants

		Counter signing officer			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	105	80.2	82.0	82.0
	Yes	23	17.6	18.0	100.0
	Total	128	97.7	100.0	
Missing	Declined	3	2.3		
Total		131	100.0		

Table 11.3 Gender Distribution of Participants

		Gender			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	70	53.4	56.0	56.0
	Male	55	42.0	44.0	100.0
	Total	125	95.4	100.0	
Missing	declined	5	3.8		
	System	1	.8		
	Total	6	4.6		
Total		131	100.0		

Table 11.4 Distribution of Participants' Ethnicity

		Ethnicity			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	White	103	78.6	95.4	95.4
	Asian	1	.8	.9	96.3
	Welsh	2	1.5	1.9	98.1
	Jewish	1	.8	.9	99.1
	Irish	1	.8	.9	100.0
	Total	108	82.4	100.0	
Missing	declined	23	17.6		
Total		131	100.0		

No BME respondents participated in the study

Table 11.5 Declared Disability Status of Participants

		Disability status			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	114	87.0	94.2	94.2
	Yes	7	5.3	5.8	100.0
	Total	121	92.4	100.0	
Missing	Declined	10	7.6		
Total		131	100.0		

Table 11.6 Declared Sexuality of Participants

		Sexuality			Cumulative Percent
		Frequency	Percent	Valid Percent	
Valid	Heterosexual	57	43.5	95.0	95.0
	Lesbian/Gay	3	2.3	5.0	100.0
	Total	60	45.8	100.0	
Missing	Declined	71	54.2		
Total		131	100.0		

Table 11.7 Descriptives of Participants' Age

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Age	120	27	62	42.73	8.012
Valid N (listwise)	120				

The youngest respondent was 27 years of age and the older 62 years and the average age was 43 years.

Table 11.8 Descriptives of Participants' Length of Service

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Length of Service	126	2	26	13.58	6.277
Valid N (listwise)	126				

11.7. It is interesting to note that whilst relatively few people declined to report their gender (4%); ethnicity (18%); or disability status (8%); more than half the participants declined to report their sexuality (54%).

11.8. To check for sample representativeness, chi squared Goodness-of Fit tests were run on socio-demographic attributes for which appropriate comparative population data were available. Expected distributions of gender; declared disability status; and ethnicity (using only 2 categories of White and BME) were calculated from the organisation's personnel data base for 2004. The results of

these representation checks are shown in tables 11.9 (gender), 11.10 (disability) and 11.11 (ethnicity).

Table 11.9 Summary of Test for Sample Representativeness in terms of Gender

Gender			
	Observed N	Expected N	Residual
Female	70	83.8	-13.8
Male	55	41.3	13.8
Total	125		

$\chi^2 = 6.84$ df=1; p=<.01

Compared to the distribution of men and women within this organisation, this sample has an over-representation of men and an under-representation of women.

Table 11.10 Summary of Test for Sample Representativeness in terms of Disability

Disability status			
	Observed N	Expected N	Residual
No	114	113.7	.3
Yes	7	7.3	-.3
Total	121		

$\chi^2 = .010$; df =1; p =.921

The distribution of declared disability and non-disability status within this sample fits almost exactly with that found in the organisation as a whole.

Table 11.11 Summary of Test for Sample Representativeness in terms of Ethnicity

Ethnicity			
	Observed N	Expected N	Residual
White	103	90.5	12.5
BME	1	13.5	-12.5
Total	104		

$\chi^2 = 13.33$; df =1; p = <.001

Compared to the distribution of White and BME personnel within this organisation, this sample has an over-representation of White and a significant under-representation of BME participants

RESPONSES TO EACH QUESTIONNAIRE OPINION ITEM

- 11.9. Frequencies and mean scores were calculated for each questionnaire item to indicate strength of feeling within the study sample (a full copy of questionnaire frequencies is given in appendix 3).
- 11.10. Only a minority of respondents (26%) expressed any positive intention to leave the organisation within the next 12 months.
- 11.11. Summaries of frequencies within the three broad categories of negative, neutral and positive responses are presented in table 11.12.
- 11.12. Table 11.13 shows the overall mean score and standard deviation for each opinion item.

Performance Appraisal Ratings: a Review and Impact Assessment

Table 11.12 Summary of Negative, Neutral and Positive Responses to each Item

1) Overall, in your opinion, how fair is the PAR rating method?	Unfair 77%	Neutral 10%	Fair 13%
2) Overall, in your opinion, how effective is the PAR rating method?	Ineffective 88%	Neutral 8%	Effective 4%
3) Do you feel that your last PAR rating accurately reflected your performance?	No 39%	Neutral 17%	Yes 44%
4) Do you feel that you had adequate feedback about your job performance throughout the year?	65%	11%	24%
5) Do the standards of behaviour shown by appraisers meet your expectations?	46%	19%	35%
6) Do you feel that the appraisal system could be made to be more objective?	19%	19%	62%
7) Are you satisfied with the current mechanisms for appealing against PAR reports?	44%	35%	21%
8) Do you feel uncomfortable in writing down information to demonstrate how you have performed against each job competence?	47%	13%	40%
9) Do you think the PAR process is an effective way of recognising individuals who perform well?	87%	5%	8%
11) Do you feel able to negotiate the number of objectives that you agree each year?	45%	19%	36%
12) Have you ever needed to ask for support in achieving your objectives?	59%	15%	26%
15) Do you intend to find another job within the next 12 months?	61%	13%	26%
16) How successful do you consider the CPS has been in introducing a performance culture?	84%	8%	8%
10) Do you think that it is better to have a large number of small objectives or a few major objectives?	Lots of Minor 10%	Neutral 21%	A few Major 69%
14) Would you like to see a change in the amount of variety in your current role?	Less 19%	Neutral 55%	More 26%

Please note that for ease of presentation, the questionnaire items do not run in exact numerical sequence

Performance Appraisal Ratings: a Review and Impact Assessment

Table 11.13 Summary of Mean Scores for each Opinion Item (using 7 point scale) Presented in Descending Order of Mean Scores

Descriptive Statistics					
	N	Min	Max	Mean	Std. Deviation
prefer lots of minor or a few major objectives?	128	1	7	5.51	1.651
could the appraisal system could be made to be more objective?	130	1	7	5.02	1.811
like less or more variety in job role?	128	1	7	4.17	1.404
did your last PAR rating accurately reflect your performance?	130	1	7	4.08	1.995
feel uncomfortable writing competence evidence?	131	1	7	3.89	2.283
can you negotiate number of objectives set?	130	1	7	3.78	2.000
do appraisers' behaviours meet your expectations?	115	1	7	3.77	2.010
satisfied with current PAR appeal mechanisms?	125	1	7	3.38	1.678
you ever needed to ask for support in achieving your objectives?	128	1	7	2.97	1.827
did you have adequate feedback about your job performance?	130	1	7	2.92	1.858
intend to find another job within the next 12 months?	128	1	7	2.91	2.107
how fair is the PAR rating method?	131	1	7	2.66	1.455
how fair is the PAR rating method?	131	1	6	2.21	1.380
how effective is the PAR rating method?	131	1	6	1.95	1.122
has CPS successfully introduced performance culture?	131	1	7	1.93	1.458
PAR an effective way of recognising good performers?	131	1	6	1.89	1.305
	101				

Bearing in mind that in general, the higher the score, the more positive the response, it is clear from tables 11.12 and 11.13 that participants expressed quite strong preferences for a few major objectives rather than lots of minor objectives to be set. It is also evident that they expressed a clear scepticism regarding the perceived effectiveness and fairness of the current PAR system. However, it is interesting to note that among the generally negative responses, almost half the respondents felt that their last PAR rating accurately reflected their performance.

11.13. The sample obtained had a high representation of White, able bodied people, approximately half of whom declined to declare their sexuality. Thus, sub group comparisons were only directly feasible for gender and collapsed job role (as described in the data preparation, section of this report). However, whilst a

Performance Appraisal Ratings: a Review and Impact Assessment

direct comparison by sexuality was not possible, it was considered appropriate to compare responses on the basis of people who did and did not disclose their sexuality, the difference in group size between Countersigning Officers and non-Countersigning Officers was quite strongly marked, but as this difference in group sizes within the sample is assumed to reflect differences within the organisational populations,, some comparisons between these two groups were made.

GENDER

11.14. There were no significant differences found in responses between men and women.

COLLAPSED JOB ROLES

11.15. The collapsed job roles used in this analysis were those of roles A and B against roles C, D and E. The only statistically significant differences found between the responses of the two groups were in relation to Q2 (How effective is the PAR rating method); Q7 (satisfaction with current PAR appeal mechanism) and Q9 (How effective is the PAR process as a way of recognising individuals who perform well). Summaries of these tests are shown in tables 11.14 to 11.16.

Table 11.14 Summary of t-test of collapsed job role differences for opinion item 2 (overall effectiveness of PAR)

	<i>n</i>	<i>Group mean</i>	<i>t</i>	<i>df</i>	<i>p</i>
<i>A & B</i>	62	2.19	2.23	125	.027
<i>C, D & E</i>	65	1.75			

Employees at pay levels A and B felt that the PAR system was more effective than employees graded at levels C, D and E.

Table 11.15 Summary of t-test of collapsed job role differences for opinion item 7 (satisfaction with current PAR appeal mechanism)

	<i>n</i>	<i>Group mean</i>	<i>t</i>	<i>df</i>	<i>p</i>
A & B	59	3.71	2.16	119	.032
C, D & E	62	3.06			

Employees at pay levels A and B were more satisfied with the mechanism for appealing against PAR reports than employees at pay levels C, D and E.

Table 11.16 Summary of t-test of collapsed job role differences for opinion item 9 (effectiveness for recognising good performers)

	<i>n</i>	<i>Group mean</i>	<i>t</i>	<i>df</i>	<i>p</i>
A & B	62	2.13	2.03	108.8 equal variances not assumed	.045
C, D & E	65	1.66			

Whilst the means for both groups indicate a general lack of faith in PAR effectiveness and satisfaction with the current mechanism for PAR appeals, participants in the higher job roles group gave even less favourable responses than those in the lower level groups.

SEXUALITY

11.16. All but two participants who disclosed their sexuality were heterosexual. The two gay participants were excluded from analysis and comparisons were made between the disclosed heterosexual participants and those participants who did not disclose their sexuality. The only statistically significant differences found between the two groups were in terms of Q1 (PAR is a fair system) and Q3 (PAR rating accurately reflected performance). Summaries of these results are presented in Tables 11.17 and 11.18.

Table 11.17 Summary of t-test of disclosed and undisclosed sexuality participants differences for opinion item 1 (PAR is a fair system)

	<i>n</i>	<i>Group mean</i>	<i>t</i>	<i>df</i>	<i>p</i>
Sexuality not disclosed	71	2.38	2.49	97.43 equal variances not assumed	.015
Heterosexual	57	3.04			

Performance Appraisal Ratings: a Review and Impact Assessment

Table 11.18 Summary of t-test of disclosed and undisclosed sexuality participants differences for opinion item 3 (last PAR accurately reflected performance)

	n	Group mean	t	df	p
Sexuality not disclosed	71	3.75	2.29	125	.024
Heterosexual	56	4.55			

Those who disclosed their sexuality as heterosexual gave consistently more positive responses than those who did not disclose their sexuality, across both of these statistically significant results.

COUNTERSIGNING OFFICERS

11.17. The only statistically significant differences found between the responses of Countersigning Officers and non-Countersigning Officers were in relation to Q10 (preferred number of objectives set) and Q16 (successful introduction of performance culture). Summaries of these tests are shown in tables 11.19 and 11.20.

Table 11.19 Summary of t-test of CSOs and Non-CSOs differences for opinion item 10 (preferred number of objectives)

	n	Group mean	t	df	p
Non-CSO	102	5.38	2.68	47.3 equal variances not assumed	.010
CSO	23	6.17			

Table 11.20 Summary of t-test of CSOs and Non-CSOs differences for opinion item 16 (successful introduction of performance culture)

	n	Group mean	t	df	p
Non-CSO	105	1.76	2.97	1.26	.004
CSO	23	2.74			

Although both groups expressed preference for a few major objectives over lots of minor objectives, the counter signing officers were even more pronounced in their preference. Whilst neither group considered the organisation to be very successful in the introduction of a performance culture, CSOs did give significantly more positive responses than non-CSOs. However, the large size discrepancy for these groups should be considered when interpreting these data.

AGE AND LENGTH OF SERVICE

- 11.18. Correlation analyses were performed to investigate any correlations between questionnaire opinion items and age, and with length of service.
- 11.19. There was a positive correlation between age and feeling uncomfortable in giving written evidence for job competencies ($r = .195$; $n=120$ $p=.03$), so older people were more likely than younger people to report discomfort.
- 11.20. Tenure was negatively correlated with intention to leave ($r = -.268$; $n=124$; $p=.003$), so newer recruits were more likely than those with longer length of service to report an intention to seek a new job within the next twelve months.
- 11.21. Tenure was also significantly negatively correlated with preferences for job variety ($r = -.202$; $n=124$; $p=.024$), indicating that the longer people had been in post the less likely they were to seek more variety in their job role.

KEY POINTS

Representation

- (a) The sample was representative in terms of disability;
- (b) Men were overrepresented in the sample;
- (c) BME respondents were significantly underrepresented in the sample.

Overall responses

- (a) Almost half of responding employees felt their PAR report reflected their performance (44%);
- (b) Overall employees felt comfortable (47%) writing down information to demonstrate how they perform against each job competence;
- (c) The majority of employees have not needed to ask for support in achieving their objectives (59%);
- (d) The majority of survey respondents intended to stay with the CPS for the next 12 months (61%). A minority of respondents (26%)

- expressed the intention to leave the organisation within the next 12 months;
- (e) A majority of employees felt that the PAR rating method was unfair (77%);
 - (f) The majority of employees felt the PAR rating method was ineffective (88%);
 - (g) Most employees 65% did not feel there was adequate feedback about their performance;
 - (h) The majority of employees (46%) did not feel the standards of behaviour showed by appraisers meet their expectations;
 - (i) A majority of staff felt that the PAR system could be made more objective (62%);
 - (j) Almost half of responding employees (44%) did not feel satisfied with the current mechanisms for appealing against PAR reports;
 - (k) The vast majority of employees (87%) did not consider the PAR process to be an effective method of recognising individuals who perform well;
 - (l) Almost half of participating employees felt unable to negotiate the number of objectives that they have to achieve each year (45%);
 - (m) Employees strongly believed that the CPS had not been successful in introducing a performance culture (84%);
 - (n) The majority of employees (69%) felt that it was better to have a few major objectives rather than a large number of small objectives;
 - (o) Most employees (55%) were neutral about whether they wanted to see a change in the amount of variety in their current roles.

Free Response Items

- (a) Staff were mainly dissatisfied with the removal of rewards such as pay, and promotion from the PAR system;
- (b) The next most important demotivator was the failure to manage poor performance;
- (c) The third major concern was that the box marking system was considered inadequate.

Demographic group differences

Pay level

- (a) Employees at pay levels A to B were more satisfied with the appeal mechanisms that employees at pay levels C, D and E;
- (b) Staff at pay levels A to B were more satisfied with the effectiveness of the PAR system in recognising performance than employees at pay levels C to E;

Sexuality

- (c) Employees who disclosed their sexuality as heterosexual were more likely than those who did not disclose their sexuality to state the PAR system was a fair system and that their PAR review accurately reflected their performance;

CSOs

- (d) CSOs were more likely than non CSOs to prefer a few major objectives and were more likely to state that the CPS has been successful in introducing a performance culture;

Age

- (e) Older people were more likely to report discomfort with providing written evidence to support their PAR reviews;

Length of Service

- (f) Newer recruits were more likely to leave. Newer recruits were also more likely to want more variety in their job role.

12. IMPACT ASSESSMENT CONCLUSIONS

ASSESSMENT OF LIKELY IMPACT

- 12.1. In this section of the report the assessment of the impact of the current performance appraisal review process is considered in light of the qualitative and quantitative evidence obtained in phase one and phase two of this research. In particular, in this section we consider whether the current PAR process could have an adverse impact upon key demographic groups.
- 12.2. Insofar as the statistical evidence is concerned, there is no evidence which demonstrates that there is a statistically significant difference in PAR box markings amongst the various demographic groups reviewed. In respect of the evidence relating to disabled employees we conclude that there are statistically significant differences, these differences require further research to determine the reason for them.
- 12.3. It is important to note that the qualitative research in this study presented significantly different findings than the quantitative data in a number of respects. For example, one African Caribbean man expressed stronger perceptions of unequal treatment than were immediately apparent from the quantitative analysis. However, this could possibly be because there were no African Caribbean men in the quantitative survey sample. There were, however some consistencies in that the perceptions expressed by the disabled worker in the qualitative study supported the findings of phase one. We also found that an employee engaged in alternative working patterns highlighted very strong perceptions of unequal treatment.
- 12.4. The Service needs to consider improvements to the current mechanism for monitoring and managing performance. A number of these improvements have already been suggested as part of the qualitative analysis. These suggestions need to be developed and refined into models for monitoring and managing

performance. More detailed consultation on the new PAR framework needs to be undertaken before a new performance appraisal system is installed.

Overall conclusion

- 12.5. Having considered the evidence from both phases of the impact assessment we conclude that there is insignificant evidence which demonstrates that there is an adverse impact against any particular demographic group arising out of the operation of the current PAR system. The Service has been proactive in frankly identifying employee views about the PAR processes and equality proofing the process. There was evidence of a strong commitment to continuous improvement.

13. RECOMMENDATIONS

13.1. It is recommended that:

- (a) key stakeholders are engaged in consultation about the performance appraisal system;
- (b) HR business partners become fully engaged in working with Areas and Directorates to eliminate poor performance fairly;
- (c) the issues raised in this review will be taken forward as part of the planned review of PAR in 2006–2007. The Human Resources Director considers developments to the PAR system within the resources available;
- (d) developmental work with disabled employees and other key groups is undertaken. This work should help identify how behaviour, social relationships and attitudes impact on the performance and motivation of key employee groups;
- (e) managers ensure that all the key issues and documents which have a bearing on the final PAR record are traceable and identifiable;
- (f) objectives should be clearly identifiable in the appraisal report so that the Countersigning Officer and others can make an assessment of how judgements have been made;
- (g) all staff have a maximum of six objectives which are commensurate with the grade of the post. It is also imperative to ensure that objectives are checked to ensure that they are SMART, with measurable timescales and links to service plans;
- (h) managers' comments are clearly linked to the competence levels, headings and standards;
- (i) the CPS will consider providing training which uses role play and coaching to help managers to more effectively manage poor performance;
- (j) greater individual and organisational responsibility needs to be taken for organisational change and development;
- (k) the Service considers staggering the PAR reporting cycle so that managers with large teams will have more time to invest in completing quality PAR reviews;

- (l) the CPS will consider whether the continued use of the box marking system is the most effective way of giving feedback on performance. In this connection the Service will consider whether the identification of development needs in relation to achieving organisational objective should be the key output of the PAR report;

- (m) CPS competencies are reviewed and updated.

Dr Bernard Horsford
Chief Executive
Sankofa Exchange Limited

14. APPENDIX 1: DETAILS OF CROSS-TABULATION ANALYSIS IN MAIN REPORT

Table A1. Cross-tabulation Analysis Disability Status * Gender

			Sex		<i>Total</i>	
			male	female		
Disability Status	non-disabled	Count	1550	3210	4760	
		Expected Count	1561.4	3198.6	4760	
	disabled	Count	100	170	270	
		Expected Count	88.6	181.4	270	
	<i>Total</i>		<i>Count</i>	1650	3380	5030

	Value	df	Asymp. Sig (2 sided)
Pearson Chi Square	2.32	1	.128

Table A2. Cross- tabulation Analysis Disability Status * Ethnicity (White & BME)

			Ethnicity		<i>Total</i>	
			White	BME		
Disability Status	non-disabled	Count	4142	618	4760	
		Expected Count	4156.2	603.8	4760	
	disabled	Count	250	20	270	
		Expected Count	235.8	34.2	270	
	<i>Total</i>		<i>Count</i>	4392	638	5030

	Value	df	Asymp. Sig (2 sided)
Pearson Chi Square	7.17	1	.007

Table A3. Cross-tabulation Analysis Disability Status * Lawyer/Non Lawyer

			Professional Status		<i>Total</i>	
			Non-Lawyer	Lawyer		
Disability Status	non-disabled	Count	3309	1451	4760	
		Expected Count	3312.1	1447.9	4760	
	disabled	Count	191	79	270	
		Expected Count	187.9	82.1	270	
	<i>Total</i>		<i>Count</i>	3500	1530	5030

Table A4. Cross-tabulation Analysis Gender * Ethnicity

			Ethnicity		<i>Total</i>
			White	BME	
Gender	male	Count	1470	180	1650
		Expected Count	1440.7	209.3	1650
	female	Count	2922	458	3380
		Expected Count	2951.3	428.7	3380
	<i>Total</i>	<i>Count</i>	4392	638	5030

	Value	df	Asymp. Sig (2 sided)
Pearson Chi Square	6.98	1	.008

Table A5. Cross-tabulation Analysis Gender * Lawyer/Non-Lawyer

			Professional Status		<i>Total</i>
			Non-Lawyer	Lawyer	
Gender	male	Count	891	759	1650
		Expected Count	1148.1	501.9	1650
	female	Count	2609	771	3380
		Expected Count	2351.9	1028.1	3380
	<i>Total</i>	<i>Count</i>	3500	1530	5030

	Value	df	Asymp. Sig (2 sided)
Pearson Chi Square	281.7	1	<.001

Table A6. Cross-tabulation Analysis Ethnicity* Lawyer/Non-Lawyer

			Professional Status		Total
			Non-Lawyer	Lawyer	
Ethnicity	White	Count	3034	1358	4392
		Expected Count	3056.1	1335.9	4392
	BME	Count	466	172	638
		Expected Count	443.9	194.1	638
	Total	Count	3500	1530	5030

	Value	df	Asymp. Sig (2 sided)
Pearson Chi Square	4.13	1	.042

	Value	df	Asymp. Sig (2 sided)
Pearson Chi Square	.181	1	.671

15. APPENDIX 2: DETAILS OF FACTORIAL ANOVA RESULTS

Table A7. Between-Subjects Factors

Factors		Value label	N
Disability Status	0	Not disabled	4760
	1	disabled	270
Gender	0	male	1650
	1	female	3380
Ethnicity	0	White	4392
	1	BME	638

Table A8. Tests of Between-Subjects Effects

Dependent Variable: PAR Rating

Source	Sum of Squares	df	Mean Square	F	Sig
Corrected Model	5.995	7	.865	3.09	.003
Intercept	1830.646	1	1830.646	6614.18	.000
Disability	2.296	1	2.296	8.29	.004
Gender	.144.859	1	.144.859	.521	.470
Ethnicity	.859	1	.859	3.10	.078
Disability*Gender	.341	1	.341	1.23	.267
Disability*Ethnicity	.459	1	.459	1.66	.198
Gender*Ethnicity	.637	1	.637	2.30	.129
Disability*Gender*Ethnicity	.517	1	.517	1.87	.172
Error	1389.97	5022	.277		
Total	34336.00	5030			
Corrected Total	1395.963	5029			

Performance Appraisal Ratings: a Review and Impact Assessment

Table A9. Estimated Marginal Means

Dependent Variable: PAR Rating

				Mean	Std Error	
DISABILITY		Non-Disabled		2.56	.013	
		Disabled		2.75	.064	
GENDER		Male		2.68	.053	
		Female		2.63	.039	
ETHNICITY		White		2.59	.018	
		BME		2.71	.063	
DISABILITY*GENDER		Non-Disabled	Male	2.55	.021	
			Female	2.57	.013	
		Disabled	Male	2.81	.103	
			Female	2.69	.076	
DISABILITY*ETHNICITY		Non-Disabled	White	2.54	.009	
			BME	2.57	.024	
		Disabled	White	2.65	.034	
			BME	2.85	.123	
GENDER*ETHNICITY		Male	White	2.57	.028	
			BME	2.78	.101	
		Female	White	2.62	.022	
			BME	2.64	.074	
DISABILITY * GENDER * ETHNICITY		Non-Disabled	Male	White	2.53	.014
				BME	2.57	.040
			Female	White	2.56	.010
				BME	2.58	.025
		Disabled	Male	White	2.61	.055
				BME	3.00	.199
			Female	White	2.68	.042
				BME	2.69	.146

16. APPENDIX 3: QUESTIONNAIRE & FREQUENCIES

(a) PAR REVIEW QUESTIONNAIRE

How to complete your written submission

The CPS is currently carrying out research into the performance appraisal process. If you have not attended a focus group and are not available for interview you may participate by making a written submission. Your response is totally confidential. We will remove any information which could identify you from your response.

Please place an X in the box that accords with your views about the PAR process. There are no right or wrong answers. If you feel that there are additional issues that you would like to us to consider or if you would like to expand on your responses to earlier questions, please use the attached sheet to do this.

Where to return your submission

Please return your completed questionnaire by email to CPS@sankofa.co.uk or by post to Dr Bernard Horsford, Sankofa Exchange Limited, Africa House, 21 Shorwell Road, Nottingham NG3 7HG. The closing date for receipt of submissions is 7 June 2005.

Background information about you

It would be useful to get some background information about you (answers are voluntary).

1. Job Title/Role: _____
2. Are you a Counter Signing Officer? Yes/No
3. Gender? _____ Male/Female
4. Age (approx)? _____
5. How would you describe your ethnic origin? _____
6. Do you have a disability? Yes/No
7. How would you describe your sexual orientation _____
8. How long have you worked for the CPS _____

Performance Appraisal Ratings: a Review and Impact Assessment

1	Overall, in your opinion, how fair is the PAR rating method?								
	<i>Very Unfair</i>	1	2	3	4	5	6	7	<i>Very Fair</i>
2	Overall, in your opinion, how effective is the PAR rating method?								
	<i>Very Ineffective</i>	1	2	3	4	5	6	7	<i>Very Effective</i>
3	Do you feel that your last PAR rating accurately reflected your performance?								
	<i>Not at All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
4	Do you feel that you had adequate feedback about your job performance throughout the year								
	<i>Not at All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
5	Do the standards of behaviour you shown by appraisers meet your expectations?								
	<i>Not at All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
6	Do you feel that the appraisal system could be made to be more objective?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
7	Are you satisfied with the current mechanisms for appealing against PAR reports?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
8	Do you feel uncomfortable in writing down information to demonstrate how you have performed against each job competence?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>

Performance Appraisal Ratings: a Review and Impact Assessment

9	Do you think the PAR process is an effective way of recognising individuals who perform well?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
10	Do you think that it is better to have a large number of small objectives or a few major objectives?								
	<i>Lots of Minor Objectives</i>								<i>A few Major Objectives</i>
11	Do you feel able to negotiate the number of objectives that you agree each year?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
12	Have you ever needed to ask for support in achieving your objectives?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
13	Overall, in your opinion, how fair is the PAR rating method?								
	<i>Unfair</i>	1	2	3	4	5	6	7	<i>Fair</i>
14	Would you like to see a change in the amount of variety in your current role?								
	<i>Yes, Much Less</i>	1	2	3	4	5	6	7	<i>Yes, Much More</i>
15	Do you intend to find another job within the next 12 months								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
16	How successful do you consider the CPS has been in introducing a performance culture?								

Please use this space to make any other written comments about the PAR process

Performance Appraisal Ratings: a Review and Impact Assessment

(b) Questionnaire frequencies

1	Overall, in your opinion, how fair is the PAR rating method?								
	<i>Very Unfair</i>	1	2	3	4	5	6	7	<i>Very Fair</i>
		33	33	35	13	11	4	2	
2	Overall, in your opinion, how effective is the PAR rating method?								
	<i>Very Ineffective</i>	1	2	3	4	5	6	7	<i>Very Effective</i>
		57	44	15	10	4	1	0	
3*	Do you feel that your last PAR rating accurately reflected your performance?								
	<i>Not at All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
		18	18	15	22	16	24	17	
4*	Do you feel that you had adequate feedback about your job performance throughout the year								
	<i>Not at All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
		38	33	13	15	15	10	6	
5*	Do the standards of behaviour you shown by appraisers meet your expectations?								
	<i>Not at All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
		22	14	17	22	8	20	12	
6*	Do you feel that the appraisal system could be made to be more objective?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
		9	4	11	25	21	22	38	
7*	Are you satisfied with the current mechanisms for appealing against PAR reports?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much</i>

Performance Appraisal Ratings: a Review and Impact Assessment

									So
		27	13	15	44	12	10	4	
8	Do you feel uncomfortable in writing down information to demonstrate how you have performed against each job competence?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
		31	15	15	17	14	8	31	
9	Do you think the PAR process is an effective way of recognising individuals who perform well?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
		73	30	11	7	7	3	0	
10*	Do you think that it is better to have a large number of small objectives or a few major objectives?								
	<i>Lots of Minor Objectives</i>	1	2	3	4	5	6	7	<i>A few Major Objectives</i>
		5	2	6	27	8	30	50	
11*	Do you feel able to negotiate the number of objectives that you agree each year?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
		22	21	16	25	10	22	14	
12*	Have you ever needed to ask for support in achieving your objectives?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
		40	25	11	19	21	8	4	
13	Overall, in your opinion, how fair is the PAR rating method?								
	<i>Unfair</i>	1	2	3	4	5	6	7	<i>Fair</i>
		54	35	18	14	5	5	0	
14*	Would you like to see a change in the amount of variety in your current role?								
	<i>Yes, Much Less</i>	1	2	3	4	5	6	7	<i>Yes, Much More</i>
		8	3	14	70	12	9	12	
15*	Do you intend to find another job within the next 12 months								

Performance Appraisal Ratings: a Review and Impact Assessment

	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
		56	13	9	17	14	7	12	
16	How successful do you consider the CPS has been in introducing a performance culture?								
	<i>Not At All</i>	1	2	3	4	5	6	7	<i>Very Much So</i>
		79	20	11	11	5	3	2	

** Please note that whilst the overall sample was comprised of 131 people, not all participants gave responses to all items, therefore the sum of responses does not total 131 in all cases. Items with missing data are identified by an asterisk.*

17. REFERENCES

Alderfer, C. P. (1986). An intergroup perspective on group dynamics. In Lorsch, J. (Ed.), *Handbook of Organisational Behaviour* (pp. 190–222). Englewood Cliffs, NJ: Prentice-Hall.

Andrew, F. (1996). *Not Just Black and White: An Exploration of Opportunities and Barriers to the Development of Black Managers in Public Services*. London: Office for Public Management.

Baldi, S., & McBrier, D. B. (1997). Do the determinants of promotion differ for Blacks and Whites? *Work and Occupations*, 24(4), 478–497.

Bass, A. R., & Turner, J. N. (1973). Ethnic group differences in relationships among criteria of job performance. *Journal of Applied Psychology*, 57(2), 101-109.

Cianni, M., & Romberger, B. (1995). Perceived racial, ethnic and gender differences in access to developmental experiences. *Group and Organization Management*, 20(4), 440–459.

Connor, H., La Valle, I., Tackey, N., & Perryman, S. (1996). *Ethnic Minority Graduates: Differences By Degrees: IFS Report 309*. Brighton, UK: Institute for Employment Studies.

Cooper, C. L., & Melhuish, A. (1984) Executive stress and health: Differences between men and women. *Journal of Occupational Medicine*, 26(2), 99–103.

Dansereau, F., Graen, G., & Haga, W. J. (1975). A vertical dyadic linkage approach to leadership within formal organizations. *Organizational Behaviour and Human Performance*, 13, 46–70.

Davison, M. J. (1997). *The Black and Ethnic Minority Women Manager: Cracking the Concrete Ceiling*. London: Paul Chapman.

Performance Appraisal Ratings: a Review and Impact Assessment

Davison, M. J., & Cooper, C. L. (1992). *Shattering the Glass Ceiling: The Woman Manager*. London: Paul Chapman.

Davison, M. J., & Cooper, C. L. (1983). *Stress and the Woman Manager*. Oxford: Martin Robertson.

Dewberry, C. (2001). Performance disparities between whites and ethnic minorities: Real differences or assessment bias? *Journal of Occupational and Organisational Psychology*, 74, 659–673.

Eichinger, R. (2000). *In search of the Holy Grail*. Paper presented at Conference on Organizational Development, London.

Ford, J. K., Kraiger, L., & Schechtman, S. C. (1986). Study of race effects in objective indexes and subjective evaluations of performance: A meta-analysis of performance criteria. *Psychological Bulletin*, 99(3), 330–337.

Greenberg, J. (1986). Determinants of perceived fairness of performance evaluations. *Journal of Applied Psychology*, 71(2), 340–342.

Greenhaus, J. H., & Parasuraman, S. (1993). Job performance attributions and career advancement prospects: An examination of gender and race effects. *Organizational Behaviour and Human Decision Processes*, 55, 273–297.

Greenhaus, J. H., Parasuraman, S., & Wormley, W. M. (1990). Effects of race on organisational experiences, job-performance evaluations, and career outcomes. *Academy of Management Journal*, 33(1), 64–86.

Hattrup, K., Rock, J., & Scalia, C. (1997). The effects of varying conceptualizations of job performance on adverse impact : Minority hiring and predicted job performance. *Journal of Applied Psychology*, 82(5), 656–664.

Performance Appraisal Ratings: a Review and Impact Assessment

Hedge, J. W., & Kavanagh, M. J. (1988). Improving the accuracy of performance evaluations: Comparison of three methods of performance appraiser training. *Journal of Applied Psychology*, 73(1), 68–73.

Horsford, B. I. (2003) *The career progression of Black managers*. Thesis submitted in partial fulfilment of PhD, Cranfield University, UK.

Ilgen, D. R., & Youtz, M. A. (1986). Factors affecting the evaluation and development of minorities in organizations. In: Rowland, K. M. & Ferris, G. R. (Eds.), *Research in Personnel and Human Resources Management* (Vol. 4, pp. 307–337). Greenwich, CT: JAI Press.

Institute of Employment Studies (IRS) (2001). *Equality in Performance Review*. Brighton, UK: Capita and Institute of Employment Studies.

Jones, T. (1993). *Britain's Ethnic Minorities*, London: Policy Studies Institute.

Kraiger, K., & Ford, J. K. (1985). A meta-analysis of ratee race effects in performance ratings. *Journal of Applied Psychology*, 70, 56–65.

Madood, T., Berthoud, R., Lakey, J., Nazroo, J., Smith, P., Virdee, S., & Beishon, S. (1997). *Ethnic Minorities in Britain: Diversity and Disadvantage*. London: Policy Studies Institute.

Landau, J. (1995). The relationship of race and gender to managers' ratings of promotion potential. *Journal of Organisational Behaviour*, 16(4), 391–400.

Leong, F. T. L. (2001). The role of acculturation in the career adjustment of Asian-American workers: A test of Leong and Chou's (1994) formulations. *Cultural Diversity and Ethnic Minority Psychology*, 7(3), 263–273.

Mishra, S. P. (1980). The influence of examiner's ethnic attributes on intelligence tests scores. *Psychology in Schools*, 17(1), 117–122.

Performance Appraisal Ratings: a Review and Impact Assessment

Pulakos, E. D., & Wexley, K. N. (1983). 'The relationship among perceptual similarity, sex and performance ratings in manager-subordinate dyads. *Academy of Management Journal*, 26, 129–139.

Pulakos, E., Schmitt, N., & Chan, D. (1996). Models of job performance ratings: An examination of rater race, ratee gender, and rater level effects. *Human Performance*, 9(2), 103–119.

Rick, J., Tamkin, P., Pollard, E., Tackey, N., & Heron, P. (1999). *Organisational and Managerial Implications of Devolved Personnel Assessment Processes*. Brighton, UK: Institute for Employment Studies, for Greater London Employers Association.

Runnymede Trust (2000). *Moving On Up? Racial Equality and The Corporate Agenda: A study of FTSE 100 Companies*. London: Runnymede Trust.

Tackey, N. D., Tamkin, P., & Sheppard, E. (2001). *The Problem of Minority Performance in Organisations*. Brighton, UK: Institute of Employment Studies, Report No. 375.

Terrell, F., Terrell, S. L., & Taylor, J. (1981). Effects of race of examiner and cultural mistrust on the WIAS performance of Black students. *Journal of Consulting and Clinical Psychology*, 49, 750–751.

Trades Union Congress (TUC) (1999). *Black and Betrayed*. London: TUC.

Trades Union Congress (TUC) (2000). *Qualifying for Racism*. London: TUC.

Uniform Guidelines on Employee Selection (1978). *Federal Register*, 43, 38290–38315.

Waldman, D. A., & Avolio, B. J. (1991). Race effects in performance evaluations: Controlling for ability, education and experience. *Journal of Applied Psychology*, 76(6), 897–901.