Analysis of Civil Service Performance Management Diversity and Inclusion outcomes data 2015-6

Steve French, Keele University

Introduction

This report has been commissioned by the PCS to provide an evaluation of the data provided by Civil Service Human Resources on the variation in performance management rating outcomes for the review year 2015-16. The report, following the structure used in this analysis last year[1], examines whether variations in performance management outcomes by gender, ethnicity, disability, sexual orientation, age, working patterns or grade are statistically significant, both by Government Department and across the Civil Service.

The report is structured as follows. In the first section the methods used to undertake the statistical analysis are explained, along with comments on the limitations of these methods as well as the performance management data provided. The second section then seeks to interpret these findings to identify trends in performance management outcomes and where these indicate a cause for concern, notably in relation to staff with *"protected characteristics"* under the 2010 Equality Act. The final section of the report seeks to undertake a degree of comparison with the 2014-2015 data, though issues of comparative data and categorisation limit the extent to which robust conclusions can reached from such a comparison.

1. Overview of the statistical analysis undertaken

The presentation of the data, as aggregated and disaggregated performance management ratings by department, does limit the extent to which statistical analysis can be undertaken. The most appropriate statistical method for analysing these data are chi-square tests. The aim of these tests is to compare the actual ratings results against an expected distribution of the ratings. The chi-square test then identifies the extent to which the actual results differ from the expected distribution. This is the so-called 'P value', which will be a score between 1 and 0. The P value allows the analyst to determine the extent to which the variations in the ratings found differ *significantly* from the expected distribution. For the purposes of this analysis the researcher has used the standard rating system of significance, identifying a P value of <0.05 to be significant (*) and identifying results of higher statistical significance (P <0.01 ** and P <0.001 ***). It should be noted that in the case of a P value of <0.001, there is less than one in a thousand chance of a statistically significant difference between actual and expect results being wrong. The P values are provided for each department for each grouping of disaggregated data (gender, ethnicity, disability, sexual orientation, age, working patterns and grade) as well as for the Civil Service as a whole, in Appendix 1.

It is important to note that even 'highly significant' P values from the chi-square test only allow the analyst to reject an assertion that there are no statistically significant differences in the ratings reported, they do not indicate causality. Therefore, the next stage of the analysis is to interpret the findings for those cases where statistical significance has been identified.

---

[1]   French, S. (2016) *Civil Service Performance Management Diversity & Inclusion outcomes data 2014-5: an analysis.* Report for the PCS trade union.

There are two important observations about the methods and data that also need to be raised before the interpretation of the data is considered. Firstly, the data provided and the chi-square tests can only be used to identify potential relationships between two variables, the performance rating and one particular (individual) characteristic of civil service staff. This means that potential interactions between these characteristics cannot be identified. As an example, the results from the analysis by gender indicate that female staff receive a higher proportion of 'Exceed' ratings. At the same time, it is also be the case that staff working part-time receive fewer 'Exceed' ratings. While chi-square tests can identify both of these outcomes separately, they cannot assess the interaction between the two. For example, it may be the case that the various PM systems make it much harder for part-time staff to secure higher box markings and, if the majority of part-time staff are women in a department, this would mean that women working part-time hours receive disproportionately lower markings, even if women, overall, receive higher box markings. Nevertheless, while this issue must be considered, where statistically significant differences are identified, possible interaction effects *should not be* used as a reason to disregard potentially discriminatory outcomes.

Secondly, there remain limitations with the data provided. The first relates to the limited information available in some departments, and consequently across the whole civil service, in respect of specific protected groups. While the data available by grade, age, gender and working patterns ranges from between 253,642 to 249,039 civil servants, the total number of staff recorded falls to 182,783 in relation to disability, 195,359 for ethnicity and only 92,012 for sexual orientation (although it should be noted that those recorded in these latter three categories have increased compared to last year's data). In these latter categories it is also the case that some smaller departments do not have sufficient numbers of staff with (or identified with) these characteristics to provide disaggregated data (identified in the Appendix 1 and 2). The second relates to discrepancies in the reporting of data over time (see appendix 2) which makes comparative analysis problematic. This is particularly the case with the data supplied for the MOJ (which it is presumed does not include NOMS data this year), while the FCO has not been able to supply disaggregated data on PM outcomes and the DH has supplied no PMR data this year (having been able to do so last year). The third issue relates to the standardisation of categories for age and grade across departments this year. This is helpful for analysis, but the categories for age are quite distinct from those used in last year's data making comparison difficult. This point will be returned to in section three of the report.

2. Interpreting the Performance Management ratings data.

In this section of the report, the performance ratings data are examined according to each distinctive characteristic provided, identifying where the analysis has revealed statistically significant differences between staff and how these differences can be best explained.

a. gender

There were highly statistically significant differences between the performance ratings by gender in 9 out of the 16 departments and for the Civil Service overall (See Appendix 1). When interpreting these data, it would appear that the explanation for this variation in ratings relates to two factors. Firstly, in most departments (except FCO, HO, BIS and DEFRA) women were more likely to receive 'Exceed' performance ratings than men (notably in MOD,

DCMS CO, HMT and DCLG) and secondly, in all departments (except the FCO) men were more likely to receive a 'Must Improve' rating.

Table 1 highlights the differences in outcomes by gender for each Department. The columns provide an indication of how women have compared to men in relation to each performance rating in each department. A positive percentage indicates the extent to which women make up the majority of staff in the performance category, while a negative score indicates that men make up the majority of staff in a particular performance category. The table also indicates how the 'Exceed' and 'Must improve' results are not always related, depending upon differences between men and women in terms of the 'achieved' box marking.

Table 1 comparative performance of men and women by performance rating

Department	% Difference in PM outcomes: women and men		
	Exceed	Achieved	Must Improve
HMRC	2.16	2,21	-4.38
DECC	1.05	4.09	-5.14
FCO	-2.94	2.87	0.07
MOJ	0.80	1.94	-2.74
MOD	3.04	-1.12	-1.92
HO	-0.10	1.85	-1.75
DCMS	5.58	-3.32	-2.26
DWP	2.33	1.61	-3.94
DFE	2.81	0.37	-3.18
CO	3.52	-1.93	-1.58
DfT	2.50	-1.62	-0.88
HMT	3.27	2.47	-5.74
BIS	-0.39	5.36	-4.97
DCLG	4.36	0.62	-4.98
DFID	0.60	0.68	-1.28
DEFRA	-0.28	2.27	-2.00
Civil Service	1.21	1.78	-2.99

As with last year's analysis, the PMR outcomes in a significant number of departments (as well as the overall result across departments) do indicate that men were less likely than women to have received an 'Exceed', and more likely than women to have received a 'Must Improve'. While this suggests that PM systems and rating procedures have been developed that do not replicate the traditional direct and indirect discrimination against women, there needs to be careful consideration as to why men are now more likely to be rated as needing improvement and less likely to secure 'Exceed' performance markings.

b. ethnicity

There were highly statistically significant differences between the performance ratings by ethnicity in 9 out of the 15 departments where disaggregated data was provided and for the

3

Civil Service overall (see Appendix 1). When interpreting these data, it would appear that the explanation for the variation relates to two factors. Firstly, in all departments, BAME staff were less likely to receive 'Exceed' performance ratings than those staff categorised as white (notably in DECC, DCMS, DFE, CO, and DCLG). Secondly, in all departments BAME staff were also more likely to receive a 'Must Improve' rating (notably in DCLG).

Table 2 highlights the differences in outcomes by ethnicity for each Department. The columns provide an indication of how BAME staff compared to white staff in relation to each performance rating in each department. A positive percentage indicates the extent to which BAME make up the majority of staff in the performance category, while a negative score indicates that white staff make up the majority of staff in a particular performance category.

**Table 2 comparative performance of BAME and 'white' staff
by performance rating**

Department	% Difference in PM outcomes: BAME and White staff		
	Exceed	Achieved	Must Improve
HMRC	-6.38	0.49	5.89
DECC	-10.58	8.45	2.13
FCO	-	-	-
MOJ	-5.09	-1.54	6.63
MOD	-5.64	2.45	3.19
HO	-7.03	2.75	4.28
DCMS	-17.70	12.74	4.96
DWP	-4.84	2.36	2.48
DFE	-12.17	7.29	4.88
CO	-12.04	3.06	8.98
DfT	-1.36	-3.72	5.08
HMT	-8.14	6.01	2.13
BIS	-8.36	3.00	5.36
DLG	-10.36	-8.15	18.51
DFID	-3.47	-0.29	3.76
DEFRA	-7.37	3.96	3.42
Civil Service	-5.80	1.82	3.98

This set of findings is of particular concern given the extent to which BAME staff do less well under all PM systems across the civil service and with respect to the size of the differences in the distribution of performance ratings. As in last year's analysis, it can certainly be argued that these data suggest discriminatory outcomes in the operation of PM in the civil service.

c. disability

There were highly statistically significant differences between the performance ratings by disability in 9 out of the 15 departments for which there were data available, as well as for the Civil Service overall (see Appendix 1). When interpreting these data, it would appear that the explanation for the variation (again) relates to two factors. Firstly, in all departments, disabled staff were less likely to receive 'Exceed' performance ratings than those categorised as non-disabled (notably in DCMS, DFE, CO, BIS, and DFID) and secondly, in all departments

(excepting DECC) disabled staff were also more likely to receive a 'Must Improve' rating (notably in DCMS and BIS).

Table 3 highlights the differences in outcomes for each Department. The columns provide an indication of how disabled staff compared to non-disabled staff in relation to each performance rating in each department. A positive percentage indicates the extent to which disabled staff make up the majority of staff in the performance category, while a negative score indicates that non-disabled staff make up the majority of staff in a particular category.

Table 3 comparative performance of disabled and non-disabled staff by performance rating

Department	% Difference in PM outcomes: disabled and non-disabled staff		
	Exceed	Achieved	Must Improve
HMRC	-5.68	0.40	5.28
DECC	-0.02	0.51	-0.49
FCO	-	-	-
MOJ	-5.03	-0.99	6.02
MOD	-9.47	2.28	7.19
HO	-9.85	2.04	7.81
DCMS	-11.96	1.08	10.88
DWP	-6.51	1.41	5.10
DFE	-9.29	4.14	5.16
CO	-9.94	8.46	1.48
DFT	-5.99	3.04	2.95
HMT	-0.05	-4.35	4.40
BIS	-11.01	0.91	10.10
DCLG	-4.89	-3.22	8.11
DFID	-11.05	6.71	4.34
DEFRA	-3.74	-0.90	4.64
Civil Service	-6.52	0.87	5.66

This set of findings is also of particular concern given the extent to which disabled staff do less well under all PM systems across the civil service and with respect to the size of the differences in the distribution of performance ratings. Again, in line with last year's analysis, there is evidence from these data to suggest discriminatory outcomes in the operation of PM in the civil service.

d. Sexual orientation

There were no *highly* statistically significant differences between the performance ratings by sexual orientation in any of the 12 departments for which there were data available and there were no statistically significant differences for the Civil Service overall (see Appendix 1). In the two departments where there were statistically significant differences (HO and DCLG), staff identifying as Lesbian Gay or Bisexual were more likely to be awarded an 'Exceed' and in DCLG were also less likely to be awarded a 'Must improve'. The data provide little evidence to suggest discriminatory outcomes in the operation of PM in the civil service on the basis of sexual orientation, although as noted above, the relatively small numbers of staff prepared to

classify their sexual orientation means that this is an area where data limitations impact upon the analysis. As Appendix 2 highlights, in the DWP only 6.8% of staff had data recorded on sexual orientation, while it was also low in the CO (16.9%) and DFT (39.5%).

e. Age

There were highly statistically significant differences between the performance ratings by age (category) in 13 out of the 16 departments, as well as for the Civil Service overall (see Appendix 1). However, given that there are five separate age categories, identifying the nature of these statistical differences is more challenging, since there are many more potential relationships between the categories. This is exacerbated by the absence of staff in the under 25 and over 55 age categories in some departments. To address this, the results for each age category were not compared against each other category, as in the previous four sections, but rather each age category was compared against the overall distribution of the performance ratings for each department. Table A3.1 (see Appendix 3) presents these findings for each department in full. The most reliable explanation for these variations between departments appears to be the differences between the performance ratings achieved by staff aged 24 and under and those over 55.

Table 4 presents these data for staff aged 24 and under. The columns provide an indication of how staff in this age category compared to all staff in relation to each performance rating in each department. A positive percentage indicates the extent to which staff aged 24 and under constitute a higher share of staff in the specific performance rating, while a negative score indicates that this staff group constitute a lower share of staff in that particular performance category. The data highlights how staff aged 24 and under in all departments listed (except the MOJ, DCMS and CO), were less likely to receive 'Exceed' ratings, with this being particularly prevalent in the DECC, MOD, DWP and HMT. Further, in 9 departments, these staff were also less likely to receive 'Must improve' ratings. This reflects the overall civil service findings when the data are aggregated.

Table 4 comparative performance of staff by age groups 24 and under to overall departmental performance ratings

Department	Exceed	Achieved	Must Improve
HMRC	-2.86	2.52	0.33
DECC	-6.98	10.71	-3.73
MOJ	5.68	-6.90	1.22
MOD	-6.04	2.61	3.43
DCMS	3.33	-2.04	-1.29
HO	-0.77	3.32	-2.55
DWP	-7.78	2.03	5.75
DFE	-1.69	2.53	-0.84
CO	2.89	-4.26	1.37
DfT	-2.07	-2.14	4.22
HMT	-6.30	4.07	2.23
DEFRA	-1.10	-0.22	1.31
Civil Service	-3.24	1.17	2.07

Table 5 presents these data for staff aged 55 and over. The columns provide an indication of how staff in this age category compared to all staff in relation to each performance rating in each department. A positive percentage indicates the extent to which staff aged 55 and over

constitute a higher share of staff in the specific performance rating while a negative score indicates that this staff group constitutes a lower share of staff in that particular performance category. The data highlight how staff aged 55 and over in all 14 departments (for which data were available) were less likely to receive 'Exceed' ratings, with this being particularly prevalent in the DECC, DCMS, CO and BIS, while they were more likely to receive 'Must improve' ratings in all departments (except DFT).

Table 5 comparative performance of staff by age group 55+ to overall departmental performance ratings

Department	Exceed	Achieved	Must Improve
HMRC	-5.84	2.40	3.43
DECC	-11.21	4.30	6.91
MOJ	-5.87	4.95	0.92
MOD	-1.02	-0.96	1.97
DCMS	-12.81	3.43	9.38
HO	-7.33	2.65	4.68
DWP	-6.54	2.86	3.69
DFE	-9.73	5.46	4.27
CO	-10.91	3.61	7.30
DfT	-6.02	6.73	-0.71
BIS	-11.93	3.94	7.99
DCLG	-7.14	-2.17	9.31
DFID	-8.45	3.04	5.41
DEFRA	-5.74	3.90	1.84
Civil Service	-5.62	2.65	2.97

While there are greater difficulties in identifying relationships between age groups based on these data, the findings are of particular concern given the extent to which staff in the youngest and eldest age categories were less likely to secure 'Exceed' box markings and more likely to receive 'Must improve' markings. This is a distinctive set of findings, as the data from last year's PMR outcomes (where the lowest age category was 16 to 29) indicated that younger staff were more likely to receive an exceed. It appears that the re-categorisation of age bands has been helpful in identifying differences for staff aged 24 and under. Similarly, the evidence from the data also reflect that a larger proportion of older staff appear to be disadvantaged by PM, as last year's analysis evaluated staff aged 60 and over, while similar significant differences apply to those aged 55 and over this year. As noted in last year's report, while there may be interaction effects that may help explain these distributions, these results provide evidence to suggest potential discriminatory outcomes according to age in the operation of PM across departments.

f. grade

There were highly statistically significant differences between the performance ratings by grade in 8 out of 16 departments, as well as statistically significant differences for the Civil Service overall (see Appendix 1). However, as with age, given that there are four separate grading categories, identifying the nature of these statistical differences is challenging, since there are many more relationships between the categories. To address this, the results for each grade were not compared against each other, but rather each grade was compared

against the overall distribution of the performance ratings for each department. Table A2.2 (see Appendix 3) presents these findings for each department in full. The most reliable explanation for these variations between departments appears to be the differences between the performance ratings achieved by staff in AA/AO grades and those in Grades 6 to 7.

Table 6 presents these data for staff in AA/AO grades. The columns provide an indication of how staff in these grades compared to all staff in relation to each performance rating across each department. A positive percentage indicates the extent to which staff in AA/AO grades constitute a higher share of staff in the specific performance rating while a negative score indicates that this staff group constitute a lower share of staff in that category. The data highlight how staff in AA/AO grades in all departments listed (except DECC, DFE and HMT) were less likely to receive 'Exceed' ratings and more likely to have received 'Must improve' ratings. These outcomes were particularly noticeable in DCMS and DCLG.

Table 6 comparative performance of staff by AA and AO grades to overall departmental performance ratings (%)

Department	Exceed	Achieved	Must Improve
HMRC	-0.63	0.39	0.24
DECC	2.42	8.46	-10.88
FCO	-7.67	6.14	1.53
MOJ	-3.57	3.96	-0.39
MOD	-0.49	-0.09	0.58
HO	-1.31	0.33	0.98
DCMS	-18.20	-11.46	29.66
DWP	-2.59	1.75	0.84
DFE	0.10	-3.87	3.77
CO	-2.98	11.55	5.38
DfT	-2.54	1.23	1.31
HMT	1.48	-3.36	1.88
DCLG	-9.83	1.97	7.86
DFID	-2.46	0.51	1.95
DEFRA	-3.61	3.06	0.54
Civil Service	-2.35	1.81	0.54

Table 7 presents these data for civil service Grades 6 to 7. The columns provide an indication of how staff in these grades compared to all staff in relation to each performance rating across each department. A positive percentage indicates the extent to which staff in Grades 6 and 7 constitute a higher share of staff in the specific performance rating while a negative score indicates that this staff group constitutes a lower share of staff in that particular category. The data highlight staff in these senior grades in all departments (except MOD and HMT) were more likely to receive 'Exceed' ratings, with this being particularly prevalent in the MOJ, DWP and DfT while they were less likely to receive 'Must improve' ratings in all departments although to a lesser degree than in last year's PMR outturn.

Table 7 comparative performance of staff by Grades 6 and 7 to overall departmental performance ratings (%)

Department	Exceed	Achieved	Must Improve
HMRC	2.14	-1.37	-0.77
DECC	0.51	-0.35	-0.16
FCO	3.37	-2.46	-0.91
MOJ	11.54	-10.86	-0.68
MOD	-0.82	0.37	0.44
HO	1.15	0.50	-1.66
DCMS	4.15	-3.07	-1.08
DWP	6.75	-5.65	-1.10
DFE	3.23	-1.92	-1.31
CO	1.11	-3.87	-2.00
DfT	6.20	-8.09	1.89
HMT	-0.39	1.46	-1.07
BIS	1.24	-0.43	-0.81
DCLG	4.34	-1.15	-3.19
DFID	0.71	0.25	-0.97
DEFRA	4.10	-3.16	-0.95
Civil Service	4.60	-3.59	-1.01

While there are similar difficulties in identifying relationships between grades as there are with age ranges, these data do highlight how performance rating outcomes appear to favour staff in Grades 6 and 7 across most departments, while at the same time being less likely to favour staff in the lowest AA and AO grades. While there may be interaction effects that may help explain these distributions, these results on grade should continue to be a concern in relation to the operation of PM systems.

g. Working patterns

There were highly statistically significant differences between the performance ratings by full or part-time contractual status in 9 out of the 14 departments for which there were data available and for the Civil Service overall (see Appendix 1). When interpreting these data, it would appear that the explanation for the variation relates to one main factor. In all departments, staff on part-time contracts were less likely to receive 'Exceed' performance ratings than those on full-time contracts. It was also the case that in a majority of departments staff on part-time contracts staff were also more likely to receive a 'Must Improve' rating, but the differences here are less pronounced compared to last year's PMR outcomes.

Table 8 highlights the differences in outcomes by working patterns for each department. The columns provide an indication of how staff on part-time contracts compared to those on full-time contracts in relation to each performance rating in each department. A positive percentage indicates the extent to which part-time staff make up the majority of staff in the performance category, while a negative score indicates that full-time staff make up the majority of staff in that particular performance category.

Table 8 comparative performance of part-time and full-time staff by performance rating

Department	% Difference in PM outcomes: part-time and full-time		
	Exceed	Achieved	Must Improve
HMRC	-6.93	6.80	0.14
DECC	-7.47	7.21	0.26
MOJ	-8.43	9.90	-1.47
MOD	-4.94	3.09	1.85
HO	-8.24	5.51	2.73
DCMS	-8.63	6.87	1.75
DWP	-8.60	8.15	0.45
DFE	-7.16	9.75	-2.59
CO	-9.47	9.00	0.48
DfT	-8.26	7.75	0.50
HMT	-4.61	7.55	-2.94
BIS	-7.92	11.86	-3.95
DFID	-9.32	9.27	0.05
DEFRA	-7.78	8.30	-0.52
Civil Service	-8.10	7.44	0.66

These findings again raise concerns given the extent to which staff on part-time contracts are less likely to secure a higher box marking under all PM systems across the civil service. Again care should be taken when assessing this data as the simple categorisation of part-time and full-time working patterns means potential differences arising from variations in contracted hours (varieties of part-time working) cannot be identified. However, there is evidence from these data to suggest that part-time workers experience discriminatory outcomes in the operation of PM in the civil service.

3. Comparison

As noted earlier in the report, there are problems in trying to undertake comparative analysis of the 2014-5 and 2015-6 PMR outcomes. Firstly, there are those departments who have not supplied data (DH) or meaningful disaggregated data (FCO) for 2015-6 as well as significant changes to the MOJ submission. Secondly, some areas where data were missing last year, notably disability and working patterns, can now be examined, but not compared. Finally, age, and to a lesser extent grade, comparisons are problematic due to changes of data categories, in addition to the added difficulty of analysing data with multiple categories.

Table A4.1 (see Appendix 4) provides a basis for limited comparison by looking at the extent to which there have been changes in the outcomes, based upon the statistical significance testing used in both analyses. This has been focused upon three protected characteristics (gender, disability and ethnicity), where potentially discriminatory outcomes have been identified, as well as working patterns. The table identifies the extent to which there may be persistently significant, and therefore potentially discriminatory, outcomes. The key issues to emerge from this analyses can be summarised as follows:

- while there is some variation in the outcomes between the two sets of data, there are persistent highly statistically significant differences in relation to gender, ethnicity and disability in the largest departments (employing more than twenty thousand employees); namely HMRC, MOD, HO and DWP. Where data exists for the MOJ this is also the case (despite the dramatic change in the composition of the data this year). The impact of these outcomes, given the size of these departments, clearly influences the overall data for the civil service; and
- while it is difficult for the reasons outlined above to identify any distinctive trends from the chi-square tests, if anything the comparison shows that there have not any substantive improvements in the potentially discriminatory outcomes from the evidence provided by the 2015-6 data. The analysis by each characteristic in section two shows that there have not been major changes in the outcomes (in terms of identifying causality) so that this year's analysis indicates that particular groups of staff are still being disadvantaged under their respective PM systems.

4. Conclusion

The statistical analysis of the 2015-16 Performance Management ratings data across 16 Departments has shown there are statistically significant differences in performance outcomes based on gender, age, grade and working patterns and, particularly, in relation to disability and ethnicity. This report, drawing upon last year's analysis, has examined the nature of these statistical relationships to identify potential discriminatory outcomes. This would imply that the interpretation of last year's data, drawing as it did upon research findings based upon performance-related appraisal systems, is sound and that important aspects of organisational structure, job content and industrial relations in the civil service, as well as the wider political discourse, need to be carefully considered in moves to reform performance management in the civil service.

Appendix 1 – Chi-Square Statistical Test Results

Department	Gender		Ethnicity		Disability		Sexual Orientation	
	P Value	Significance	P Value	Significance	P Value	Significance	P Value	Significance
HMRC	0.00000	***	0.00000	***	0.00000	***	0.21132	
DECC	0.01090	*	0.01177	*	0.99114		0.79652	
FCO	0.27936		n.d.		n.d.		n.d.	
MOJ	0.00000	***	0.00000	***	0.00000	***	n.d.	
MOD	0.00000	***	0.00000	***	0.00000	***	0.10341	
HO	0.00003	***	0.00000	***	0.00000	***	0.00610	**
DCMS	0.29996		0.00722	**	0.03163	*	0.63141	
DWP	0.00000	***	0.00000	***	0.00000	***	0.24534	
DFE	0.00061	***	0.00000	***	0.00016	***	0.53167	
CO	0.16643		0.00211	**	0.37146		0.44960	
DfT	0.00061	***	0.00000	***	0.00000	***	0.18650	
HMT	0.00143	**	0.03269	*	0.44340		0.15482	
BIS	0.00000	***	0.00178	**	0.00000	***	n.d.	
DLG	0.00694	**	0.00000	***	0.12367		0.01744	*
DFID	0.42329		0.04344	*	0.01582	*	n.d.	
DEFRA	0.00100	**	0.00007	***	0.00000	***	0.70643	
Civil Service	0.00000	***	0.00000	***	0.00000	***	0.00510	**

This table provides the outcomes of chi-square tests used to analyse the extent to which the PMR outcomes are statistically significant for gender, ethnicity, disability and sexual orientation for each of the departments where data is available and for the aggregated Civil Service Data.

Significance of the chi-square test, measured by the P value is then assessed using standard statistical approach of indicating significance using the asterix rating system - statistically significant as P < 0.05 (*), P < 0.01 (**) and statistically highly significant as P < 0.001 (***). Cells are left blank where results are no statistically significant

Where no or insufficient data were supplied by a department to undertake a chi-square significance test, this is indicated by n.d.

Department	Working Pattern		Age		Grade	
	P Value	Significance	P Value	Significance	P Value	Significance
HMRC	0.00000	***	0.00000	***	0.00000	***
DECC	0.15892		0.00072	***	0.97274	
FCO	n.d.		n.d.		0.00000	***
MOJ	0.00000	***	0.00000	***	0.00000	***
MOD	0.00000	***	0.00000	***	0.00130	**
HO	0.00000	***	0.00000	***	0.00319	*
DCMS	0.41271		0.03752	*	0.00000	***
DWP	0.00000	***	0.00000	***	0.00000	***
DFE	0.00013	***	0.00000	***	0.01928	*
CO	0.06454		0.00000	***	0.19681	
DfT	0.00000	***	0.00000	***	0.00000	***
HMT	0.32227		0.00674	**	0.95682	
BIS	0.00023	***	0.00000	***	0.60502	
DCLG	n.d.		0.00000	***	0.00004	***
DFID	0.02204	*	0.00000	***	0.34964	
DEFRA	0.00000	***	0.00000	***	0.00002	***
Civil Service	0.00000	***	0.00000	***	0.00000	***

This table provides the outcomes of chi-square tests used to analyse the extent to which the PMR outcomes are statistically significant by working pattern, age and grade for each of the departments where data is available and for the aggregated Civil Service Data.

Significance of the chi-square test, measured by the P value is then assessed using standard statistical approach of indicating significance using the asterix rating system - statistically significant as P < 0.05 (*), P < 0.01 (**) and statistically highly significant as P < 0.001 (***). Cells are left blank where results are no statistically significant

Where no or insufficient data were supplied by a department to undertake a chi-square significance test, this is indicated by n.d.

Appendix 2 – Comparison of response data across PMR reports and by provision of data for selected protected characteristics

Department	2014-5	2015-6			
	Total responses	Total responses	% ethnicity	% disability	% LGBT
HMRC	62,689	62,883	74.4	61.0	50.2
DECC	1,400	1,374	74.2	81.8	75.7
FCO	3,930	2,569	-	-	-
MOJ	**70,436**	20,225	77.7	75.5	-
MOD	33,439	30,030	84.5	57.1	66.2
HO	20,311	21,438	91.5	92.3	83.9
DCMS	407	472	71.8	73.1	71.0
DWP	85,361	80,853	76.8	83.2	8.6
DFE	3,266	3,269	79.4	79.0	61.6
CO	1,805	1,800	33.8	23.4	16.8
DfT	1,820	12,211	70.7	77.1	39.5
HMT	1,005	1,163	95.7	95.9	74.7
BIS	3,099	2,068	54.7	83.1	57.2
DCLG	1,569	1,445	83.2	84.2	78.5
DFID	1,983	1,888	70.8	34.6	50.6
DEFRA	2086	6,539	79.6	70.7	47.5
DH	**1,767**	n.d	-	-	-
Civil Service	293,373	250,227	78.1	73.1	36.8

Appendix 3

Table A3.1 comparative performance of staff by age group to overall departmental performance ratings

Department	Age Group	Exceed	Achieved	Must Improve
HMRC	24 and under	-2.86	2.52	0.33
	25-34	3.18	-1.93	-1.25
	35-44	3.83	-2.13	-1.70
	45-54	0.88	0.06	-0.94
	55+	-5.84	2.40	3.43
DECC	24 and under	-6.98	10.71	-3.73
	25-34	4.21	-3.07	-1.14
	35-44	-0.77	1.52	-0.74
	45-54	4.03	-3.65	-0.38
	55+	-11.21	4.30	6.91
MOJ	24 and under	5.68	-6.90	1.22
	25-34	6.07	-4.88	-1.19
	35-44	1.64	-1.31	-0.33
	45-54	-1.15	0.89	0.26
	55+	-5.87	4.95	0.92
MOD	24 and under	-6.04	2.61	3.43
	25-34	2.86	-2.18	-0.67
	35-44	3.52	-2.24	-1.27
	45-54	-1.39	2.67	-1.28
	55+	-1.02	-0.96	1.97
DCMS	24 and under	3.33	-2.04	-1.29
	25-34	8.51	-6.13	-2.38
	35-44	-2.35	3.95	-1.60
	45-54	-3.08	2.00	1.08
	55+	-12.81	3.43	9.38
HO	24 and under	-0.77	3.32	-2.55
	25-34	5.96	-3.47	-2.50
	35-44	1.95	-1.30	-0.65
	45-54	-1.82	1.86	-0.05
	55+	-7.33	2.65	4.68
DWP	24 and under	-7.78	2.03	5.75
	25-34	4.88	-3.32	-1.56
	35-44	3.48	-1.59	-1.89
	45-54	1.35	-0.21	-1.15
	55+	-6.54	2.86	3.69
DFE	24 and under	-1.69	2.53	-0.84
	25-34	7.55	-4.28	-3.27
	35-44	1.47	-0.36	-1.10
	45-54	-2.65	0.64	2.01
	55+	-9.73	5.46	4.27

Department	Age Group	Exceed	Achieved	Must Improve
CO	24 and under	2.89	-4.26	1.37
	25-34	7.03	-2.66	-4.37
	35-44	-2.68	0.52	2.15
	45-54	-7.21	5.15	2.06
	55+	-10.91	3.61	7.30
DfT	24 and under	-2.07	-2.14	4.22
	25-34	5.12	-6.97	1.85
	35-44	3.44	-3.23	-0.21
	45-54	-0.59	1.48	-0.88
	55+	-6.02	6.73	-0.71
HMT	24 and under	-6.30	4.07	2.23
	25-34	4.61	-2.57	-2.03
	35-44	-3.58	1.78	1.80
	45-54	-5.60	2.81	2.79
	55+	-	-	-
BIS	24 and under	-	-	-
	25-34	15.74	-10.86	-4.88
	35-44	3.36	-0.61	-2.75
	45-54	-4.19	4.04	0.15
	55+	-11.93	3.94	7.99
DCLG	24 and under	-	-	-
	25-34	12.98	-5.91	-7.07
	35-44	2.61	2.51	-5.12
	45-54	-5.11	2.15	2.97
	55+	-7.14	-2.17	9.31
DFID	24 and under	-	-	-
	25-34	6.48	-5.84	-0.64
	35-44	1.45	0.36	-1.81
	45-54	-3.86	3.33	0.53
	55+	-8.45	3.04	5.41
DEFRA	24 and under	-1.10	-0.22	1.31
	25-34	6.94	-5.83	-1.11
	35-44	1.55	-0.13	-1.42
	45-54	-1.34	0.54	0.80
	55+	-5.74	3.90	1.84
Civil Service	24 and under	-3.25	1.15	2.10
	25-34	5.09	-3.64	-1.45
	35-44	3.14	-1.70	-1.44
	45-54	0.08	0.63	-0.72
	55+	-5.62	2.65	2.96

Table A3.2 comparative performance of staff by grade to overall departmental performance ratings

Department	Grade	Exceed	Achieved	Must Improve
HMRC	AA/AO	-0.63	0.39	0.24
	EO	-0.14	-0.28	0.42
	HEO/SEO	0.65	-0.05	-0.60
	Grade 6/7	2.14	-1.37	-0.77
DECC	AA/AO	2.42	8.46	-10.88
	EO	-2.37	0.85	1.52
	HEO/SEO	-0.25	-0.29	0.53
	Grade 6/7	0.51	-0.35	-0.16
FCO	AA/AO	-7.67	6.14	1.53
	EO	-3.21	1.91	1.31
	HEO/SEO	1.79	-1.30	-0.49
	Grade 6/7	3.37	-2.46	-0.91
MOJ	AA/AO	-3.57	3.96	-0.39
	EO	2.80	-3.47	0.67
	HEO/SEO	5.37	-6.22	0.85
	Grade 6/7	11.54	-10.86	-0.68
MOD	AA/AO	-0.49	-0.09	0.58
	EO	0.36	0.43	-0.79
	HEO/SEO	0.74	-0.20	-0.54
	Grade 6/7	-0.82	0.37	0.44
HO	AA/AO	-1.31	0.33	0.98
	EO	0.51	0.02	-0.54
	HEO/SEO	0.71	-0.62	-0.08
	Grade 6/7	1.15	0.50	-1.66
DCMS	AA/AO	-18.20	-11.46	29.66
	EO	-4.18	2.28	1.91
	HEO/SEO	-7.12	8.75	-1.63
	Grade 6/7	4.15	-3.07	-1.08
DWP	AA/AO	-2.59	1.75	0.84
	EO	1.65	-0.88	-0.77
	HEO/SEO	2.82	-2.63	-0.18
	Grade 6/7	6.75	-5.65	-1.10
DFE	AA/AO	0.10	-3.87	3.77
	EO	-1.62	0.87	0.74
	HEO/SEO	-1.84	1.49	0.34
	Grade 6/7	3.23	-1.92	-1.31

Department	Grade	Exceed	Achieved	Must Improve
CO	AA/AO	-2.98	11.55	5.38
	EO	-1.27	6.40	2.35
	HEO/SEO	-0.12	-0.93	0.16
	Grade 6/7	1.11	-3.87	-2.00
DfT	AA/AO	-2.54	1.23	1.31
	EO	-2.00	4.17	-2.17
	HEO/SEO	5.63	-5.06	-0.57
	Grade 6/7	6.20	-8.09	1.89
HMT	AA/AO	1.48	-3.36	1.88
	EO	-0.83	-0.92	1.75
	HEO/SEO	0.92	-2.58	1.66
	Grade 6/7	-0.39	1.46	-1.07
BIS	AA/AO	-0.02	-3.94	3.96
	EO	0.09	-0.29	0.20
	HEO/SEO	-1.44	1.06	0.38
	Grade 6/7	1.24	-0.43	-0.81
DCLG	AA/AO	-9.83	1.97	7.86
	EO	-1.90	-6.15	8.05
	HEO/SEO	-1.49	2.82	-1.32
	Grade 6/7	4.34	-1.15	-3.19
DFID	AA/AO	-2.46	0.51	1.95
	EO	1.96	-3.43	1.47
	HEO/SEO	-1.72	0.65	1.07
	Grade 6/7	0.71	0.25	-0.97
DEFRA	AA/AO	-3.61	3.06	0.54
	EO	-1.19	1.32	-0.13
	HEO/SEO	1.67	-1.76	0.09
	Grade 6/7	4.10	-3.16	-0.95
Civil Service	AA/AO	-2.35	1.81	0.54
	EO	0.93	-0.62	-0.31
	HEO/SEO	1.87	-1.57	-0.30
	Grade 6/7	4.60	-3.59	-1.01

Appendix 4

Table A4.1 Comparison of significance testing for selected characteristics for 2014-5 and 2015-6 PMR outcomes

Department	Gender		Ethnicity		Disability		Working patterns	
	2014-5	2015-6	2014-5	2015-6	2014-5	2015-6	2014-5	2015-6
HMRC	***	***	***	***	***	***	***	***
DECC	*	*	*	*				
FCO			**	n.d	***	n.d		n.d
MOJ	***	***	***	***	n.d	***	n.d	***
MOD	***	***	***	***	***	***	***	***
HO	***	***	***	***	***	***	***	***
DCMS	*			**	n.d	*	n.d	
DWP	***	***	***	***	***	***	***	***
DFE	**	***	***	***	***	***	**	***
CO	**		*	**	n.d			
DfT	**	***	***	***		***		***
HMT	*	**		*	n.d			
BIS	**	***	***	**	***	***	**	***
DCLG	**	**	***	***	**		*	n.d
DFID			*	*		*		*
DH		n.d	***	n.d	***	n.d	n.d	n.d
DEFRA	**	**	***	***	*	***	-	***
Civil Service	***	***	***	***	***	***	***	***

This table compares the outcomes of chi-square tests used to analyse the extent to which the PMR outcomes are statistically significant for gender, ethnicity, disability and working patterns for each of the departments where data is available and for the aggregated Civil Service data.

The significance level is indicated using the asterix rating system - statistically significant as $P < 0.05$ (*), $P < 0.01$ (**) and statistically highly significant as $P < 0.001$ (***). Cells are left blank where results are no statistically significant

Where no or insufficient data were supplied by a department to undertake a chi-square significance test, this is indicated by n.d.