



This is a peer-reviewed, post-print (final draft post-refereeing) version of the following published document:

**Millmore, Mike, Biggs, David M and Morse, Laura (2007)
Gender differences within 360-degree managerial
performance appraisals. *Women in Management Review*, 22
(7). pp. 536-551. doi:10.1108/09649420710825715**

Official URL: <http://dx.doi.org/10.1108/09649420710825715>

DOI: <http://dx.doi.org/10.1108/09649420710825715>

EPrint URI: <https://eprints.glos.ac.uk/id/eprint/1245>

Disclaimer

The University of Gloucestershire has obtained warranties from all depositors as to their title in the material deposited and as to their right to deposit such material.

The University of Gloucestershire makes no representation or warranties of commercial utility, title, or fitness for a particular purpose or any other warranty, express or implied in respect of any material deposited.

The University of Gloucestershire makes no representation that the use of the materials will not infringe any patent, copyright, trademark or other property or proprietary rights.

The University of Gloucestershire accepts no liability for any infringement of intellectual property rights in any material deposited but will remove such material from public view pending investigation in the event of an allegation of any such infringement.

PLEASE SCROLL DOWN FOR TEXT.

This is a peer-reviewed, post-print (final draft post-refereeing) version of the following published document:

Millmore, Mike and Biggs, David M. and Morse, Laura (2007). *Gender differences within 360-degree managerial performance appraisals*. *Women in Management Review*, 22 (7), 536-551.

Published in *Women in Management Review*, and available online at:

<http://www.emeraldinsight.com/doi/full/10.1108/09649420710825715>

We recommend you cite the published (post-print) version.

The URL for the published version is <http://dx.doi.org/10.1108/09649420710825715>

Disclaimer

The University of Gloucestershire has obtained warranties from all depositors as to their title in the material deposited and as to their right to deposit such material.

The University of Gloucestershire makes no representation or warranties of commercial utility, title, or fitness for a particular purpose or any other warranty, express or implied in respect of any material deposited.

The University of Gloucestershire makes no representation that the use of the materials will not infringe any patent, copyright, trademark or other property or proprietary rights.

The University of Gloucestershire accepts no liability for any infringement of intellectual property rights in any material deposited but will remove such material from public view pending investigation in the event of an allegation of any such infringement.

PLEASE SCROLL DOWN FOR TEXT.

Gender differences within 360-degree managerial performance appraisals

Mike Millmore. University of Gloucestershire, The Park, Cheltenham, GL50 2RH.

David Biggs. University of Gloucestershire, Francis Close Hall, Cheltenham, GL50 4AZ.

Laura Morse. Wiltshire Fire & Rescue Service, Manor House, Potterne, Devizes, Wiltshire, SN10 5PP.

Abstract

Purpose of this paper

The study examines the impact of gender on the performance assessments of managers arising from the 360-degree scheme operated within the UK headquarters of a large multi-international financial services organisation.

Design/methodology/approach

A questionnaire approach was used to collect data on 66 managers (33 males and 33 females) against 4 broad capabilities on the 360-degree appraisal system.

Data was gathered on each of the 66 managers from eight different sources including the individual being appraised, three of their peers, three of their direct reports (subordinates) and their manager (supervisor).

Findings

Performance ratings were either gender neutral or higher for female than for male managers. Within the case company there was no evidence of unfavourable discrimination against female managers, if anything the reverse with female managers showing superior performance compared to their male counterparts.

Research limitations/implications

As with all cross-sectional research causality cannot be confirmed and difficulties in accessing 360-degree appraisal information for a large number of managers led to constraints on research methodology.

Practical implications

The implication for human resource management (HRM) is that the 360-degree appraisal system did not necessarily fulfil the degree of objectivity claimed by its adherents and that possible adverse influence may be inherent within the 360 degree rating system of managers particularly.

Key Words

360-degree appraisal, gender bias, performance appraisal, human resource management

Word Count (Excluding references): 5,624

Introduction

Sex equality in the workplace has been an important human resource agenda item in work organisations since the mid-1970s when legislation was introduced to tackle a history of discriminatory practices, largely directed against female employees. The legal framework has largely been directed at tackling detrimental discriminatory practices that segregated women horizontally into women's work, comprising lower status roles predominantly in the secondary sector, and vertically into the lower echelons of career hierarchies in the primary sector. Thus as a result of the Sex Discrimination Act 1975 it became unlawful for employers to discriminate against women in the opportunities afforded through recruitment, promotion, transfer or training.

More than 30 years on it is debatable as to how successful the gender equality 'movement' has been. Substantial increases in female participation rates during this period have been largely in part-time jobs within the secondary labour market (Taylor, 2005) and horizontal and vertical occupational segregation as well as a gender pay gap are still evident (Willey, 2003). Irrespective of the legal framework making direct and indirect sex discrimination unlawful, both overt and covert discrimination arguably contribute to this lack of progress and continue to frustrate gender equality at work. At one level, some organisations may deliberately adopt a non-compliance approach to discriminatory legislation and produce a business case to justify their actions (Dickens, 2000). At another level, within those organisations who seek to comply with sex discrimination legislation employees may operate to subvert these good intentions. The world of work provides considerable scope, particularly for those in positions of authority, to covertly discriminate against certain individuals or groups.

Literature Review

Gender bias in performance appraisal: an overview

As identified above, the performance appraisal process is an important dimension of HRM and can play a pivotal role in determining the life chances of employees.

According to Baruch (1996), it is two main purposes. First, as a source of information used by management to make decisions about promotions, salaries, training needs and training support. Second, it is used as a feedback tool for employees, facilitating personal improvement and development. Given the pivotal role played by performance appraisal it is understandable that great stress is placed on its objectivity such that performance ratings are not skewed by unlawful discriminatory variables such as age, gender and race.

Despite this anticipated premium on objectivity, a major and recurrent criticism of performance appraisal systems is their openness to subjective, discriminatory outcomes (Brown and Heywood, 2005). Perhaps because of the statutory framework applying to discrimination at work, subjectivity is more likely to arise from the day-to-day behaviours and practises of those involved in the process of appraisal than the policies and procedures developed to guide that practise. The latter is subject to public scrutiny to a greater extent whilst the former carries with it the possibility that the appraiser will not always follow the letter let alone the spirit of the espoused organisational position. Both parties of the appraisal system may have the scope to follow their own personal agendas that reflect social attitudes and organisational cultures built up over decades (Boyden and Paddison, 1986; Ng and Chui, 2001; Willey, 2003; and Wilson, 2000). This leads to a presumption that unlawful discrimination is both possible and probable across a number of human resource practices, including performance appraisal, and will act to the detriment of certain employee groups. This contention is supported by available evidence of discrimination at work in areas such as age (Duncan and Loretto, 2004; Snape and Redman, 2003; Urwin, 2006), disability (Dibben, James and Cunningham, 2001; Woodhams and Danieli, 2000), gender (Higginbottom and Roberts, 2002; IDS, 2004; Duncan and Loretto, 2004) and race (Creegan, Colgan, Charlesworth and Robinson, 2003; Equal Opportunities Review, 2003).

With respect to gender discrimination, Aguinis and Adams (1998) argue that despite the steady increase of female managers, evident over the past three decades (Bartol, 1999), their positions in organisations involve lower-paid roles with less power and authority than their male counterparts. One barrier facing female employees wishing to overcome vertical occupational segregation is the

existence of gender-based stereotypes made up of undesirable traits (Schein, 1978). Dickerson and Taylor (2000) illustrate that these over generalisations are dangerous because despite their inaccuracy they can lead those in positions of power to limit the opportunities of females. Klonoff and Landrine's (1995) study covering a variety of occupations revealed that 40% of females studied had been denied a promotion due to their gender.

However, such findings are not consistently evident from reported research findings and the next section briefly reviews difficulties confronting research into gender bias in performance appraisal that may help to explain these inconsistencies.

Gender bias in performance appraisal: conflicting conclusions

An initial difficulty with research evidence of gender bias in performance appraisal assessments is that it is inconclusive with three general stances emerging. First, is the conclusion that given identical qualifications or performance there is a tendency to give males more favourable performance ratings than females. Nieva and Gutek (1980) argue that such outcomes are attributable to two sources. The first relates to the extent to which the attitudes and motivations of females themselves act as impediments to high achievement levels. The second concerns those external factors such as organisational structures, general stereotypes and the attitudes of others that militate against their advancement some of which find expression in discriminatory performance ratings.

Nieva and Gutek (1980) and Haefner (1977) argue that female employees who perform well may not always receive due credit because their performance is attributed to characteristics other than ability. For example, where males and females have been assessed as performing equally well in male-related tasks it is attributed to skill for males but luck for female employees (Deaux and Emswiller, 1974). However, perversely, where a female fails on a male-related task it is attributed to an internal source such as lack of ability whereas a male's failure on a female-related task is perceived as coming from a temporary external cause, such as task difficulty (Feather and Simon, 1975).

Second, is the contrary conclusion that performance ratings favour females. Studies such as those by Jacobson and Effertz (1974) and Scherer, Owen and Brodzinski (1991) found that females received more favourable performance ratings than males. Such findings may simply reflect that females were actually outperforming their male counterparts. An alternative view, known as the 'gender contrast' effect, occurs where raters, surprised that female performance ratings for traditional male roles exceed those predicated by typically held stereotypes, over-rate that performance (Heilman, Martell and Simon, 1988).

Third, is the conclusion derived from other studies that there are no gender differences in performance ratings (Mobley, 1982; Griffeth and Bedeian, 1989). One reason suggested for this is that over time many occupations, for example management, have become gender neutral rather than being viewed as stereotypically female or male-oriented thereby removing a potential source of bias (Griffeth and Bedeian, 1989).

However, although they may be applicable, those findings reported on above do not relate specifically to 360-degree appraisal systems. This leads to a second difficulty with research evidence of gender bias in performance appraisal assessments in that there is little published material relating directly to 360-degree appraisal. What is available is directed at the individual forms of appraisal that commonly make up 360-degree appraisal.

360-degree appraisal

A relatively recent development in the performance appraisal arena has been the increasing adoption by organisations of multi-rater feedback or 360-degree appraisal systems (Fletcher, 1999; McCarthy and Garavan, 2001). One of the reasons for its growing popularity is its perceived potential for generating more objective assessments of employee performance (Atwater, Waldman and Brett, 2002; Boice and Kleiner, 1997) particularly when compared to the constraints of one-way feedback that characterises organisational practice (Garavan, Morley and Flynn, 1997).

Mount and Judge (1998) characterise 360-degree feedback systems as an evaluation of an individual's performance by multiple raters from multiple levels in the organisation. "Although procedures vary, typically the individual is rated by others who interact frequently with the individual, who are knowledgeable about the individual's performance and whose opinions are valued by the individual" (Mount and Judge, 1998:557).

The majority of 360-degree systems include feedback from peers, subordinates, supervisors and self-ratings. However, depending on the nature of the business, raters from outside the organisation, such as customers, may also be included. The target manager is normally rated on various behavioural dimensions or competencies, in particular leadership and teamwork (Fletcher and Baldry, 2001).

Edwards (1996) believes that the accuracy and credibility of 360-degree performance evaluations are enhanced by expanding the sources of assessment. However, this view is not universally accepted such that irrespective of the appraisal process being used there remains the potential for bias in the constituent performance ratings. As already identified, one significant area of potential bias is that associated with discriminatory performance ratings based on gender. The study reported on here not only tests this possibility but also adds to the literature base that addresses gender bias in 360-degree appraisal.

In the case company managerial performance assessments are derived from four sources: the individual manager themselves; their peers; their subordinates and their supervisor. As pointed out above, these dimensions are generally dealt with individually in the literature and not as part of a broader remit investigating gender bias in 360-degree appraisal. We now go on to explore the scope for gender bias in and between these components.

Self-appraisal

Self-appraisal in a 360-degree appraisal process refers to the practice where the ratee rates their own performance. According to Brutus, London and Martineau (1999), self-appraisal refers to the type of feedback in which the individual is the source, the evaluator and the recipient of the information.

A significant finding from the literature is that rating errors are more common in self-appraisal in comparison to other appraisals (Fox, Caspy and Reisler, 1994). Over-raters refer to those individuals whose self-ratings are more positive than others' ratings. Those whose self-ratings are more negative than others' ratings are known as under-raters. Individuals whose self-ratings are similar or about equal to others' ratings are seen as in-agreement.

Jansen and Vloeberghs (1999) explain that over-rating by an individual is more common than under-rating, which suggests that people are generally optimistic about their competences. However, according to Lenney (1977) in the majority of work situations, males tend to over-estimate their performance and females under-estimate their performance. Deaux (1979), for example, found that male heads of department stores viewed themselves as performing significantly better than their female counterparts. Baruch (1996) suggests that the tendency of females to under-rate their performance is related to lower levels of self-esteem.

However, the inconsistency of research findings is again apparent with self-appraisal. Shore and Thornton (1986) found only a few studies where females rated themselves lower than males and some more recent studies indicate a trend towards females rating themselves higher than males (Bartol, 1999). Tsui and Gutek's (1984) study of middle managers indicated that female managers' self-ratings of performance effectiveness were significantly higher than their male colleagues. There is also a body of research that has indicated that there are no significant gender differences between self-ratings. Wohlers, Hall and London (1993) concluded in their study that male and female managers, from a variety of different organisations and jobs, did not differ in their self-ratings.

In conclusion, there is little evidence that female self-ratings are lower than those of males. The few studies available on gender effects in self-ratings suggest a tendency towards higher self-ratings for women or equality in ratings.

Peer appraisal

The involvement of subordinates and peers in 360-degree performance appraisal challenges the dominance of superiors as the most common source of performance assessments. Whilst appraisal by superiors remains the case in most schemes because of their top-down, hierarchic nature (Brutus et al, 1999), Fedor, Bettenhausen and Davis (1999) argue that with organisations increasing their emphasis on teamwork and flatter structures, supervisors are no longer the most appropriate judges of performance. May (2006) suggests such developments help explain the adoption by organisations of 360-degree appraisal systems as the basis for performance management.

Peers are viewed as a valid source of performance feedback on managers' performance because in working closely with them they have the opportunity to observe and accurately evaluate their behaviour (Brutus et al, 1999). "Also, because ratings are usually collected from a number of peers, they are more defensible from a legal standpoint, than ratings by the superior alone" (Fletcher and Baldry, 2001:124). However, May (2006) cautions that unlike traditional supervisor-led appraisals, little research has been conducted on the validity and reliability of peer ratings. This is an important consideration given that peer ratings are likely to exhibit the same types of bias attributed to supervisor and self-ratings such as gender (Drexler, Beehr and Stetz, 2001; Farh, Cannella and Bedeian, 1991; and Heslin, 2005).

It has been argued that peer ratings are more likely to be affected by pro-male bias because their performance ratings will have fewer consequences than supervisors' ratings, in that supervisors are more aware of the legal implications involved. Pro-male bias can also occur in peer ratings when their male colleagues see women as competitors. According to Hagen and Kahn (1975) peers have greater dislike for women managers when they are perceived as direct

competitors for scarce organisational rewards, such as high-level management positions. Contrary to this research, Tsui and Gutek's (1984) results showed peer ratings were higher for female managers rather than male managers, suggesting that women managers form better relationships with others or that peer competition was not an issue in the specific organisation studied, but could be present in other organisations.

Subordinate appraisals

Subordinate appraisals, or upward appraisals, are the furthest from the traditional approaches to performance appraisal because they operate counter to the traditional top-down flow of most organisations. Mount (1984) explains that from a measurement perspective subordinate ratings represent a valid source of performance assessment. Subordinates may be in a better position than other rater groups to provide meaningful assessments of certain managerial behaviours and the multi-rater basis of upward appraisal, as with peer appraisal, has the potential to remove the potential for bias inherent within single-rater appraisal systems (Atwater and Yammarino, 1993; Redman and Snape, 1992). However, this should not obscure the fact that within anonymous rating systems some subordinates have confessed to over-rating managers in order to present them in a better light (Bernardin and Beatty, 1987).

Nieva and Gutek (1980) argue that sex bias can also be present in subordinate ratings, especially if subordinates have a tendency to perceive women managers as possessing less power than their male colleagues. Osborn and Vicars' (1976) findings indicate that subordinates express a more favourable attitude toward female managers compared to male managers. This was especially true with those in higher management levels.

Supervisor appraisals

With respect to supervisor assessments, Deaux's (1979) research on multi-source feedback systems indicates that supervisors' attitudes were more favourable towards male subordinates than females. However, Tsui and Gutek (1984) found

that supervisors liked their female subordinates as much as their male subordinates. Two factors may have influenced this. Firstly, increased interaction between supervisor and subordinate may overcome any initial discomfort the supervisor had with managing a subordinate of a different gender. Secondly, the actual competence of the ratee may out-weigh any preconceptions the supervisor may have had regarding gender.

In contrast to Deaux's (1979) generalised findings, Furnham and Stringfield (2001) have more recently identified a number of significant differences in the ratings of subordinates conducted by female and male managers. They found "a significant sex of ratee effect such that male employees received lower ratings than female ratees" and that "female managers (raters) gave male reports (ratees) lower scores than they gave female reports (ratees) (Furnham and Stringfield, 2001: 284).

Inter-relationships between self, peer, subordinate and supervisor appraisals

Within the multi-rater mix of 360-degree appraisal it is conceivable that gender differences may surface when comparing the self-ratings of females and males with their ratings by others, in this case peers, subordinates and supervisors. Garavan et al (1997) illustrate that an advantage of 360-degree appraisals is that they allow individuals who are being rated to compare their self-perceptions with the perceptions of others with regard to their own behaviour. They continue to explain that 360-degree feedback is therefore based on two assumptions, that the awareness of any differences between how individuals see themselves, and how others see them will firstly, enhance self-awareness and secondly, maximise performance. "The differences in perspectives between the rater and the individual manager are not treated as an error, but as an important source of information which can enhance personal learning" (Garavan et al, 1997: 137).

London and Wohlers' (1991) research indicated that the correlation between self-ratings and subordinate ratings was higher for female than male managers. They suggested that this was because women are more likely to be concerned about interpersonal relationships and how they affect their jobs and careers than men.

Baruch (1996) produced similar results as women's self-ratings correlated closer to supervisors' ratings than men's self-ratings, indicating that either the self-evaluations of women are more accurate, or they have more of a tendency to agree with their supervisors. "The difference may be explained by lower levels of competitiveness among women, a tendency towards greater co-operation when compared to men, or higher levels of modesty in women" (Baruch, 1996: 55). London, Larsen and Thisted (1999) support Baruch's (1996) view, that compared to men, women are more likely to be sensitive to others' views of them and also more aware of their own strengths and weaknesses as seen by other people.

The research

Gender bias in performance appraisal is an area of concern for organisations promoting equal opportunities. Performance appraisal outcomes play a critical role in human resource decision-making. However, subjectivity and the scope for arbitrary discrimination in performance ratings, including on the basis of gender, has been a recurrent criticism levelled at performance appraisal processes. 360-degree appraisal has been welcomed by some because it is thought to produce more objective assessments of ratee performance (Atwater et al, 2002; Boice and Kleiner, 1997; McCarthy and Garavan, 2001). However, if it can be shown that individual differences, such as gender, can also affect 360-degree appraisal ratings, or the way they are interpreted, then it is possible that multi-rater systems if not constructed and implemented carefully may simply multiply the extent or rater error rather than reduce it (Edwards and Ewen, 1996).

In order to test out these two contradictory positions the 360-degree appraisal scheme used by the case company to assess the performance of its managers was investigated for evidence of gender bias. The investigation was based around four questions generated from themes emerging from the literature on gender bias in performance appraisal ratings and from the analytical possibilities presented by available data:

1. Are there any significant differences between the overall performance rating scores achieved by female and male managers on the 360-degree appraisal scheme?

2. Are the self-ratings of one gender significantly different to the other?
3. Do different rater groups (supervisors, peers and subordinates) rate female and male managers differently?
4. How do the respective self-ratings of female and male managers compare against those of other rater groups?

Research Design

Introduction

As has been identified in the literature review, the results of research into performance appraisal reveal a patchwork quilt of findings from which it is difficult to establish a clear pattern of gender bias. This may of course be partly explained by the factors identified earlier (laboratory Vs field studies, differences in methods of performance appraisal, different time frames). However, given that there is a strong chance that any gender discrimination is likely to be associated with prevailing organisational cultures it is arguably unreasonable to expect search for consistency and generalisability from a panoply of research findings. An alternative approach is to explore gender bias on an organisation-by-organisation basis using case study analysis to establish the reality of gender discrimination in different organisational contexts and accept the limitations of generalisation from a single case. Unfortunately there is limited reporting of contemporary case study research into gender discrimination in performance appraisals in the literature.

The study reported on here arguably goes some way to addressing identified research difficulties and hopefully provides a useful counterpoint to previous findings emerging from the literature. The research team were given access to all the raw data generated by the 360-degree appraisal process for a hierarchy of managers, both male and female, operating at different levels within the UK headquarters of a large multi-international financial services organisation (*the case company*). In consequence reported findings flow from a contemporary, organisational-based case study with the only limits to confidentiality being those inherent in the data collection methods written into the 360-degree performance appraisal system operated by the case company.

Sample

The sample comprised 66 managers (33 males, 33 females), representing 92% of the total population of 72 managers in the case company, the UK headquarters of a large multi-international financial services organisation. The participants worked in a range of different divisions or branches of the organisation. However, no notable differences between these divisions and branches were found suggesting that the sample were highly representative of all managers within the organisation.

Data collection

Data from the 360-degree appraisal system was gathered by the organisation on each of the managers. Data was collected against 4 broad capabilities comprising 15 competencies that were considered important by the organisation and included: Business Acumen; Customers and Quality; Expertise; Personal Leadership; Achievement; Developing Talent; Perspective; Creativity and Innovation; Problem-solving; Resilience; Collaboration; Motivation; Adaptability; Persuasive Communication; and Planning and Organising.

For each of the 66 managers, eight sources of data were collected via questionnaires. Typically, these questionnaires were completed by the individual being appraised, their manager (supervisor), three of their peers and three of their direct reports (subordinates). In total, 528 separate ratings were taken. Nevertheless, in order that the study could utilise parametric statistics and data from each of these four sources was considered equally, average ratings were taken with the peer and subordinate rating. Thus, statistical analysis could be completed using the gender of the managers as the independent variable.

Findings

Question 1: Are there any significant differences between the overall performance rating scores achieved by female and male managers on the 360-degree appraisal scheme?

Overall scores used were a combination of self, supervisor, peers' and subordinates' scores (See Table 1). The data was found to be normally distributed using a Kolmogorov-Smirnov Test ($KS=0.53$, $p=n/s$) and as such parametric tests were used on the data. This data illustrated that females had an overall higher score ($M=137.3$) than males ($M=130.0$). This difference was found to be statistically significant using a t-test ($t=2.1$, $p<0.05$). This result indicates that females were rated significantly higher than males in the overall scores achieved on the 360-degree appraisal system.

INSERT TABLE 1 HERE

Question 2: Are the self-ratings of one gender significantly different to the other?

This research question was posed to test whether gender affected self-ratings. Indeed, self-appraisal ratings were higher in females ($M=146.2$) than males ($M=133.7$) but this difference was not statistically significant using an independent samples t-test ($t=1.9$, $p=n/s$) suggesting no difference between the self-rating scores and gender.

Question 3: Do different rater groups (supervisors, peers and subordinates) rate female and male managers differently?

Supervisors, peers and subordinates ratings on the 360 degree instrument were all assessed in terms of gender differences. Although most ratings were slightly different none of these differences was statistically significant with the exception of subordinate ratings. Subordinate ratings demonstrated a difference between males ($M=138.7$) and females ($M=153.3$). An independent t-test demonstrated that this difference was statistically significant ($t=2.0$, $p<0.05$). This suggests that

only subordinate ratings of managers were significantly different in terms of gender and that supervisor and peer differences were not statistically different.

Question 4: How do the respective self-ratings of female and male managers compare against those of other rater groups?

This question was posed in order to examine whether female or male managers rated themselves higher or lower in comparison to the ratings provided by other raters. Table 1 demonstrates that both female and male managers rate themselves higher than their supervisors and peers, but lower than their subordinates. Differences between all of these scores were tested using an ANOVA and found to be statistically significant in males ($F=4.6$, $p<0.01$) and females ($F=7.8$, $p<0.01$) indicating a significant difference between self-ratings and the other rating groups. Post-hoc analysis was performed to ascertain differences in specific scores utilising Tukey tests. These results indicate that self-ratings for male scores were significantly different to peer-ratings ($F=4.6$, Tukey HSD=18.2, $p<0.05$) but were not significantly different to ratings by supervisors and subordinates. The self-ratings for female scores differed significantly from the ratings provided by supervisors ($F=7.8$, Tukey HSD=22.5, $p<0.05$) and peers ($F=7.8$, Tukey HSD=20.5, $p<0.05$) but were not significantly different to subordinate ratings.

Discussion

It has been demonstrated that the overall performance rating produced through the 360-degree appraisal system was higher for female than for male managers (question 1). This is at odds with the traditional view of sex discrimination, where females are subjected to negative, sex-based stereotypes, as found by Deaux and Emswiller (1974) and Haefner (1977). Taken at face value this may reflect that female managers were consistently out performing their male counterparts on the work-related competencies assessed and that this has been objectively identified through the 360-degree appraisal process. This is consistent with the effort being made by management professions to ensure increased equality and objectivity in employee performance evaluations (Fletcher, 1999). However, unless reverse discrimination is in operation, this superior performance by female managers requires explanation.

One possible explanation relates back to the suggestion that female managers have something more to prove than male managers and, as a result, may work harder or differently to achieve results (Greener, 2007). If this is the case then at least their efforts are being recognised through the 360-degree appraisal system in the case company (Fletcher, 1999). An alternative explanation is provided by the 'gender contrast' effect (Heilman et al, 1988) where the over-rating of females performing roles perceived as male-oriented, such as management, is evidenced. The argument here is that raters are surprised by the actual performance of females in comparison to sex-based stereotypes, particularly, given the perceived barriers that they need to surmount if they are to succeed, leading to over-compensation in any ratings of performance.

It may also be that the espoused strengths of the 360-degree appraisal process as an objective approach to performance appraisal are being realised within the case company. The more accurate the information supplied to raters the less likelihood there is that sex-based stereotypes will affect the performance ratings given (Fletcher, 1999; Nieva and Gutek, 1980). The very nature of 360-degree appraisal's multi-rater structure is designed to eliminate bias by ensuring that the majority of managerial activity is exposed to scrutiny so that a holistic picture of

their performance emerges (Atwater and Yammarino, 1993; Brutus et al, 1999; Fedor et al, 1999; Fletcher and Baldry, 2001).

The findings reported against question two (self-ratings) are consistent with more recent research that demonstrate little or no difference in the self-ratings of females and males across a variety of organisations (Bartol, 1999; Skipper and Bell, 2006). This may suggest that female managers today are more confident and possess higher self-esteem compared with 20-30 years ago (Aycan, 2004). It is also an outcome to be hoped for from the equality-centred and positive action agenda flowing from the sex discrimination legislation introduced in the 1970s.

The third question concerned the extent to which gender differences in the ratings of female and male managers emerged from the three stakeholder groups: supervisors; peers; and subordinates. No significant differences were found in supervisors' ratings of female and/or male managers' performance, echoing the work of Tsui and Gutek (1984). This could be explained by the fact that the competence level of ratee managers outweighs the affects of any potential preconceptions of gender held by their rater supervisors or that favourable relationships between rater and ratee inhibit stereotyping based on gender (Tsui and Gutek, 1984).

Similarly, no significant differences were identified between the ratings given to female and male managers by their peers. Given that competition between peers can be a cause of gender bias in peer-ratings (Hagen and Kahn, 1975) it points either to the culture in the case company being relatively uncompetitive or that competitive tendencies are more constructively channelled through the ready availability of development opportunities. As discussed with supervisor-ratings, it could also be a function of good relationships between ratees and their peers (Tsui and Gutek, 1984).

In contrast, the findings for subordinate-ratings support Osborn and Vickers' (1976) research in that their subordinates rated female managers significantly higher than male managers. As with the significantly higher overall performance scores achieved by female managers this may be a true reflection of their

competence rather than some kind of reverse discrimination. Alternatively, it may be that gender differences in managerial behaviour lead subordinates to feel more comfortable and form closer relationships with female managers who exhibit more concern for development through interpersonal relationships than with males who seek advancement through independence and self-sufficiency (London and Wohlers, 1991).

In relation to question four, and consistent with the findings of Jansen and Vloeberghs (1999), when compared to other rater groups there was a tendency for both female and male managers to 'over-rate' themselves. Female managers were different to supervisors and peers whereas males were only different to peers. Male self-ratings were generally in agreement with those made by their superiors and subordinates but evidenced over-rating in comparison with peer assessments. One explanation for this is that the nature of the working relationship means that managers generally are more likely to be aware of the perceptions of them held by their superiors and subordinates than their peers. This reflects the fact that discussions around their targets and performance are more likely to be held with superiors and subordinates with peer contact being on a different footing.

For female managers, not only did they rate their performance higher than their peers but their self-ratings were also significantly higher than their supervisor ratings. These findings contrast with those of London and Wohlers (1991) who found females' self-evaluations correlated more closely to supervisor-ratings than was the case with male self-evaluations. The findings from this, more recent, study therefore challenge the notion that female managers are likely to be more co-operative with others and have higher levels of modesty (Baruch, 1996). Rather this greater tendency to over-rate may reflect greater confidence in evaluating their own performance and increasing levels of self-esteem.

Conclusions

Overall in the case company there is no evidence of any statistically significant gender bias against female managers as exhibited by rating scores derived from a 360-degree appraisal system based on four rating sources - managers' self-ratings and the ratings of their supervisors, peers and subordinates. Indeed it was found that when the ratings of all these groups were aggregated, the performance of female managers was rated significantly higher than the performance of their male counterparts. This can be attributed to female managers receiving higher rating scores from their own self-appraisal and from the ratings of their peers and subordinates. Only managers' supervisors rated male managers higher than female managers but their difference in performance ratings was not found to be statistically significant.

These findings arguably have significance for the case company. They are contrary to the view that the glass ceiling remains evident in organisations other than those in sectors where female employees predominate, such as education and health. The findings would also seem to support those who argue that 360-degree appraisal systems are more accurate and credible because multiple ratings iron out the greater potential for bias inherent in superior-subordinate performance appraisal systems.

Whilst this might suggest that the case company operates effective, gender-neutral performance appraisal systems this cannot necessarily be attributed to the operation of 360-degree appraisal. Clearly as a cross-sectional, rather than longitudinal, study we have no comparative data on the objectivity or otherwise of the appraisal system operating prior to the introduction of 360-degree appraisal. It may well have been that the previous system produced comparable performance assessments for female managers relative to their male counterparts. It should also be noted that the disaggregated stakeholder ratings shown in Table 1 show a lack of gender bias, at least against female managers, across the rating groups. This might suggest that for the case company other performance appraisal systems, if operated, would have been similarly free from gender bias. However, on the basis of available evidence it can at least be argued that the current 360-

degree appraisal system is not discriminating against female managers to their detriment. Perhaps what is more significant is the possibility that the higher performance ratings of female managers might be attributable to reverse discrimination!

From the data available it has not been possible to establish the reasons lying behind the higher performance ratings received by female managers. A particular difficulty arising from the confidential nature of the performance appraisal process is that apart from the managers themselves, the gender of raters (superiors, peers and subordinates) was not collected and therefore analysis of rater and ratee gender comparisons was not possible. For the sake of argument it could be that it was the over-rating of female managers by their female superiors, peers and subordinates that produced their higher performance scores. Equally, as argued earlier, it may have been that for the case organisation its female managers were on average more competent than their male colleagues.

Unfortunately therefore, although supporting the gender-neutral position, the case study findings do little to resolve the inconsistent evidence on gender bias in performance appraisal reported in the literature and only serve to demonstrate that the relative paucity of research in this area needs to be addressed.

References

- Aguinis, H. and Adams, S. (1998) Social-role versus structural models of gender and influence use in organizations: a strong inference approach, *Group & Organization Management*, Vol. 23, No. 4, pp. 414-447.
- Atwater, L. and Yammarino, F. (1993) Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology*, Vol. 45, No. 1, pp. 141-164.
- Atwater, L., Waldman, D. and Brett, J. (2002) Understanding and optimising multi-source feedback, *Human Resource Management*, Vol. 41, No. 2, pp. 193-208.
- Aycan, Z. (2004) Key Success Factors for Women in Management in Turkey. *Applied Psychology: An International Review*, Vol. 53, No. 3, pp. 453-477.
- Bartol, K. (1999) Gender influences on performance evaluations, in Powell, G. (ed) *Handbook of Gender & Work*, London, Sage Publications.
- Baruch, Y. (1996) Self performance appraisal vs direct-manager appraisal: a case of congruence, *Journal of Managerial Psychology*, Vol. 11, No. 6, pp. 50-65.
- Bernardin, H. and Beatty, R. (1987) Can subordinate appraisals enhance managerial productivity?, *Sloan Management Review*, Vol. 28, No. 4, pp. 63-73.
- Boice, D. and Kleiner, B. (1997) Designing effective performance appraisal systems, *Work Study*, Vol. 46, No. 6, pp. 197-201.
- Boyden, T. and Paddison, L. (1986) Banking on Equal Opportunities, *Personnel Management Journal*, September, pp. 42-46.
- Brown, M. and Heywood, J. (2005) Performance appraisal systems determinants and change, *British Journal of Industrial Relations*, Vol. 43, No. 4, pp. 659-679.
- Brutus, S., London, M. and Martineau, J. (1999) The impact of 360-degree feedback on planning for career development, *The Journal of Management Development*, Vol. 18, No. 8, pp. 676-693.
- Creegan, C., Colgan, F., Charlesworth, R. and Robinson, G. (2003) Race equality policies at work: employee perceptions of the "implementation gap" in a UK local authority, *Work, Employment and Society*, Vol. 17, No. 4, pp. 617-640.
- Deaux, K. (1979) Self-evaluations of male and female managers, *Sex Roles*, Vol. 5, No. 5, pp. 571-580.
- Deaux, K. and Emswiller, T. (1974) Explanations of successful performance on sex-linked tasks: what is skill for the male is luck for the female, *Journal of Personality & Social Psychology*, Vol. 29, No. 1, pp. 80-85.
- Dibben, P., James, P. and Cunningham, I. (2001) Senior management commitment to disability, *Personnel Review*, Vol. 30, No. 4, pp. 9-19.
- Dickens, L. (2000) Still wasting resources? Equality in employment, in Bach, S. and Sisson, K. (eds) *Personnel Management: A Comprehensive Guide to Theory and Practice*, 3rd edition, Oxford, Blackwell, pp. 137-169.

- Dickerson, A. and Taylor, M. (2000) Self-limiting behaviour in women, *Group & Organization Management*, Vol. 25, No. 2, pp. 191-211.
- Drexler, J.A., Beehr, T.A. and Stetz, T.A. (2001) Peer appraisals: differentiation of individual performance on group tasks, *Human Resource Management*, 40, pp. 333-345.
- Duncan, C. and Loretto, W. (2004) Never the right age? Gender and age-based discrimination in employment, *Gender, Work and Organisations*, Vol. 11, No. 1, pp. 95-115.
- Edwards, M. (1996) Improving performance with 360-degree feedback, *Career Development International*, Vol. 1, No. 3, pp. 5-8.
- Edwards, M. and Ewen, A. (1996) *360 Degree Feedback*, New York, Amacom, American Management Association.
- Equal Opportunities Review (2003) Economic activity rates of disabled people and ethnic minorities, *Equal Opportunities Review*, No. 121, September, pp. 7-8.
- Farh, J., Cannella, A.A. and Bedeian, A.G. (1991) Peer ratings: the impact of purpose on rating quality and user acceptance *Group and Organizational Studies*, Vol.16, pp. 367-386.
- Feather, N. and Simon, J. (1975) Reactions to male and female success and failure in sex-linked occupations: impressions of personality, causal attributions, and perceived likelihood of different consequences, *Journal of Personality and Social Psychology*, Vol. 31, No. 1, pp. 20-31.
- Fedor, D., Bettenhausen, K. and Davis, W. (1999) Employees' dual roles as raters and recipients, *Group & Organization Management*, Vol. 24, No. 1, pp. 92-120.
- Fletcher, C. (1999) The implication of research on gender differences in self-assessment and 360 degree appraisal. *Human Resource Management Journal*, Vol. 9 No. 1, pp. 39-46.
- Fletcher, C. and Baldry, C. (2001) Multi-source feedback systems: a research perspective, in Robertson, I. and Cooper, C. (eds) *Key Issues in Industrial & Organizational Psychology: Personnel Psychology & HRM*, West Sussex, John Wiley & Sons Ltd.
- Fox, S., Caspy, T. and Reisler, A. (1994) Variables affecting leniency, halo and validity of self-appraisal, *Journal of Occupational and Organizational Psychology*, Vol. 67, No. 1, pp. 45-56.
- Furnham, A. and Stringfield, P. (2001) Gender differences in rating reports: female managers are harsher raters, particularly of males, *Journal of Managerial Psychology*, Vol. 16, No. 4, pp. 281-288.
- Garavan, T., Morley, M. and Flynn, M. (1997) 360 degree feedback: its role in employee development, *Journal of Management Development*, Vol. 16, No. 2, pp. 134-147.
- Greener, I. (2007) The Politics of Gender in the NHS: Impression Management and 'Getting Things Done. *Gender, Work & Organization*, Vol. 14, No. 3, pp. 281-299
- Griffeth, R. and Bedeian, A. (1989) 'Employee performance evaluations: effects of ratee age, rater age and ratee gender, *Journal of Organizational Behaviour*, Vol. 10, No. 1, pp. 83-90.
- Haefner, J. (1977) Sources of discrimination among employees: a survey investigation, *Journal of Applied Psychology*, Vol. 62, No. 3, pp. 265-270.
- Hagen, R. and Kahn, A. (1975) Discrimination against competent women, *Journal of Applied Social Psychology*, Vol. 5, No. 3, pp. 362-376.

- Heilman, M., Martell, R. and Simon, M. (1988) The vagaries of sex bias: conditions regulating the undervaluation, equiavaluation and overvaluation of female job applicants, *Organizational Behaviour and Human Decision Processes*, Vol. 41, No. 1, pp. 98-110.
- Heslin, P.A. (2005) Conceptualizing and evaluating career success *Journal of Organizational Behaviour*, Vol. 26, No. 2, pp. 113-127.
- Higginbottom, K. and Roberts, Z. (2002) EOC urges local authorities to confront lack of senior women, *People Management*, Vol. 8, No. 12, 13 June, p.7.
- Jacobson, M. and Effertz, J. (1974) Sex roles and leadership: perceptions of the leaders and the led, *Organizational Behaviour & Human Performance*, Vol. 12, No. 3, pp. 383-397.
- Jansen, P. and Vloeberghs, D. (1999) Multi-rater feedback methods: personal and organizational implications, *Journal of Managerial Psychology*, Vol. 14, No. 6, pp. 455-476.
- Klonoff, E. and Landrine, H. (1995) The schedule of sexist events: a measure of lifetime & recent sexist discrimination in women's lives, *Psychology of Women Quarterly*, Vol. 19, No. 4, pp. 439-472.
- Lenney, E. (1977) Women's self-confidence in achievement settings, *Psychological Bulletin*, Vol. 84, No. 1, pp. 1-13.
- London, M., Larsen, H. and Thisted, L. (1999) Relationships between feedback and self-development, *Group and Organization Management*, Vol. 24, No. 1, pp. 5-28.
- London, M. and Wohlers, A. (1991) Agreement between subordinate and self-ratings in upward feedback, *Personnel Psychology*, Vol. 44, No. 2, pp. 375-390.
- May, G.L. (2006) The effect of social style on peer evaluation ratings in project teams, *Journal of Business Communications*, Vol. 43, No. 1, pp. 4-20.
- McCarthy, A. and Garavan, T. (2001) 360° feedback process: performance, improvement and employee career development, *Journal of European Industrial Training*, Vol. 25, No. 1, pp. 5-32.
- Mobley, W. (1982) Supervisor and employee race and sex effects on performance appraisals: a field study of adverse impact and generalizability, *Academy of Management Journal*, Vol. 25, No. 3, pp. 598-606.
- Mount, M. (1984) Psychometric properties of subordinate ratings of managerial performance, *Personnel Psychology*, Vol. 37, No. 4, pp. 687-701.
- Mount, M. and Judge, T. (1998) Trait, rater, and level effects in 360-degree performance ratings, *Personnel Psychology*, Vol. 51, No. 3, pp. 557-577.
- Ng, C. and Chiu, W. (2001) Managing Equal Opportunities for women: sorting the friends from the foes, *Human Resource Management Journal*, Vol. 11, No.1, pp.75-88.
- Nieva, V. and Gutek, B. (1980) Sex effects on evaluation, *Academy of Management Review*, Vol. 5, No. 2, pp. 267-276.
- Osborn, R. and Vicars, W. (1976) Sex stereotypes: an artefact in leader behaviour and subordinate satisfaction analysis, *Academy of Management Journal*, Vol. 19, No. 3, pp. 439-449.
- Redman, T. and Snape, E. (1992) Upward and onward: can staff appraise their managers? *Personnel Review*, Vol. 21, No. 7, pp. 32-46.

- Schein, V. (1978) Sex role stereotyping, ability and performance: prior research and new directions, *Personnel Psychology*, Vol. 31, No. 2, pp. 259-268.
- Scherer, R., Owen, C. and Brodzinski, J. (1991) Rater and ratee sex effects on performance evaluations in a field setting, *Management Communication Quarterly*, Vol. 5, No. 2, pp. 174-192.
- Shore, L. and Thornton, G. (1986) Effects of gender on self- and supervisory ratings, *Academy of Management Journal*, Vol. 29, No. 1, pp. 115-129.
- Skipper, C.O. and Bell, L.C. (2006) Assessment with 360° Evaluations of Leadership Behavior in Construction Project Managers. *Journal of Management in Engineering*, Vol. 22 No. 2, pp 75-80.
- Snape, E. and Redman, T. (2003) Too old or too young? The impact of perceived age discrimination, *Human Resource Management Journal*, Vol. 13, No. 1, pp. 78-89.
- Taylor, S. (2005) *People Resourcing*, 3rd edn. London, CIPD.
- Tsui, A. and Gutek, B. (1984) A Role set analysis of gender differences in performance, affective relationships, and career success of industrial middle managers, *Academy of Management Journal*, Vol. 27, No. 3, pp. 619-635.
- Urwin, P. (2006) Age discrimination, legislation and human capital accumulation, *Employee Relations*, Vol. 28, No. 1, pp. 87-97.
- Willey, B. (2003) *Employment Law in Context*, Harlow: Financial Times, Prentice Hall.
- [Wilson, E. (2000) Inclusion and exclusion and ambiguity – the role of organizational culture, *Personnel Review*, Vol. 29, No.3.
- Wohlers, A., Hall, M. and London, M. (1993) Subordinate rating managers: organizational and demographic correlates of self/subordinate agreement, *Journal of Occupational and Organizational Psychology*, Vol. 66, No. 3, pp. 263-275.
- [Woodhams, C. and Danieli, A. (2000) Disability and diversity – a difference to far?, *Personnel Review*, Vol.29, No. 3.

Table 1: Gender differences in appraisal rating scores

Rating Group	Male Mean Scores	Female Mean Scores
Self	133.7	146.2
Supervisors	132.2	123.9
Peers	115.5	125.6
Subordinates	138.7	153.3
Overall Scores	130.0	137.3