Performance appraisal: Stressful for some.

Dr. Geoff Carter

School of Management, Griffith University, Gold Coast campus, Australia

Email: g.carter@griffith.edu.au

and

A/Prof Brian Delahaye, School of Learning and Professional Studies, Queensland University of Technology, Brisbane, Australia.

Email: b.delahaye@qut.edu.au
ABSTRACT: Performance appraisal has been generally accepted as being stressful for both appraisers and appraisees. Using pre- and post-measures of the three catecholamine neurotransmitters, namely adrenaline, noradrenaline and cortisol, this paper examined the extent that a performance appraisal interview is stressful for appraisees. It was found that the event was highly stressful for unsuccessful appraisees, both male and female and, further, that unsuccessful appraisees expected a stressful event. Gender differences were found in the study. Males appear to be more physiological responsive while for females, the psychological, rather than the physiological costs may be higher. Finally, the paper explores a number of issues for managers resulting from these findings.

Key words: performance appraisal, stress, gender differences

Performance appraisal is the mechanism by which organisations seek to assess and develop staff, thereby enhancing performance and distributing rewards (Fletcher, 2001). Several writers (Delahaye, 2005; Fisher, Schoenfeldt and Shaw, 1999; Lee, 1996) have suggested that there are two basic purposes for performance appraisals in organisations – developmental and administrative. Administrative reasons for performance appraisal include making decisions for salary increments, promotions and retrenchments and these appraisals are predominantly retrospective in nature. Developmental performance appraisals predominantly concentrate on identifying the developmental needs of an individual for successful present and future performance (Delahaye, 2005). This paper examines the extent to which the administrative performance appraisal is stressful to the appraisees, whether there are any gender differences and the implications of the findings for managers.

PERFORMANCE APPRAISAL IS STRESSFUL

As Nesbit and Wood (2002) point out, as pivotal as performance appraisal is to motivation, organisations face substantial problems in effectively implementing performance appraisal systems and there is widespread dissatisfaction with their use. Pettijohn, Pettijohn and Taylor (2000) report that employees and supervisors alike often perceive the performance appraisal with fear and loathing. Many managers are uncomfortable with the appraisal process and employees dislike receiving the appraisals (Milliman, Nason, Zhu & De Cieri, 2002). Adding to this pressure is the belief held by many appraisees that their future may be under threat. This is a reasonable concern given that one of the roles of performance appraisal is of an administrative nature where salary increases and/or
While there are many inherent suspicions, as indicated in the literature already described, there appear to be no studies indicating that there is a direct link between performance appraisal and stress. Such suppositions do seem reasonable in the light of several studies on the effect of work upon employees. In particular, links have been found between stressors and work demands and mood (see, for example, Lundberg, Hasson and Magnusson 1989). The assumption that performance appraisal is a source of stress does appear reasonable. Therefore the overarching research question for this project was: To what extent is performance appraisal a stressful event to appraisees?

This overarching research question can be examined further. Tattersall and Morgan (1997) report that individuals may interact differently to a stressful event such as performance appraisal. In particular, a number of writers (for example, Frankenhaeuser, Lundberg & Forsman, 1980; Rauste-von Wright, von Wright & Frankenhaeuser, 1981) suggest that males and females may experience stress differently. Further, a consensus of opinion exists that males have a more psychophysiological responsive body than females and hence that the physiological cost of stress is higher for males than for females (see, for example, Fibiger, Singer & Miller, 1984; Rauste-von Wright et al 1981).

On the other hand, however, several extensive reviews have supported the notion that males receive better reviews than females do (Dobbins, Cardy & Truxillo, 1989; Kalin & Hodgins, 1984; Neiva & Gutek, 1980; Ruble & Ruble, 1982). Specifically, it was found that raters with traditional stereotypes of women evaluate the performance of female ratees less favourably and that these behaviours occurred when the purpose of the appraisal was administrative, that is, when dealing with pay, promotions, transfers and retrenchments (Dobbins et al, 1989). Thus it could be argued that, for females, the performance appraisal is a more stressful event than for males. For the purposes of this study, though, it should be noted that the literature does not give unequivocal support for such gender
differences under stressful circumstances. For example, Bergman and Magnusson (1979) believe that physiological cost may be higher for males but that the psychological costs to females may be higher.

**MEASURING STRESS**

Cannon (1936) originally defined stress as an “emergency response” designed to save humans via the fight or flight mechanism. Selye (1976) heralded a new direction in stress research by positing that, when a stressor is applied to a human, she or he will emit global psychophysical reactions. Mason (1975) found that reaction to uncertainty, ambiguity or fear is characterised by the release of catecholamine neurotransmitters, namely adrenaline, noradrenaline and cortisol. Several other researchers (Henry 1993; Henry & Stephens, 1979; Sourkes, 1985) have noted three consistent indicators of stress - adrenaline, noradrenaline and cortisol. These catecholamines are rapidly metabolised and their outputs found in urine.

Most of the physiological literature suggests that adrenaline is the marker that best reacts to psychological state (for example, Rauste-von Wright et al 1981). Adrenaline activation occurs as part of the flight or fight mechanism in that, once danger or anger is recognised, secretion of the hormone adrenaline follows almost immediately. Goodman and Gillman (1975) list the main effects as an increase in blood pressure, heart rate and gastric secretions as well as constriction of blood vessels in the skin, mucosa and kidneys and an increase in blood flow to the skeletal muscles. All the above effects are designed to equip the body with immediate power to resist, flee or attack.

Noradrenaline constricts salivary glands, decreases heart rate, reduces blood supply to the skin and reduces blood volume thereby making the heart work harder (Goodman & Gillman 1970). Frankenhaeuser (1979) reports that, while adrenaline appears to be most responsive to mental stress, noradrenaline responds more to physical exercise. However, there is considerable overlap between adrenaline and noradrenaline and several studies report that noradrenaline is produced as a response to stress (see, for example, Meijman, Frings-Dresen, Kuiper & Kuiper, 1993). Further, Lundberg and
Frankenhauser (1980) found that noradrenaline was associated with perceived effort to succeed and self-imposed stress – behaviours that are often associated with performance appraisal.

Just as adrenaline and noradrenaline work to give the body power to resist noxious and potentially dangerous stimuli, the neurotransmitter, cortisol works to repair the damage done when such agents are perceived to have damaged or traumatised the body in some way (Goodman and Gillman 1975). Researchers, such as in Fibiger and Singer (1984) and Lundberg et al (1989), have found high correlations between stress and cortisol during stressful events, for example on high stress days at work and during public speaking engagements. Cortisol is associated with increasing positive mood and affect (Henry, 1993). However, too little cortisol is associated with irritation, tenseness and tiredness. Thus, a balance of cortisol is necessary to ensure health.

Overall, then, the three neurotransmitters of adrenaline, noradrenaline and cortisol have been extensively labelled as the classic stress hormones. Further, these neurotransmitters can be detected in both blood and urine samples. For the purposes of this study, it is important to note that these three neurotransmitters are liberated by an individual’s perception of stressors (Spielberger and Sarason, 1975), as well as through physical trauma.

**RESEARCH METHODOLOGY**

Ethics approval for this research was given by the supervising university. The participants were fifty employees (male=28, female=22) of several large companies from Queensland, Australia. The participants' ages ranged from 21-52 years with a mean age of 35.2 years. Employment locations varied from golf courses to building construction to real estate sales and furniture construction. The administrative performance appraisals were undertaken in the workplace by the appropriate supervisor and were part of the official performance management process of the organisation concerned. Almost all appraisees were office workers not engaged in manual labour. Participants had an annual average salary of $ AUD 25,459.00. The project was discussed with all participants before
the performance appraisal when the appraisees were assured that the urine samples would be used only for the project and not for other purposes, for example drug testing. Urine samples were taken before the interview, appraisees were to drink 375 mls of mineral water during the interview and provided another urine sample after the interview. The samples were assayed at an official pathology laboratory. Participants were also contacted the next day and no complaints were received.

As mentioned earlier, compared to the developmental appraisal, the administrative performance appraisal is most probably the most stressful. The critical issue in the administrative performance appraisal is, of course, whether the result for the appraisee can be defined as successful. In this study, success was defined as obtaining a pay rise as a result of the performance appraisal and failure (non-success) occurred when the employee did not receive a pay rise. For the study, it was important that the appraisees experienced a similar appraisal event. Performance reviews have a generally agreed on process (Delahaye, 2005; Schuler, Dowling, Smart & Huber, 1992) and this process was adhered to in this project. In particular, the interviews in this project were based on the work of Deblieux (2003), Grote (2002) and Tharenou (1991) and used the following steps - explain the purpose and benefits of the appraisal system; outline the structure of the meeting and the Performance Appraisal Checklist (the checklist covered three areas – job knowledge, job abilities and job compensation, with each area being represented by 10 items, each measured on a five point Likert scale); negotiate the main job responsibilities; identify the performance standards; negotiate the goals for each responsibility; agree on an action plan for each goal; jointly complete the Checklist; review the total workload and skill requirements; outline the method of review and the outcomes.

RESULTS

Results presented in this section focus on the effects of gender, performance, and time of measurement on neurotransmitter activation levels. A multivariate analysis of variance (MANOVA) was used to calculate the effects of the independent variables (gender, performance and time) on the dependent variables (rate of neurotransmitter excretion, i.e. adrenaline, noradrenaline and cortisol).
A mixed factorial analysis of variance design was utilised, with two between-group variables, gender and success, and one within-group variable, time. The four independent groups; successful males, unsuccessful males, successful females, and unsuccessful females had sample sizes of 10, 18, 18, and 4 respectively.

The Catecholamine changes

This section will describe the results from each of the catecholamine measures – adrenaline, noradrenaline and cortisol - of the groups.

Adrenaline.

The MANOVA results from the adrenaline measure indicated significant main effects for each of the independent variables; gender, $F(1,46) = 52.25$, $p<0.05$, success, $F(1,46) = 6.75$, $p<0.05$, and time, $F(1,46) = 55.07$, $p<0.05$. A significant three-way interaction between the variables was also found, $F(1,46) = 23.86$, $p<0.05$. This means that, over the time of the performance appraisal interview, the adrenaline levels changed significantly for both the gender group and the successful/unsuccessful group. The data was then analysed for each of the four sub-groups – that is, successful females, unsuccessful females, successful males and unsuccessful males - over the time of the interview (ie, between the pre- and post-tests). Table 1 shows the interaction of the means for each group over time.

<table>
<thead>
<tr>
<th>Table 1: Means and Standard Deviations for Adrenaline Across Gender and Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Adrenaline Pre-test</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Adrenaline Post-test</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Note. Significant differences Pre and Post are marked with a different superscript.
A significant change was found for unsuccessful (low performing) males ($F(1,46) = 80.76, p<0.05$) with the means showing an increase of 10.88. There was also a significant change for unsuccessful (low performing) females ($F(1,46) = 52.07, p<0.05$) with the means an increase of 20.25. Therefore, both male and female low performing (unsuccessful) groups experienced statistically significant increases in adrenaline levels, indicating an increase in stress.

The surprising finding was from the two successful (high performing) groups. There was a significant change for successful (high performing) females ($F(1,46) = 28.44, p<0.05$). However, the mean for the successful females decreased from 87.06 to 80.00 – a decrease of 7.06. These females began the interview feeling anxious and become calmer during the course of the appraisal, indicating that they found the interview less stressful than did any other group. The second surprise was for the high performing men as the mean difference for successful (high performing) males was not significant, $F(1,46) = 3.25, p>0.05$ (a slight actual increase from 111.20 to 114.40). The interview had no stressful effect on successful males.

Noradrenaline

The second analysis compared noradrenaline levels in the pre and post tests for different genders and success levels. Table 2 shows the descriptive statistics for each group.

Table 2: Means and Standard Deviations for Noradrenaline across Gender and Performance Appraisal

<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>Performance Appraisal</th>
<th>M</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noradrenaline Pre-test</td>
<td>Male</td>
<td>High</td>
<td>172.30</td>
<td>30.43</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>212.78</td>
<td>35.64</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>High</td>
<td>166.17</td>
<td>31.87</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>198.25</td>
<td>42.45</td>
<td>4</td>
</tr>
<tr>
<td>Noradrenaline Post-test</td>
<td>Male</td>
<td>High</td>
<td>173.20</td>
<td>27.34</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>212.89</td>
<td>37.22</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>High</td>
<td>167.33</td>
<td>30.41</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>197.00</td>
<td>42.50</td>
<td>4</td>
</tr>
</tbody>
</table>
A significant main effect for success was found ($F(1,46) = 9.77$, $p<0.05$) indicating that regardless of gender and time, unsuccessful participants had higher levels of noradrenaline than did successful participants. As can be seen in Table 2, both unsuccessful males and unsuccessful females commenced and ended the interview with significantly higher levels of noradrenaline. All other main effects and interactions for this analysis were non-significant.

**Cortisol**

The final analysis involved the comparison of cortisol levels over time for the four groups. No significant main effects were found for gender, success or time in this analysis. However, a significant two-way interaction between gender and time was discovered ($F(1, 46) = 9.09$, $p<0.05$). For males the decrease was non-significant ($F(1,46) = 1.67$, n.s.) but for females the increase was significant ($F(1,46) = 5.64$, $p<0.05$). Thus, females experience significant cortisol activation over the course of the performance appraisal, indicating that they find it a stressful event.

**DISCUSSION**

Overall, the findings can be summarised as follows:

1. Males commenced the interview with significantly higher levels of adrenaline than females.

   Frankenhaeuser (1979) reports that adrenaline appears to be most responsive to mental stress. These higher levels of commencing adrenaline appear to lend some support to the contention of Fibiger et al (1984), Rauste-von Wright et al (1981) and Barnett et al (1987) that males have more psychophysiological responsive bodies than females. This result may indicate that the physiological cost of stress to males in the performance appraisal process is higher for males than for females.

2. Based on the measures of adrenaline, both successful and unsuccessful males commenced at similar levels of stress and both successful and unsuccessful females commenced at similar levels of stress. However, the measures of noradrenaline give a more complex overlay to this simple, gender based relationship. Unsuccessful candidates, both male and female, commenced and finished the interview with significantly higher levels of noradrenaline. This
result suggests that the unsuccessful candidates expected a stressful event. Lundberg and
Frankenhauser (1980) found that noradrenaline was associated with perceived effort to
succeed and with self-imposed stress. It is reasonable to suggest that the unsuccessful
candidates experienced self-imposed stress before the interview – although such a contention
assumes that the unsuccessful candidates had some inkling that they would be unsuccessful.

3. Lack of success in the performance interview increased levels of stress, as measured by
adrenaline. Bearing in mind that adrenaline is the marker that best reacts to the psychological
state (Rauste-von Wright et al 1981) this result indicates that the process and/or the content of
the interview increased the stress levels of the unsuccessful appraisees.

4. Both successful and unsuccessful females experienced significant increases in cortisol over
the course of the performance appraisal, indicating that they do find the experience a stressful
event. Cortisol works to repair the damage when noxious or potentially dangerous stimuli
have been perceived to traumatise the body in some way. Cortisol has also been associated
with increasing positive mood and affect (Henry 1993), although cortisol increases were also
associated with irritation, tenseness and tiredness (Lundberg et al 1989). So, the increases in
cortisol for females presents a complex picture but may lend support to the findings of
Bergman and Magnusson (1979) that the psychological, rather than the physiological, costs to
the females of the performance appraisal process may be higher. Also of interest is the finding
of Dobbins et al (1989) that females receive less favourable ratings when the appraisal was
administrative, as was the case in this study. Perhaps both successful and unsuccessful
females felt this pressure of a less favourable assessment during the interview and responded
with the excretion of the more complex of the catecholamines, cortisol.

5. The successful males had no changes in their catecholamines indicating that, of all the groups,
they found the performance appraisal process non-stressful.

ISSUES FOR MANAGERS
Generalising from the results of this small study (n=50) should be undertaken with care. However, some cautionary messages for managers should be raised. Given the pivotal nature of performance appraisal in the strategic HRM (SHRM) process of any organisation (see, for example, Delahaye, 2005 and Fisher et al, 1999), these cautionary messages do need to be highlighted.

One of the key contemporary SHRM issues, is the retention of staff (Donaldson 2005). Talent is the driver capable of increasing or limiting the capability or capacity of the organisation (Sullivan, 2005). Therefore, the need to manage the performance appraisal process for underperforming staff, both male and female, is critical. Contrary to the usual practice, even in administrative performance appraisal, the appraiser will need to demonstrate more subtle and developmental abilities than that for the successful appraisees. In particular, it is recommended that managers consider the importance of providing the unsuccessful appraisees with hope for the future. Specifically, managers should indicate that another, more developmental session, will follow in a few days to help the unsuccessful staff member to improve on the deficiencies identified and ensure that such a developmental intervention does indeed eventuate.

Most importantly, though, results in this study indicate that women react in a physically different fashion to males and this result has some important messages for workplace discrimination and gender equity. Thus performance systems that are masculine, i.e. confrontational, numeric and quantitative may have a significant negative impact on women. In addition, managers need to recognise that performance appraisal systems are not one size fits all. Best practice may mean that performance systems need to be more open and flexible (Tattersall & Morgan, 1997) in order to account for individual differences. For example, from this research, managers can expect males to display more overt signs of stress, such as sweating and fidgeting. Some safe, task-orientated lead-in time, for example asking the male respondent to list the achievements for the year, will ‘burn up’ some of the accumulated adrenaline. However, females may need more time to at the end of the interview to allow the cortisol levels to wind down. Hence, performance systems may need to be
“looser” and more flexible, with less emphasis on measurement in order to foster creativity and productivity (Scatterfield & Mulenhard, 1997).

Contemporary managers need to keep in mind that gender equity in the appraisal process does not mean treating all staff exactly the same and measuring the exact same criteria. Staff will perform well in different areas, and poorly in others. In this study female high performers exhibited calm confident relationships at work, were happy with co-worker evaluations and perceived as team players. Similar findings in Boswell and Boudreau (2000) tied team behaviour to satisfaction and productivity. Anecdotally such women were often said to be the “glue” that held the workgroups together. Contrast this to the high performing men who were insecure, competitive and results driven. Such men were not team players, yet they were described as giving direction, leaders of the group and operationalised the vision of the firm. So we have leaders and team players, both necessary to the firm’s performance, yet males were often seen as being more valuable because of their congruence with the vision of the firm. Also given the nature of the appraisal document, i.e. it had a lot of questions about leadership, vision and goals, it is little wonder competitive leadership is rewarded and co-operative team work disparaged. Finally, given the central and interactive nature of performance appraisal, these findings, although tentative, could have implications for in other HRM functions, such as workplace discrimination, gender equity and workplace health and safety.

References


