

**The Impact of Rater's Individualism/Collectivism on  
Discomfort With Peer Performance Evaluation**

Maria Carolina Saffie-Robertson

A Thesis

in

The John Molson School of Business

Presented in Partial Fulfilment of the Requirements  
For the Degree of Master of Science (Administration) at  
Concordia University  
Montreal, Quebec, Canada

March 2008

© Maria Carolina Saffie Robertson, 2008



Library and  
Archives Canada

Bibliothèque et  
Archives Canada

Published Heritage  
Branch

Direction du  
Patrimoine de l'édition

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file* *Votre référence*  
*ISBN: 978-0-494-40985-5*  
*Our file* *Notre référence*  
*ISBN: 978-0-494-40985-5*

**NOTICE:**

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

**AVIS:**

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

---

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

  
**Canada**

## **Abstract**

### **The Impact of Rater's Individualism/Collectivism on Discomfort with Peer Performance Evaluation**

**Maria Carolina Saffie Robertson**

A study was conducted to determine the relationship between rater's individualism/collectivism, discomfort with a peer evaluation system, and rater leniency. It was hypothesized that collectivism would be positively related to discomfort with evaluating a peer, while the relationship would be negative in the case of individualism and discomfort. This study also attempted to corroborate previous findings establishing a positive relationship between discomfort and rater leniency. In order to test the hypotheses, the Discomfort with Peer Evaluation scale (DPE) was developed. One hundred and five undergraduate students participated in this research. The data supported a positive relationship between collectivism and discomfort. Contrary to what was expected, data supported a positive relationship between individualism and discomfort. The data collected confirmed the existence of a positive relationship between discomfort and rater leniency. These results highlight the impact of cultural background on performance appraisals. Further research is needed to determine why both collectivism and individualism are positively related to discomfort.

## **Acknowledgements**

This thesis wouldn't have seen the light of day if it hadn't been for the guidance, support and assistance of many. I would like to thank my supervisor, Prof. Stéphane Brutus, for his help and patience. I'll forever be in debt for his encouragement in times of despair and frustration. My deepest gratitude also goes to the members of the committee, Prof. Linda Dyer and Prof. Terri Lituchy for all their support and commitment to this project. I honestly appreciate the time and effort they dedicated.

It is extremely important to acknowledge the contributions of my professors and friends at Concordia University. I'm particularly grateful for the listening ears and open hearts of Anamaria Kosa, Jason Lupachow and Marco Morelli. I would also like to thank my long time friend, Paulina Asenjo, for her company and advices.

I know that I am immensely fortunate to have the unconditional support of both my immediate and my extended family. Without my husband's strength and vision I probably wouldn't have been able to live and study abroad, away from my home and my loved ones. I love you much more than words can express. Finally, I would like to thank my Wendys and my Juan Carloses. To say that my parents and siblings have helped me during this process would be an understatement. All I've ever achieved is thanks to them and their never ending love and encouragement. Thank you for helping me find my path.

To my husband, my parents,  
my sister and my brother.

## Table of Contents

Chapter		
	List of Figures and Tables	vii
I.	Introduction	1
II.	Literature Review	12
	a. Cultural Dimensions	12
	b. Performance Appraisal and Individualism/Collectivism	15
	c. Discomfort and Leniency	18
	d. Discomfort and Reaction to Feedback	20
III.	Methodology	24
	a. Sample	24
	b. Procedure	24
	c. Measures	25
IV.	Results	34
V.	Discussion	46
	a. Implications for Practice	48
	b. Limitations	50
	c. Conclusions	52
VI.	References	54
VII.	Appendix 1: DPE Scale	65

## List of Figures and Tables

### **Figures**

Diagram 1: Model	11
Diagram 2: Hypotheses	23

### **Tables**

Table 1: DPE Communalities	29
Table 2: DPE Pattern Mix	29
Table 3: DPE Reliability Analysis	29
Table 4: Correlations and Reliabilities	36
Table 5: DPE as Mediator of the Collectivism-Leniency Relationship	37
Table 6: DPE as Mediator of the Individualism-Leniency Relationship	38
Table 7: Relationship between Discomfort and Average Evaluation	39
Table 8: DPE Subscales Relationship with Collectivism	41
Table 9: DPE Subscales Relationship with Individualism	42
Table 10: Relationship between Discomfort and Reaction to Feedback	43
Table 11: DPE as Mediator of the Collectivism-Reaction to Feedback	44
Table 12: DPE as Mediator of the Individualism-Reaction to Feedback	44
Table 13: DPE subscales and Leniency Relationship	45

## **Introduction**

To estimate how much money is invested each year around the world in performance appraisals is extremely difficult, given the fact that this technique is one of the most commonly used tools for the management of human resources. The results obtained through performance appraisals form the basis for many management decisions such as promotions, terminations, payment increases and even training and development (Bernardin & Villanova, 2005).

In spite of its wide spread use, performance evaluations are a continual source of disappointment for organizations because they are usually received by employees with considerable resistance (Banks & Murphy, 1985), and because of issues dealing with low reliability and validity (Bernardin & Villanova, 2005). Performance appraisal's reputation has not improved with time. Researchers have investigated these allegations only to find that performance appraisal reliability is influenced by many factors, that vary from rater's agreeableness and conscientiousness (Bernardin, Cooke, & Villanova, 2000) and internal politics (Longenecker, Sims & Gioia, 1987), to the social context (Levy & Williams, 2004) and ratees' participation in the appraisal (Cawley, Keeping, & Levy, 1998).

Although several factors have been identified through recent studies, researchers have speculated that there are many other factors yet to be determined that are linked to performance appraisal reliability. This research proposes that two new factors should be incorporated to the list of issues that impact the reliability of performance appraisals: rater's culture and the target of the evaluation.



## The Impact of Discomfort on Performance Appraisal's Reliability

As described by Fletcher (2001, p. 473), performance appraisals are a "... variety of activities through which many organizations seek to assess employees and develop their competence, enhance performance and distribute rewards." The appraisal is usually done by the supervisor in a three-step process (Murphy & Cleveland, 1991). First, the rater obtains information regarding the ratee's performance. This information can be obtained through direct and indirect observation. The gathering of the information should be done after determining which behaviours are relevant to the organizational goals. It is important to mention that these relevant behaviours should be linked to specific tasks that are part of the ratee's job (Murphy & Cleveland, 1991). Second, the rater applies judgement to the obtained information in order to combine it and integrate it into a consolidated mass of information regarding job performance (Murphy & Cleveland, 1991). Finally, the rater evaluates the ratee, assessing whether the worker's job performance is good, poor or average according to the organizational standards (Murphy & Cleveland, 1991). Once the performance appraisals are completed, the outcomes help organizations determine which employees should be promoted, transferred, terminated, disciplined, or trained (Bernardin & Villanova, 2005; Malos, 1998).

In 2004, Levy and Williams argued that performance appraisals are immersed in a social context that impacts the effectiveness of the evaluation. The authors identified three main groups of social context variables: Distal variables (organizational culture, economic conditions and HR strategies, among others), Process proximal variables (organization's policies, task characteristics and leadership, to mention some) and Structural proximal variables (such as appraisal goals and purposes, frequency of the

appraisal and performance dimensions). All these variables could have an impact on the effectiveness of the appraisal, which the authors divide into three main outputs: rater errors and biases, rating accuracy and appraisal reactions (Levy & Williams, 2004).

One of the social variables that have been pinpointed as having a direct influence on rating accuracy is rater discomfort with the appraisal. Rater discomfort refers to the unease that a rater can feel when he or she is forced to evaluate the work performance of others. As explained by Smith, Harrington & Houghton (2000, p. 21) previous "...findings suggest that the performance appraisal process is uncomfortable for many raters." This discomfort with performance appraisal process can be rooted in a discomfort with performing any or even all the actions and behaviours that a rater is expected to execute as part of his/her role as an evaluator. These actions and behaviours include actions such as directly and indirectly observing ratee's behaviour and filling out the evaluation form. As suggested by Villanova, Bernardin, Dahmus & Sims (1993), rater discomfort can be present at any of the mentioned three stages of the appraisal. Therefore, the rater can be uncomfortable with monitoring performance, giving feedback and/or actually filling out the appraisal of an employee's work. In other words, rater discomfort is the degree to which the rater is uncomfortable with the "... enactment of a subset of role requirements..." (Villanova et al., 1993, p. 791)

One of the first attempts to measure rater discomfort with the performance appraisal was conducted by Villanova et al. (1993), who developed the Performance Appraisal Discomfort Scale (PADS). This 20-item scale measures rater's discomfort towards a top-down evaluation system, in an attempt to find "...a partial explanation for the occurrence of lenient ratings." (Villanova et al., 1993, p. 790). Leniency can be

defined as the tendency of some raters to give evaluations that are higher than what the ratee deserves (Villanova et al., 1993). Rater leniency has been studied since the early 1950s (Austin & Villanova, 1992). In fact, in 1954 Guilford hypothesized that it would be a stable tendency on raters. As Kane, Bernardin, Villanova, & Peyrefitte (1995) pointed out, although leniency was identified almost 60 years ago, much of the research done so far only started in the 1990s, leniency proved to have a deep influence on appraisal's accuracy.

The effects of rater's leniency are varied and extended. At the immediate level, rater leniency alters the accuracy of the appraisal (Kane et al., 1995). This diminished accuracy has an impact on the validity of the appraisal findings, which could impact the effectiveness of the human resources management decisions based on appraisal results (Kane et al., 1995). Furthermore, rater leniency could have an impact on employees' perception of fairness and significance of the appraisal process, while increasing "...employee dissatisfaction with both performance management and reward systems." (Bernardin, Cooke, & Villanova, 2000, p. 232) Rater leniency can even have legal implications for organizations when decisions such as terminations are challenged in court (Bernardin et al., 2000; Kane et al., 1995).

Literature on rating inaccuracy has concluded that rater leniency is closer to being a deliberate distortion of performance evaluations than to unintended mistakes (Longenecker, Sims & Gioia, 1987; Murphy & Cleveland, 1995; Tziner, Murphy & Cleveland, 2005). Raters would purposely alter ratings for reasons that were not related to the ratee's job performance. As Tziner & Murphy (1999) commented, leniency could be a resource for raters that want to protect themselves from criticisms and to promote

their position within the group. Furthermore, researchers noticed that leniency would appear in an effort by raters to avoid potentially unpleasant situations or for political reasons that had little or nothing to do with actual performance (Longenecker et al., 1987; Smith et al., 2000). In fact, Harris (1994) identified five categories of negative consequences that are associated with performance appraisals and that are avoided by raters. One of these categories refers to the damage of the rater-ratee relationship. Harris explains (1994, p. 741) that "... a common concern expressed by managers is that making accurate (i.e., lower than the employee expects) performance ratings or giving negative performance feedback will hurt their relationship with the subordinate." Furthermore, Harris (1994) explains that many raters fear that giving realistic and appropriate ratings or providing performance feedback will discourage ratees instead of motivating them.

Villanova et al. (1993) relied on these previous findings that had related leniency with performance appraisal settings to hypothesize a relationship between discomfort and leniency. Through this study, Villanova et al. (1993) discovered that rater's discomfort with the performance appraisal predicts rater's leniency. The authors concluded that when raters were uncomfortable with the performance appraisal, leniency was higher than when raters are comfortable with the evaluation process (Villanova, Bernardin, Dahmus & Sims, 1993).

Given the magnitude of the possible consequences and the pervasiveness of rater leniency, researchers have focused on the study of rater discomfort and its relation to rater leniency. In fact, Villanova et al. (1993) findings were later corroborated by the results obtained by Tziner and Murphy in 1999. Using PADS on 29 managers, the authors concluded that higher discomfort with the performance appraisal was positively

related to higher ratings. In other words, raters who presented higher levels of discomfort with the appraisal were likely to give higher ratings than raters with lower levels of discomfort.

Since the consequences of rater discomfort could prove to be significant for the success of a performance appraisal system, researchers have investigated possible antecedents or variables that could predict rater discomfort. The causes of rater discomfort can be divided into two main branches: structural factors and personal characteristics of the rater.

#### Structural factors affecting rater discomfort

One branch of research has focused on structural factors, or in other words, items that surround the performance appraisal system. In 2000, Smith et al. investigated three possible predictors of rater discomfort, concluding that rater's beliefs in the importance of the performance appraisal, communication reticence, and the time raters have been supervising the ratees, are all related to rater's discomfort with the performance appraisal.

A structural factor that has not yet been tackled is the 'direction' of the appraisal, meaning if the rater evaluates top-down, bottom-up, or laterally. Until very recently, performance appraisals were conducted in a top-down direction, where the supervisor who would evaluate the work performance of the subordinate (Gillespie, 2005). In the past decades, organizations have applied a new system that includes performance feedback from other sources. This technique, called multi-source feedback can include the evaluation by the supervisor, peers, self, subordinates, and, even clients. This method

would give the employee a "... multi-dimensional view of his or her performance that reveals discrepancies, consistencies, strengths and developmental areas." (Gillespie, 2005, p. 362)

Peer evaluations in particular have caught the attention of scholars, because these evaluations have the potential to be a better predictor of performance than traditional supervisor ratings (Costigan, Insinga, Kranas, Ilter, Kurechov, & Berman, 2005). Costigan et al. (2005) argued that co-workers have more opportunities to effectively observe peer's performance without the usual constraints of the hierarchical nature of the relationship between workers and supervisors. As commented by Drexler et al. (2001) there are different measurement forms (ranging from ratings to comments) and uses (administrative decisions, developmental purposes or both) for peer evaluations, but the function of the appraisal is the same across organizations: "... to provide job incumbents with valid information to allow them to maintain or improve performance or to provide the basis for administrative decisions." (p. 334).

The rationale behind this relatively new performance appraisal approach is that by including more than one point of view of the same performance, different and valuable information will be added (Drexler et al., 2001). Indeed, in a meta-analysis done in 2001, Conway, Lombardo & Sanders found that adding evaluations of peers and subordinates adds validity to the appraisal. The authors noticed that "...for individual ratees there will be disagreements (as noted in the past), but these results suggest that the disagreements may well provide valid information." (Conway et al., 2001, p. 297) In fact, previous research has shown that because of normal daily interactions, each source has different points of view that make them ideal to observe specific behaviours. For instance, research

has shown that peers are the more suited source to evaluate a worker's team performance, but are not an appropriate source to evaluate a worker's supervision capabilities or leadership styles (London & Smither, 1995). The underlying assumption behind this conclusion is that every source provides an accurate evaluation, implying that disagreements are due to different points of view and not to raters' inaccuracies.

As Drexler et al. (2001) explained, research on peer evaluation is vast. Studies have been conducted on various aspects of peer evaluations, such as attitudes towards peer evaluations, correlations between peer and supervisor ratings, and incremental validity of peer evaluations (Drexler et al., 2001; Conway et al., 2001). Although peer evaluations have been researched extensively, some areas remain uninvestigated. That is the case of the possible relationship between peer appraisals and rater discomfort. Even though the target of the appraisal could have an impact on the rater discomfort with the evaluation, so far researchers have focused on developing and using scales that only measure discomfort with the traditional top-down evaluation. Since multi-source feedback has become increasingly popular, it is necessary to fill the gap in the existing literature by evaluating the possible effects of other appraisal directions (bottom-up, lateral) on rater discomfort.

#### Rater's personal characteristics

A second branch of research has focused on rater discomfort caused by rater's own personal characteristics. Bernardin, Cooke, & Villanova (2000) used the NEO Five Factor Inventory in an effort to match personality traits with rater discomfort. The authors

found that Agreeableness and Conscientiousness are related to rater discomfort and therefore to rater leniency. Individuals with high levels of agreeableness would tend to provide higher ratings, while highly conscientious individuals would give lower ratings (Bernardin et al., 2000)

Although the studies by Smith et al. (2000) and by Bernardin et al. (2000) identify variables that cause rater discomfort with the performance appraisal, (and in consequence, the rater's leniency), more research is needed in order to identify other proximal or distal variables (as defined by Levy and Williams, 2004) that could have an impact on rater's discomfort.

This study proposes that rater's culture may have an influence on the rater's discomfort with the performance appraisal. More specifically, this paper suggests that individualism/collectivism could explain rater's discomfort in both giving and receiving feedback related to work performance.

To believe that culture could have an impact on rater's discomfort is to recognize that as globalization increases performance appraisal and feedback have become more complex. Managers have to deal with employees from their own countries and cultures as well as with employees from different countries and cultures (Milliman, Nason, Gallagher, Huo, Von Glinow & Lowe 1998). Performance evaluations have to be done for employees from different backgrounds that may react to this process in diverse ways. In fact, even though the results obtained by Conway et al. (2001) encourage the use of multi-source appraisals in organizations, these findings should not be considered conclusive. This meta-analysis included only a couple of studies done outside of North



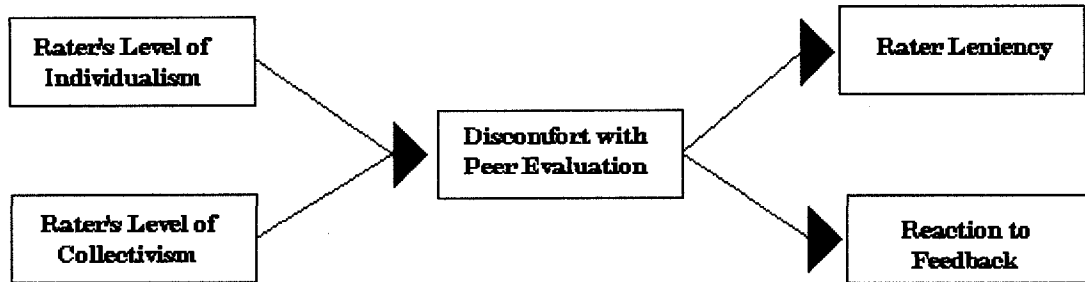
America. Conway et al. (2001) did not consider culture as a factor that could affect the results obtained.

Some researchers have established that the relationship between culture and performance appraisal is extremely important (Milliman et al., 1998; Brett, Tinsley, Janssens, Barsness, & Lytle, 1997; Fletcher & Perry 2001). Although these researchers have recommended managers to not apply the same practices in every country, they have also acknowledged that more research is needed in order to recognize in what extent, how and which cultural dimensions affect the performance appraisals (Milliman et al., 1998; Sully de Luque & Sommer, 2000). This research is a response to that call, bringing some light to the relationship between one cultural dimension, individualism/collectivism, and one aspect of the performance appraisal, evaluation reliability.

Therefore, this thesis has two main objectives: The first objective is to test the model presented in Diagram 1. This model proposes that rater's discomfort can be partially explained by the rater's level of individualism/collectivism. The rater's culture will not only impact his/her discomfort with the appraisal system, but also the tendency of the ratings (leniency) and the reactions towards feedback regarding the rater's own performance. In other words, individualism/collectivism will have an affect on the rater's level of discomfort with the performance appraisal, which will impact the ratings given by the rater to peers and the rater's reaction towards feedback of his/her own performance.

**Diagram 1**

**Proposed Model: The Effect of Individualism/Collectivism on Discomfort,  
Leniency and Reaction to Feedback**



The second objective is to develop a scale to measure rater's discomfort with horizontal performance appraisal, in other words, a scale to measure the discomfort a rater feels when confronted with evaluating his/her peers. Two scales exist to measure rater's discomfort with performance appraisal, Performance Appraisal Discomfort Scale (PADS, Villanova, Bernardin, Dahmus & Sims, 1993), and the Performance Appraisal Self-Efficacy Scale (PASES, Bernardin & Villanova, 2005). However, both scales measure discomfort when the rater conducts a top-down evaluation. Since there are no scales to measure rater's discomfort outside the traditional downward performance evaluation, this research would fill a gap in the existent literature. It is particularly important to develop such a scale since as mentioned earlier, the use of multi-source feedback evaluations is becoming more common. Both researchers and practitioners can use this new tool as a criterion to decide whether or not including peer evaluation in the performance appraisal system is feasible or relevant in different contexts, industries and businesses.

## Literature Review

### Cultural Dimensions

Culture has been defined many times and in many different ways (Milliman et al., 1998), but the most accepted one seems to be similar to the one proposed by Trompenaars and Hampden-Turner (1998, p. 13) as "...a shared system of meanings [that] dictates what we pay attention to, how we act and what we value." Milliman et al. (1998, p. 162) go a bit further defining culture as "... a set of basic assumptions shared by individuals with the same national origin (...) to be manifest in values, behaviours and artefacts..."

Hofstede made one of the first systematic attempts to understand and classify culture according to different dimensions. As Hofstede (1983) explains, he obtained data on cultural differences 'by accident'. While he was working as a psychologist for IBM, he surveyed almost 120,000 workers of the corporation across 40 different countries. Hofstede noticed that some questions dealt with employees' perceptions instead of their actions and realised that these were the values that could define a culture. Hofstede (1983) defined four dimensions: Individualism versus Collectivism, Strong or Weak Uncertainty Avoidance, Masculinity versus Femininity, and Large or Small Power Distance. Years later, Hofstede determined a fifth dimension: Long versus Short Term Orientation (Hofstede, 2002b).

Although there seems to be a never ending ongoing debate over the real existence of cultural dimensions (Hofstede, 2002a; McSweeney, 2002a & 2002b), researchers have used these dimensions and others, such as the Neutral versus Emotional, Specific versus

Diffuse, and Achievement versus Ascription dimensions proposed by Trompenaars and Hampden-Turner (1998), in an attempt to understand and give suggestions to managers that have to deal with the complexities of managing a culturally diverse work force.

One of the most commonly used dimensions is Individualism/Collectivism. As Oyserman, Coon & Kimmelmeier (2002) pointed out, the word individualism is relatively old and can be traced to the French Revolution. Originally, "... individualism was first used to describe the negative influence of individual rights on the well-being of the commonwealth." (Oyserman et al., 2002, p. 3) Collectivism, on the other hand, has been the word use to describe the exact opposite phenomenon, the bonding of individuals for the sake of the group (Oyserman et al., 2002).

The individualism versus collectivism dimension as defined by Hofstede (1983) is the extent to which the interest of the individual comes over and above the interest of a group. In societies with high level of individualism, the ties between individuals are very loose, and people feel a responsibility only to themselves (Hofstede, 1983). On the other hand, in societies with high collectivism, individuals are responsible and concern of the wellness of the group as a whole (Hofstede, 1983; Sivadas, Bruvold, & Nelson, 2008).

Individualism and collectivism in this context were conceptualized as exact opposite concepts, describing the relationship between an individual and other people in a certain environment of the same continuum, (Hofstede, 1983). Hofstede (1984) measured the level of individualism/collectivism of a culture on a scale from 0 to 100, where 0 represents the lowest score (high individualism or low collectivism) and 100 represent the highest score possible (low individualism or high collectivism).

Hofstede's model and theory are more than 20 years old, but are still in use and recent studies have demonstrated that the findings are valid. In 2003, Chirkov, Ryan, Kim and Kaplan checked the individualistic aspect of 4 countries: South Korean, Russian, Turkish and North American. The results are similar to those obtained by Hofstede, establishing that "... Koreans perceived their own culture as relatively collectivistic, Americans viewed theirs as relatively individualistic, and Russians emerged as somewhat of a mixed model." (Chirkov et al., 2003, p. 103).

What seems clear is that although globalization is permeating societies across the world, there are certain cultural values that do not change and that are different across borders. The individualistic or collectivistic nature of a society seems to be one of these factors that are stable in spite of globalization, economic liberalization and other cross-border phenomena.

Recent research has debated the conceptualization of individualism and collectivism as complete opposite concepts. The core of this issue could be rooted in the level of analysis. While individualism may be the exact opposite of collectivism at the group level, at the individual level these constructs intertwine (Singelis, 1994; Singelis, Triandis, Bhuvuk, & Gelfand, 1995; Oyserman et al., 2002). In fact, at the individual level, low collectivism may not be equivalent to high individualism and vice versa, but two related yet different constructs (Singelis, 1994). Triandis (1989) proposed that at the individual level, people are both allocentric (collectivistic) and idiocentric (individualistic). Depending on the situation that a person is confronted to, he or she will refer to either collectivistic or individualistic aspects (Singelis, 1994). Having both characteristics at the individual level allows people to be culturally sensitive and to adapt

according to the cultural context (Singelis, 1994). In other words, "... it seems likely that two aspects of self in relation to the collective can coexist, even though most prior attempts to measure individualism-collectivism have assumed a single bipolar dimension." (Singelis, 1994, p.583)

Therefore, it is more appropriate to measure individualism/collectivism as separate constructs when the samples are smaller and homogenous, and when the level of analysis is the individual, as it is the case of this study.

#### Performance Appraisal and Individualism/Collectivism

As mention earlier, individualism/collectivism is the relationship between the interests of the individual and those of a group (Hofstede, 1983). Individualistic cultures emphasize more on personal achievements in performance evaluation while collectivist cultures are more likely to emphasize team-based achievements (Milliman et al., 1998). For instance, Milliman et al., (1998) found that employees in Spain, a collectivistic society, place great importance on emotions and personal relationships in the work place, and they believe that performance appraisal can never be able to give completely objective feedback.

Unlike individualistic societies, work in a collectivist environment is not related as much to an act of self-fulfilment or self-expression. Since usually an employee's priority is to fulfil the obligations according to the employment contract (Mendonca & Kanungo, 1996), the focus of the performance evaluation shifts according to culture. For example, in collectivistic societies performance appraisal systems tend to pay less

attention to job objectives to focus on group work and harmony (Mendonca & Kanungo, 1996).

In fact, Sully de Luque and Sommer (2000), suggest that in collectivist societies, performance appraisal and feedback will be oriented towards the group, while in individualistic societies, evaluation and feedback will be addressed to the employee, with no regards to the group were he/she works.

Furthermore, for Fletcher and Perry (2001), evaluation and feedback in collectivist societies will not only be group oriented, but will have a more positive connotation and will be focused on maintaining healthy relationships within the group and the organization. These authors commented that individualistic societies will have appraisals and feedback focused on job performance and relationships will not be as important as in collectivist societies (Fletcher & Perry, 2001). Moreover, Milliman et al., (2002) suggested that evaluation in individualistic cultures is directed to improve the individual's performance and self-identity, while in collectivistic societies, appraisal is directed to consider group's harmony and relationships.

Researchers have determined that highly individualistic societies tend to deal directly with conflict situations in performance appraisals while collectivist cultures are more likely to deal with conflict in a more indirect manner (Milliman et al., 1998). Therefore, high collectivist societies are more likely to minimize open criticism and attention to mistakes in conflict situations, denominated 'agreement management' (Milliman et al., 1998). The relation between the employer and employee in these highly collectivist societies typically involves protection in exchange for loyalty (Fletcher & Perry, 2001).

In highly collectivistic societies, actions that have the potential to cause conflict within the group are avoided. Previous research has identified as important for highly collectivist employees to reduce differences between members of the collective and maintain group harmony (Fletcher & Perry, 2001). Therefore, it is expected that collectivistic individuals will tend to be more uncomfortable with rating members of the group, aiming to preserve and encourage the cohesion of the group. More individualistic raters will be more comfortable with rating the performance of peers, with no particular regards towards protecting the harmony and unity of the group. Therefore, it is expected that highly collectivistic raters will present a higher level of discomfort with the performance evaluation of peers. On the other hand, individualistic raters are expected to be comfortable with rating the performance of peers, since they will not perceive the evaluation process as a source of possible disharmony or conflict within the work group. Hypotheses 1a and 1b propose the following:

**Hypothesis 1a:** Collectivism will be positively related to rater's discomfort with peer performance appraisal system.

**Hypothesis 1b:** Individualism will be negatively related to rater's discomfort with peer performance appraisal system.



### Discomfort and Leniency

As previously mentioned, research such as the studies by Villanova, Bernardin, Dahmus and Sims in 1993, Tziner and Murphy in 1999, and Bernardin and Villanova in 2005, found a direct and positive relationship between discomfort with performance evaluation and leniency. Looking for the causes of rater leniency, the study by Villanova et al. (1993) was the first research to link rater discomfort with leniency. Research had identified that leniency had little to do with unconscious errors as earlier believed, but was closer to being a deliberate distortion of performance ratings (Longenecker, Sims & Gioia, 1987; Murphy & Cleveland, 1995; Tziner, Murphy & Cleveland, 2005). The reasons why raters would deliberately alter ratings vary from impression management to political reasons (Longenecker, Sims & Gioia, 1987; Smith, Harrington & Houghton, 2000).

Given these results, research had already related leniency to other rater's characteristics such as rater's personality and rater's leadership style (Villanova, Bernardin, Dahmus & Sims, 1993). Villanova et al. (1993) relied on these previous findings that had related leniency with performance appraisal settings to hypothesize a relationship between discomfort and leniency. Through this study, Villanova et al. (1993) discovered that rater's discomfort with the performance appraisal predicts rater's leniency. The authors concluded that when raters were uncomfortable with the performance appraisal, leniency was higher than when raters are comfortable with the evaluation process (Villanova, Bernardin, Dahmus & Sims, 1993). The reason for this link between discomfort and leniency could be given by rater's perception that rating alteration could be a way to protect oneself from criticisms and to promote one's position

within the group (Tziner & Murphy, 1999). This view is supported by Longenecker et al. (1987, p. 183) who concluded that "...appraisals have a significance that reaches far beyond the few hours it takes to conduct them." The awareness of the implications of the appraisal mixed with ulterior motives such as controlling destinies, gaining influences and other political considerations taint the appraisal process, elevating discomfort and therefore, ratings (Longenecker, Sims & Gioia, 1987).

Villanova et al. (1993, p. 797) found that "... leniency can be predicted by rater individual differences in performance appraisal discomfort... [which] ... appears to be a relatively stable rater characteristic..." Six years later, Tziner and Murphy (1999) studied the relationship between discomfort and leniency in an effort to corroborate previous findings. The authors concluded that raters who indicated having higher levels of discomfort regarding the performance appraisal were more likely to give higher ratings (Tziner & Murphy, 1999).

Although the study by Tziner and Murphy (1999) had a very small sample (29 managers), these results are supported by the findings by Villanova et al. (1993) and Bernardin and Villanova (2005). Therefore, it is expected to find the same tendency of higher evaluations ratings by raters that are highly uncomfortable with the process.

**Hypothesis 2:** The discomfort of the rater with the performance evaluation process will have an impact on the leniency of the performance evaluation ratings. Specifically, the higher the level of discomfort a rater feels towards the appraisal system, the more lenient his/her ratings will be.

From the proposed hypotheses 1a, 1b and 2 as well as from the proposed model, it can be expected to find a mediation of discomfort for the relationship between individualism/collectivism and tendency of the evaluation.

**Hypothesis 3a:** Discomfort will mediate the relationship between collectivism and rater leniency.

**Hypothesis 3b:** Discomfort will mediate the relationship between individualism and rater leniency.

#### Discomfort and Reaction to Feedback

Previous research has determined that the factors that influence the reaction of a ratee to feedback regarding his/her own performance are many. One of the factors that affect reaction to feedback is the way the feedback is given or presented to the ratee. Atwater & Brett (2006) determined that the format of the feedback has an effect on the reaction towards feedback. Ratees would react more positively to feedback given as scores and comparative information than when that same feedback information was in the form of text (Atwater & Brett, 2006).

Ratee's personal characteristics also have been pinpointed as having an effect on reaction to feedback. Shrauger and Rosenberg (1970) determined that ratee's self-esteem also has an impact on reaction to feedback. Individuals with high self-esteem tend to react more positively to feedback than people with low self-esteem. In a study following

similar lines, Smith and Sarason (1975) determined that social anxiety has an effect on reaction to feedback: individuals with high levels of social anxiety will perceive the same feedback as more negative than individuals with low social anxiety. Furthermore, Smither et al. (2005) determined that rater's personal characteristics such as breadth of interest, emotional stability, responsibility and sociability would influence the individual's reactions towards feedback. Given the fact that the previously mentioned research has determined that some personal characteristics have an impact on reaction to feedback, it is expected that to find other personal differences that will also influence the reaction to feedback. This research proposes that the level of discomfort with the appraisal can be one of those individual characteristics that will have an effect on a rater's reaction to feedback regarding his/her own performance.

As mentioned earlier, previous research as determined that there is a link between rater discomfort and leniency (Villanova, Bernardin, Dahmus & Sims, 1993; Tziner & Murphy, 1999; Bernardin & Villanova 2005), which is one area of the performance appraisal system. Since rater discomfort refers to the level of distress that a rater feels towards the performance appraisal system as a whole, it would be reasonable to expect a link between the level of discomfort of a rater and other areas of the performance appraisal system. In particular, it is expected to find a similar relationship between rater's discomfort level and rater's reactions towards the feedback of his/her own performance.

Although no previous research has studied this relationship in particular, some related studies have shed light on the subject, implying that there could be a relationship between discomfort and reaction to feedback. In a study conducted in 1988, Russell and Goode found a positive relationship between rater's reactions to feedback of their own

performance with system satisfaction. In this study, system satisfaction was conceived as the degree of acceptance of the peer evaluation system. Furthermore, McEvoy and Buller (1987) found that positive raters' reactions towards feedback were related with good experiences with past peer rating experiences. In other words, raters were comfortable with receiving feedback when they perceived prior experiences with the appraisal as satisfactory. Although both studies researched system satisfaction, this variable is closely related to rater discomfort with the evaluation system. Therefore it is expected to find a similar relationship as the ones previously found but this time between discomfort and reaction to feedback.

**Hypothesis 4:** Rater's reactions towards feedback regarding their own performance will be affected by the rater's level of discomfort. It is expected to find that the higher the discomfort with the performance evaluation, the higher the discomfort with receiving feedback will be.

From hypotheses 1a, 1b and 4, it is expected to find that discomfort will mediate the relationship between individualism/collectivism and reaction to feedback. In fact, for highly collectivistic raters, dealing with feedback regarding their performance could be as disagreeable as rating others. Highly collectivistic raters will not only avoid giving but also receiving feedback dealing with work performance. The following hypotheses follow the propositions by Sully de Luque and Sommer (2000) who described that feedback behaviour should be different across cultures. Therefore, different cultures will imply different feedback needs and sources. This means that the cultural background will

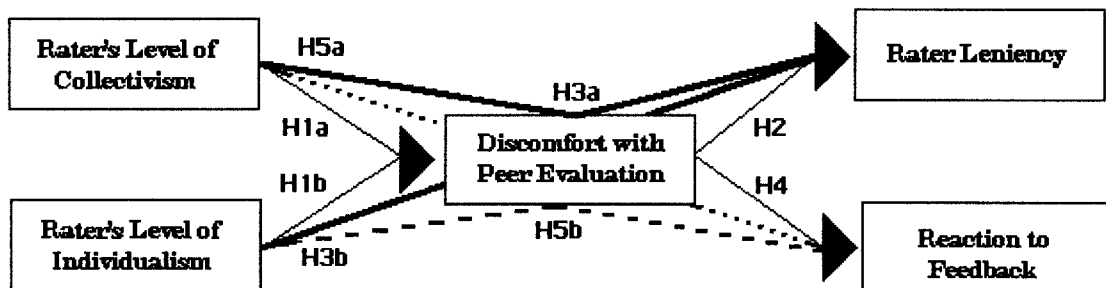
affect the determination with which an employee will seek feedback (Sully de Luque & Sommer, 2000).

**Hypothesis 5a:** Discomfort will mediate the relationship between collectivism and reaction to feedback.

**Hypothesis 5b:** Discomfort will mediate the relationship between individualism and reaction to feedback.

**Diagram 2**

**Hypotheses**



## **Methodology**

### Sample

Undergraduate students from a second year business course were invited to participate. The total sample used in this study consisted of 105 students (of 5 different sections, with classes given by 4 different professors), of which 57 were male and 48 were female. On average, the age for this sample was 25 years old, while the reported GPA was 3.5.

As part of the course requirements, students had to form teams in order to develop a group project. The size of the group varied between 3 and 6 members, with the average being 4 members per group. This team had to work during the semester on a research that dealt with the study of individual behaviour in formal organizations. Students could research topics such as motivation, leadership and group behaviour, among others. By the end of the semester, the group had to hand in a printed version of the project, as well as to present it orally to the rest of the class.

### Procedure

At the beginning of the semester, participants were asked to fill out an online survey designed by the researchers, previously approved by the University Human Research Ethics Committee (HREC). Students were asked to log in and complete the questionnaire that included demographic questions (such as age, gender, GPA and ethnicity), as well as scales to measure their level of individualism/collectivism and impression management. Students were told that their responses would contribute to a

study regarding evaluator discomfort, but no other specific information regarding the study was given. There was no mention about the hypotheses or the links that would be tested between individualism/collectivism, rater discomfort and rater reaction to feedback. This was done in order to avoid responses tainted by social desirability.

Weeks later and as part of their course, the students had to evaluate their group members through an online evaluation system. This second questionnaire consisted of an existing system that has been used at Concordia University for several years. The original questionnaire was modified in order to include questions regarding the students' discomfort with the peer evaluation system, and their reactions towards the feedback they receive regarding their own work performance during the group project.

Although anonymity was ensured, the students' ID number was recorded in order to match the information obtained through both instruments. The course instructor never had any access to the collected data and after the matching was done, all identification was erased.

### Measures

*Rater's Individualism/Collectivism:* To measure the individualism/collectivism of the evaluator, the Self-Construal Scale designed by Singelis (1994) was used. Each item is answered as a Likert scale with 7 points (1=Strongly Agree, 7=Strongly Disagree). This scale has been widely used since its development in 1994 so it seemed appropriate to use it in this context. The scale includes items such as: "I have respect for the authority figures with whom I interact"; "I'd rather say "No" directly than risk being misunderstood"; and "I enjoy being unique and different from others in many respects".



The Cronbach  $\alpha$  were 0.73 for individualism (independent) and 0.70 for collectivism (interdependent) subscales.

*Rater discomfort with peer evaluations:* A scale had to be developed in order to could measure this new scenario of peer evaluation, that hasn't been considered in the existent rater discomfort scales. Therefore, to measure rater discomfort with horizontal evaluations, a scale DPE (Discomfort with Peer Evaluation) was constructed. Items from the PADS (Villanova et al., 1993) plus some items from the Performance Appraisal Self-Efficacy Scale (PASES, Bernardin & Villanova, 2005) were modified to make them appropriate to measure rater's discomfort with peer evaluations in an academic environment. Extra items that do not appear in any of the previously mentioned scales were added too, in an attempt to make the scale more complete. This new rater discomfort scale asks the participants to think about the process of rating the performance of a peer or group member and evaluating their level of discomfort in a 5 point Likert Scale, where 1= No Discomfort and 5= High Discomfort (See Appendix 1). The reliability of the DPE scale was 0.84.

This scale includes 3 subscales, which resemble the three-step process by Murphy & Cleveland (1991). The first subscale corresponds to the discomfort of collecting information that will be used for evaluation purposes. This subscale was called "Collecting Info" and it includes the actions for obtaining information regarding the employee's performance. This first subscale is very similar to the first step of the model by Murphy and Cleveland (1991). The two items that form this subscale are related to the

gathering and use of information for appraisal purposes. “Collecting Info” has a Cronbach  $\alpha=0.83$ .

The second subscale was named “Rating” because it assesses the discomfort of actually rating a peer, therefore including the subset of actions of both the second and third stage of the model by Murphy and Cleveland (1991) that deal with evaluating a peer. In this stage and in order to consolidate the obtained information regarding job performance, the rater applies judgement to combine it and integrate it (Murphy & Cleveland, 1991). After the information is judged and pooled, the rater evaluates the ratee, assessing whether the worker’s job performance (Murphy & Cleveland, 1991). The alpha of the “Rating” subscale was 0.77.

Finally, the third subscale measured the discomfort felt by the rater after the rating process was over. In this stage, students can review the ratings they were given by their team members and have to deal the possible consequences of the ratings they gave. Therefore, this subscale was named “Post Rating”, with  $\alpha= 0.75$ .

### Factor Analysis

To test if the scale worked appropriately, a principal factor analysis was done (Whitley, 2001). It is important to mention that the sample size is relatively small for this kind for analysis, therefore the results of this factor analysis can only be considered exploratory.

Since the items may be correlated, an oblique rotation (Promax) was used in this analysis. Table 1 shows the communality matrix for the DPE scale. As it can be noticed, the common variance of seven of the ten items is above 0.60 and two of the remaining

show an extraction higher than 0.50. Only item 3 of the DPE scale shows a lower extraction, with 0.35. These results are similar to the ones obtained at the reliability analysis of the scale, that shows that all the items are useful and linked to each other, therefore forming a tight scale (see table 3).

Taking 0.35 as the cutting point, it can be noticed that the Pattern Matrix (table 1) shows two main factors. The first factor includes items 3 to 10, basically both subscales Rating and Post Rating as one factor. The second factor consists of items 1 and 2, previously categorized as the Collecting Information subscale. Even though these results show that the Rating and Post Rating subscales are correlated, it was decided not to merge them and to use them as they were designed, for these items relate to similar but yet different aspects of the performance appraisal. The fact that the reliabilities for both subscales are above 0.74, reaffirms the notion that using the subscales as previously designed for further exploratory analysis does not contradict the results obtained through the factor analysis. These results verify that the created DPE works suitably for this sample.

**Table 1: Communalities**

	Initial	Extraction
DPE1: Collecting Info	1	0.75
DPE2: Using observations	1	0.82
DPE3: Evaluating perf.	1	0.35
DPE4: Assigning ratings	1	0.58
DPE5: Distributing points	1	0.51
DPE6: Writting feedback	1	0.63
DPE7: Talking to peer	1	0.60
DPE8: Giving suggestions	1	0.57
DPE9: Working again	1	0.60
DPE10: Develop friendship	1	0.41

Extraction Method:

Principal Component Analysis.

**Table 2: Pattern Matrix**

	Component	
	1	2
DPE1: Collecting Info		<b>0.85</b>
DPE2: Using observations		<b>0.97</b>
DPE3: Evaluating perf.	<b>0.39</b>	
DPE4: Assigning ratings	<b>0.72</b>	
DPE5: Distributing points	<b>0.50</b>	
DPE6: Writting feedback	<b>0.70</b>	
DPE7: Talking to peer	<b>0.77</b>	
DPE8: Giving suggestions	<b>0.70</b>	
DPE9: Working again	<b>0.84</b>	
DPE10: Develop friendship	<b>0.70</b>	

Rotation Method:

Promax with Kaiser Normalization.

a.- Rotation converged in 3 iterations

**Table 3: Reliability Analysis DPE Scale ( $\alpha=0.84$ )**

	Scale Mean if Scale Variance			Cronbach's Alpha if Item Deleted
	Item Deleted	if Item Deleted	Corrected Item-Total Correlation	
DPE1:	35.32	35.03	0.49	0.83
DPE2	35.20	36.83	0.31	0.85
DPE3	35.18	34.00	0.46	0.84
DPE4	35.42	33.19	0.66	0.82
DPE5	35.20	32.74	0.59	0.82
DPE6	35.37	31.29	0.69	0.81
DPE7	35.59	31.03	0.63	0.82
DPE8	35.36	32.54	0.64	0.82
DPE9	36.06	30.50	0.55	0.83
DPE10	35.81	32.77	0.43	0.84

*Rater leniency:* In the existing online system the students have to evaluate the performance of their group members according to 4 criteria: Cooperation, Conceptual Contributions, Practical Contributions, and Work Ethic. In each category there is a brief description of what the criteria measures so that the rater can distribute 50 points among the group members according to their performance in each dimension. There is also a section for students to write comments regarding qualitative aspects of their peers' performance, and 2 questions dealing with an overall evaluation of each peer according to his or her performance. These last two items ("Overall, I am satisfied with the performance of this team-mate in the group project" and "The performance of this team-mate has been, in general, excellent") that were measured with a Likert scale of 7 points, were 1=Strongly Agree and 7=Strongly Disagree, were used to estimate the tendency of the evaluation (leniency). The reliability of these two items was 0.96.

To measure leniency, three different variables were calculated. The first measure was named Minimum Evaluation. It consisted of the minimum evaluation given by a rater in both of the two items previously described to any of his/her team members. This measure therefore corresponds to the lowest evaluation given by a rater during the evaluation process. The second measure was called Average Evaluation and it corresponded to the evaluations given by a rater to every group member in both items that were later aggregated and then the average evaluation was calculated. In other words, this measure consists of the average evaluation of performance given by a rater to his/her team members. Finally, Evaluation Variance calculated the difference between the lowest and the highest evaluation given by a rater in both items was calculated. This measure reflects the span of the evaluation given by a rater within his/her group.

*Reactions to Feedback:* The online evaluation system included the scale Reactions to Feedback by Smither, Wohlers, & London (1995). This scale deals with rater's reactions to feedback regarding his/her own performance. Some of the items included in this questionnaire are: "I found that the feedback that I received was clear", "I found that the feedback that I received was applicable", and "I found that the feedback that I received was useful", that were measured by a 5 point Likert scale (1=Strongly Agree and 5=Strongly Disagree). The reliability of this scale was 0.92.

The students have to access the online evaluation system to review the evaluations regarding their own performance after every team member has completed the rating process. Only then the students are asked to fill out the Reactions to Feedback scale. Usually this happens at the very end of the semester when the course is over. Therefore, the number of students that actually complete this last questionnaire decreases considerably. In this case, only 14 students filled out the Reactions to Feedback scale, which means that the data regarding this scale is available for only 14% of the sample.

*Control variables:* Five control variables were included in this study, Age, Gender, GPA, Years in the Program and Impression Management. Age and Gender were controlled to ensure that hypothesized relationships between variables were common to both genders and to all ages, therefore eliminating alternative explanations for obtained results dealing with differences in demographics. As it can be seen on table 4, the sample was equally represented by both genders. Regarding age, the sample has a mean of 25 years, with the youngest student being 18 years old, and the oldest 42 years old (see Table 4).

Years in the Program and GPA were also controlled. In the case of Years in the Program, previous research has suggested that there may be a relationship between discomfort and experience. Smith, Harrington & Houghton (2000) argued that experience could affect communication reticence, which is positively related to discomfort. Therefore and to check for a direct influence of experience on discomfort, Years in the Program was controlled.

Regarding GPA or performance, this variable was controlled following the same logic as Years in the Program: it could be the case that discomfort or even leniency may vary according to the rater's own performance. Since previous research has determined that agreeableness and conscientiousness have an effect on leniency (Bernardin, Cooke & Villanova, 2000) it may be the case that other rater's personal characteristics may also have an effect on leniency and/or discomfort.

Table 4 shows the means and standard deviation for both Years in the Program and GPA. It is important to mention that this reported GPA varied from 1.77 to 4.03, while Years in the Program varied from 0.5 to 6, therefore showing a variance between respondents.

Social desirability in the form of impression management was included as a control variable. This variable was controlled because it has been argued that sometimes subjects tend to respond or react in a way that will make them look good, instead of answering truthfully (Paulhus, 1991). Asking students to reveal their cultural background or their level of discomfort with the evaluation of peers may be sensitive issues for some and although anonymity was assured, it could be the case that these scales created a need in some students to present themselves not as they are but as they think they are expected

to be. Therefore and in order to verify that the data was not tainted by social desirability, 12 items from the Balanced Inventory of Desirable Responding (BIDR) by Paulhus (1991) were included in the questionnaire. This scale is measured in a 5 point Likert scale (1=Strongly Agree and 5=Strongly Disagree), and includes items such as: “I have never dropped litter on the street” and “I have done things that I don’t tell other people about”. The reliability of this scale was 0.73.



## Results

Table 4 presents the correlation between the variables used in this study. It also shows the reliabilities for each scale and subscale used, as well as descriptive statistics for each variable. Among other relationships, these results reveal a positive relationship between collectivism and discomfort, as well as between individualism and discomfort (see Table 4). It can be noticed too that there is a positive relationship between discomfort and both average and minimum evaluation, while the relationship between discomfort and evaluation variance is negative (see Table 4).

Given that all three measures of leniency presented the exact same results, it was decided to present the results obtained with only one of the measures. In order to make the presentation of the results easier to read and understand, evaluation tendency will be conceptualized as the average evaluation given by the rater. It is important to mention that this has been the measure used in previous studies dealing with rater leniency (Villanova, Bernardin, Dahmus & Sims, 1993)

Self reported GPA, Age, Years in the program, Gender and Impression Management were the control variables in this study.

It is interesting to notice on Table 4 that the average level of discomfort was 3,94. Considering that “1” represented being comfortable with evaluating peers while “5” represented being extremely uncomfortable with the appraisal process, the average discomfort is quite high. Regardless of the cultural background, on average all raters were uncomfortable with rating a peer. Furthermore, the standard deviation is 0.63, which shows that the variance of the level of discomfort felt by the raters was relatively small.

Of the three subscales included in the discomfort measure (DPE), raters presented the highest average discomfort with collecting information for the appraisal (4.13).

It can be observed on Table 4 that most of the control variables were not significantly correlated to the variables in the model. One exception is the positive correlation between gender and the first subscale of the DPE. This correlation shows that men are uncomfortable with observing behaviour and collecting information that will be later used to evaluate performance. Women on the other hand do not show discomfort with this task of the peer evaluation system. A second exception is the negative correlation found between age and the level of individualism. This result suggests that older raters present lower levels of individualism.

**Table 4**  
**Correlations & Reliabilities**

	Mean	Std. Deviation	1	2	3	4	5	6	7	8	8a	8b	8c	9	10
1. Years	1.68	1.05	---												
2. GPA	3.57	4.17	0.05	---											
3. Age	24.90	5.84	0.29(**)	0.07	---										
4. Gender	0.46	0.50	0.04	-0.04	.068	---									
5. Impression Management	2.84	0.57	-0.15	-0.06	-0.16	0.14	0.73								
6. Collectivism	5.20	0.63	-0.09	0.06	0.02	-0.01	-0.14	0.70							
7. Individualism	5.19	0.71	-0.07	0.06	-0.21(*)	0.13	0.06	0.19	0.73						
8. DPE	3.94	0.63	0.12	0.08	-0.06	0.12	-0.09	0.23(*)	0.26(**)	0.84					
8a. Collecting Info	4.13	0.69	0.13	-0.03	0.00	0.21(*)	-0.10	0.00	0.27(**)	0.54(**)	0.83				
8b. Rating	4.10	0.72	0.11	0.06	-0.10	0.04	-0.17	0.30(**)	0.24(*)	0.90(**)	0.42(**)	0.77			
8c. Post Rating	3.69	0.85	0.08	0.11	-0.03	0.11	0.01	0.17	0.18	0.89(**)	0.24(*)	0.65(**)	0.75		
9. Average Evaluation	5.77	1.39	-0.14	-0.08	0.02	-0.05	-0.08	0.24(*)	0.05	0.28(**)	0.08	0.24(*)	0.28(**)	0.96	
10. Reactions to Feedback	2.80	0.85	0.05	-0.05	0.08	-0.46	0.25	0.17	0.00	-0.13	0.03	-0.13	-0.17	0.40	0.92

\*\* Correlation is significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

$\alpha$ : Cronbach's alpha

In order to test hypothesis 1a, a regression analysis was conducted. Hypothesis 1 proposed that the relationship between collectivism and discomfort with the peer evaluation system would be positive. The support for this hypothesis can be found on the second regression of Table 5, the regression between collectivism and discomfort. The significant increase in R square ( $\Delta R^2 = 0.05$ ,  $\beta=0.23$ ,  $p<.05$ ) between the regression with the control variables and the regression that includes collectivism supports the significance of this relationship. Furthermore, the beta obtained is significant and positive, suggesting that the higher the rater's collectivism, the higher the rater's discomfort with the evaluation process. Therefore, hypothesis 1a is supported.

**Table 5**

**Test for DPE as Mediator of the Collectivism-Leniency Relationship**

		Beta	R <sup>2</sup>	$\Delta R^2$ *
<b>Collectivism -&gt; Leniency: Average Evaluation</b>				
Model 1	Control		.04	.04
Model 2	Collectivism	.23*	.09	.05*
<b>Collectivism -&gt; Discomfort (DPE)</b>				
Model 1	Control		.06	.06
Model 2	Collectivism	.23*	.11	.05*
<b>Discomfort (DPE) -&gt; Leniency: Average Evaluation</b>				
Model 1	Control + Collectivism		.09	.09
Model 2	DPE	.29*	.16	.07*
<b>Collectivism and Discomfort (DPE) -&gt; Leniency: Average Evaluation</b>				
Model 1	Control + DPE		.14	.14
Model 2	Collectivism	.16	.16	.02

\* p<.05

\*\* p<.01

Hypothesis 1b proposed that there would be a negative relationship between individualism and discomfort with the peer evaluation. The second regression of Table 6 shows that even though the relationship between these two variables is significant, the direction not as expected. There is a significant increase in the R square ( $\Delta R^2 = 0.06$ ,  $\beta=0.24$ ,  $p<.05$ ) between the regression with the control variables and the regression that includes individualism, but the beta obtained is positive and not negative as hypothesized. This evidence suggests that the relationship between individualism and discomfort is significant and positive, suggesting that the higher the rater's individualism, the higher the rater's discomfort with the evaluation process. Therefore, hypothesis 1b is not supported.

**Table 6**

**Test for DPE as Mediator of the Individualism -Leniency Relationship**

		Beta	R <sup>2</sup>	$\Delta R^2$ *
<b>Individualism -&gt; Leniency: Average Evaluation</b>				
Model 1	Control		.04	.04
Model 2	Individualism	.07	.05	.01
<b>Individualism -&gt; Discomfort (DPE)</b>				
Model 1	Control		.06	.06
Model 2	Individualism	.24*	.12	.06*
<b>Discomfort (DPE) -&gt; Leniency: Average Evaluation</b>				
Model 1	Control + Individualism		.05	.05
Model 2	DPE	.33*	.15	.10*
<b>Individualism and Discomfort (DPE) -&gt; Leniency: Average Evaluation</b>				
Model 1	Control + DPE		.14	.14
Model 2	Individualism	-.01	.14	.00

\* p<.05  
 \*\* p<.01

As found by previous research, hypothesis 2 proposed that there would be positive relationship between discomfort and average evaluation. As it can be noticed on Table 7, there is a significant relationship between these two variables. The significant increase in R square ( $\Delta R^2 = 0.10$ ,  $\beta=0.33$ ,  $p<.05$ ) confirms that the results are similar to those found in previous research (Villanova, Bernardin, Dahmus & Sims, 1993; Tziner & Murphy, 1999; Bernardin & Villanova 2005). According to the beta obtained, this relationship is positive, which means that higher discomfort will result in higher average evaluation, thus supporting hypothesis 2.

**Table 7**

**Test for Relationship between Discomfort and Average Evaluation**

		Beta	R <sup>2</sup>	$\Delta R^2$ *
<b>Discomfort -&gt; Average Evaluation</b>				
Model 1	Control		.04	.04
Model 2	DPE	.33*	.14	.10*

\* p<.05  
 \*\* p<.01

In order to test the mediation hypotheses 3a and 3b, the mediation approach proposed by Baron and Kenny (1986) was used. This approach consists of 4 steps, each one requiring a different regression equation. Step one requires a regression equation to show that the independent variable affects the dependent variable. Step two consists of a regression equation to correlate the independent variable to the mediator. Step three tests the complete model by regressing the mediator (as the independent variable) to the dependent variable, while controlling for the independent variable. The fourth and final

step provides support for complete mediation and requires that the effect of the independent variable on the dependent variable to be zero when the mediator is controlled.

All four steps proposed by Baron and Kenny (1986) were calculated through hierarchical regressions. The test for discomfort as a mediator for the relationship between collectivism and average evaluation can be seen on Table 5. First, the relationship between collectivism and average evaluation is significant ( $\Delta R^2 = 0.05$ ,  $\beta=0.23$ ,  $p<.05$ ); second, there is a significant relationship between collectivism and discomfort ( $\Delta R^2 = 0.05$ ,  $\beta=0.23$ ,  $p<.05$ ); third, when collectivism is controlled, a significant relationship between discomfort and average evaluation can be appreciated ( $\Delta R^2 = 0.07$ ,  $\beta=0.29$ ,  $p<.05$ ); and finally the effect of collectivism on average evaluation when discomfort is controlled is not significant ( $\Delta R^2 = 0.02$ ,  $\beta=0.16$ , ns). Therefore, these results support the complete mediation of DPE in the collectivism – average evaluation relationship. Hypothesis 3a was supported.

Table 6 presents the same steps of the Baron and Kenny's procedure (1986) but this time with discomfort as the mediator of the relationship between individualism and average evaluation. In this case the mediation of discomfort was not supported because there was no support to the first step of the model: The relationship between individualism and average evaluation is not significant ( $\Delta R^2 = 0.01$ ,  $\beta=0.07$ , ns), as it can be seen on the first regression of Table 6. The other 3 steps of the Baron and Kenny's model (1986) are supported by the results: there is a significant relationship between individualism and discomfort ( $\Delta R^2 = 0.06$ ,  $\beta=0.24$ ,  $p<.05$ ); the relationship between discomfort and average evaluation is significant when individualism is controlled ( $\Delta R^2 =$

0.10,  $\beta=0.33$ ,  $p<.05$ ); and the effect of individualism on average evaluation when discomfort is controlled is not significant ( $\Delta R^2 = 0.00$ ,  $\beta= -0.01$ , ns). Since the first step is not supported it implies that discomfort is not a mediator of the relationship between individualism and average evaluation. Therefore, hypothesis 3b was not supported.

In order to investigate these unexpected results, exploratory analyses were calculated in order to analyse the influence of both collectivism and individualism on all three subscales of the DPE. Table 8 shows the results obtained for collectivism while Table 9 has the results for individualism.

**Table 8**

**Test for DPE Subscales Relationship with Collectivism**

		Beta	R <sup>2</sup>	$\Delta R^2$ *
<b>Collectivism -&gt; Collecting Information</b>				
Model 1	Control		.07	.07
Model 2	Collectivism	0	.07	.00
<b>Collectivism -&gt; Rating</b>				
Model 1	Control		.07	.07
Model 2	Collectivism	.30*	.16	<b>.09*</b>
<b>Collectivism -&gt; Post Rating</b>				
Model 1	Control		.03	.03
Model 2	Collectivism	.18*	.06	.03

\*  $p<.05$

\*\*  $p<.01$



**Table 9****Test for DPE Subscales Relationship with Individualism**

		Beta	R <sup>2</sup>	ΔR <sup>2</sup> *
<b>Individualism -&gt; Collecting Information</b>				
Model 1	Control		.07	.07
Model 2	Individualism	.26*	.14	.07*
<b>Individualism -&gt; Rating</b>				
Model 1	Control		.07	.07
Model 2	Individualism	.22*	.12	.05*
<b>Individualism -&gt; Post Rating</b>				
Model 1	Control		.03	.03
Model 2	Individualism	.16	.06	.03

\* p&lt;.05

\*\* p&lt;.01

While collectivism is only related to discomfort with rating ( $\Delta R^2 = 0.09$ ,  $\beta=0.30$ ,  $p<.05$ , see Table 8), individualism is related to both discomfort with collecting information ( $\Delta R^2 = 0.07$ ,  $\beta=0.26$ ,  $p<.05$ ) and rating ( $\Delta R^2 = 0.05$ ,  $\beta=0.22$ ,  $p<.05$ , see Table 9). It seems that even though both individualism and collectivism are related to discomfort with peer evaluation, the root or reason behind that discomfort is different. Further analysis is needed to solve this issue since the present data does not allow a more in depth analysis of the causes of discomfort.

Even though the sample size to test hypotheses 4, 5a and 5b was extremely small ( $n=14$ ), regressions were conducted as well to test these hypotheses. Results for hypothesis 4 can be found on Table 10. Results for the relationship between discomfort and reaction to feedback are not significant ( $\Delta R^2 = .01$ ,  $\beta= -0.14$ , ns), therefore, hypothesis 4 was not supported.

**Table 10**

**Test for Relationship between Discomfort and Reaction to Feedback**

		Beta	R <sup>2</sup>	ΔR <sup>2</sup> *
<b>Discomfort -&gt; Reaction to Feedback</b>				
Model 1	Control		.54	.54
Model 2	DPE	- .14	.53	.01

\* p<.05

\*\* p<.01

Tables 11 and 12 display the results obtained for mediation hypotheses 5a and 5b. The results show no support for either mediation model. The results do not support the mediation of discomfort for the relationship of either collectivism or individualism to reaction to feedback. Further analysis is needed in order to verify if these results are product of a reduced sample size or if there is no relationship between discomfort with the performance appraisal and discomfort with feedback on the rater's own performance. Hypotheses 5a and 5b were not supported.

**Table 11****Test for DPE as Mediator of the Collectivism-Reaction to Feedback Relationship**

	Beta	R <sup>2</sup>	ΔR <sup>2</sup> *
<b>Collectivism -&gt; Reaction to Feedback</b>			
Model 1 Control		.53	.53
Model 2 Collectivism	.36	.63	.10
<b>Collectivism -&gt; DPE</b>			
Model 1 Control		.06	.06
Model 2 Collectivism	.23*	.11	.05*
<b>DPE -&gt; Reaction to Feedback</b>			
Model 1 Control + Collectivism		.63	.63
Model 2 DPE	-.08	.64	.01
<b>Collectivism and DPE -&gt; Reaction to Feedback</b>			
Model 1 Control + DPE		.54	.54
Model 2 Collectivism	.35	.64	.10

\* p&lt;.05

\*\* p&lt;.01

**Table 12****Test for DPE as Mediator of the Individualism - Reaction to Feedback Relationship**

	Beta	R <sup>2</sup>	ΔR <sup>2</sup> *
<b>Individualism -&gt; Reaction to Feedback</b>			
Model 1 Control		.53	.53
Model 2 Individualism	.55	.65	.12
<b>Individualism -&gt; DPE</b>			
Model 1 Control		.06	.06
Model 2 Individualism	.24*	.11	.05*
<b>DPE -&gt; Reaction to Feedback</b>			
Model 1 Control + Individualism		.65	.65
Model 2 DPE	-.36	.71	.06
<b>Individualism and DPE -&gt; Reaction to Feedback</b>			
Model 1 Control + DPE		.54	.54
Model 2 Individualism	.72	.71	.17

\* p&lt;.05

\*\* p&lt;.01

An exploratory analysis was done to investigate which of the DPE subscales were related to average evaluation. It can be noticed on Table 13 that discomfort with collecting information to be used for the evaluation is not related to average evaluation. Discomfort with rating and post rating are both positively related to average evaluation.

**Table 13**

**Test for DPE subscales and Leniency Relationship**

		<b>DPE: Collecting Info</b>			<b>DPE: Rating</b>			<b>DPE: Post Rating</b>		
		Beta	R <sup>2</sup>	ΔR <sup>2</sup> *	Beta	R <sup>2</sup>	ΔR <sup>2</sup> *	Beta	R <sup>2</sup>	ΔR <sup>2</sup> *
<b>DPE -&gt; Average Evaluation</b>										
Model 1	Control		.04	.04		.04	.04		.04	.04
Model 2	DPE	.10	.05	.01	.28*	.11	.07*	.32*	.14	.10*

\* p<.05

\*\* p<.01

## **Discussion and Conclusion**

The results obtained confirmed previous findings that related discomfort with rater leniency (Villanova, Bernardin, Dahmus & Sims, 1993; Tziner & Murphy, 1999). It could be observed that higher levels of discomfort with the performance appraisal would result in leniency. The data also supported the hypothesis of the existence of a positive relationship between collectivism and discomfort. These results imply that raters with higher levels of collectivism will present higher levels of discomfort with the performance appraisal. These conclusions are especially relevant when joined. In fact, it was found that discomfort mediated the relationship between collectivism and evaluation tendency. Therefore, the collectivism level of a rater has in fact a tangible impact on performance appraisals. Higher levels of collectivism are related to higher levels of discomfort with peer appraisal, which are related to rater leniency. This evidence suggests that exporting multi-source feedback practices to collectivistic cultures may undermine the validity of the whole evaluation system, by introducing leniency and therefore tainting 'real' evaluation scores. These results could be the first step towards more research dealing with the effects of cultural background on performance appraisal and other commonly used human resources management tools.

The positive relationship found between individualism and discomfort was unexpected, for it shows that both individualistic and collectivistic raters are uncomfortable with rating peers. Although it was not in the scope of this research, it can be hypothesized that these findings are due to the fact that discomfort is different for individualistic and collectivistic raters for it relates to different aspects of the evaluation

process. In other words, it is possible that individualism relates to a discomfort with the actual process of evaluating (such as observing behaviour, accessing the online peer evaluation system, dedicating time to write feedback), while collectivism is related to discomfort with the social implications of rating a peer (for example, assigning ratings that oppose peer's expectations, evaluating performance regardless of personal like or dislike for the evaluated peer). On one hand, the discomfort experienced by individualistic raters could come from considering the expenditure of time and dedication on evaluating peers as an annoyance and aggravation, since it is a task that is probably not considered as part of their role or responsibility. On the other hand, collectivistic raters may perceive that the social implications of evaluating a peer are too great and important, therefore creating a discomfort with the appraisal process. This line of thought follows the conclusions reached by Fried, Tieggs and Bellamy (1992), who found that raters with high discomfort would choose not to perform the appraisal at all when given the option. Fried et al. (1992) did not analyse the cultural background of the raters, so more research is needed to clarify this aspect, but it could be possible that although the causes of discomfort differ between these groups, discomfort is present on both collectivistic and individualistic raters.

The presented theory is correct and the causes of discomfort are in fact different for collectivistic and individualistic raters, it could explain the results obtained for hypotheses 3a and 3b. As mentioned earlier, the mediation of discomfort was supported for the relationship between collectivism and evaluation tendency, but it was not supported for individualism and evaluation tendency. Although both collectivistic and individualistic raters are uncomfortable with rating peers, only collectivistic raters allow

this discomfort to have an impact on the evaluations. In fact even though the individualistic raters show discomfort with the evaluation, there is no relationship between this discomfort and the given ratings. In other words, individualistic evaluators do not seem to allow discomfort to influence their judgement, while collectivistic raters allow discomfort to affect their given rating through leniency. These results could suggest that while it is not particularly relevant to reduce the discomfort of individualistic raters, decreasing the discomfort of collectivistic raters is vital for the validity of a peer appraisal.

Unfortunately the present study could not reach valid conclusions for any of the 3 hypotheses that dealt with reaction to feedback. The small sample size obtained for this variable made it impossible to verify if hypotheses 4, 5a and 5b were not supported because of a real lack of relationship between the variables or because the sample was too small to detect any relationships. Further research is needed in order to clarify this issue.

### Implications for Practice

The findings of the present study suggest that special attention should be given to reduce the discomfort experienced by collectivistic raters, since the discomfort felt by this group of raters will impact the ratings through leniency. It seems very relevant to try to keep the discomfort level as low as possible through the facilitation of the rating process. Collectivistic raters are uncomfortable with assigning ratings, evaluating peers and providing written feedback, among other actions included in the evaluation process. Therefore, it could prove effective to try to decrease the level of discomfort with this

process by simplifying the rating form itself, in order to make the evaluation as smooth as possible. This could mean that in collectivistic countries evaluation forms should be very simple and should refrain from complicated measures (such as forced rankings) and mandatory written feedback. In this study, the evaluation form used a forced-ranking process in which raters had to distribute a total amount of points among group members. It may be that collectivistic raters feel that this is a tough and difficult task. Therefore, a way to reduce discomfort could be through evaluation forms that use Likert scales or similar measures instead of systems where it is required to assign ratings.

Another possibility is that even though raters were assured anonymity, they might have felt that in small working groups assuring anonymity is more of a good intention than a reality. Therefore, in order for collectivistic raters to feel more comfortable with evaluating peers, it could be useful to not only ensure anonymity before the rating process begins but also during and after the evaluation. The appraisal system used in this study allowed ratees to have direct access to their ratings, with no filter or intermediary. Raters may have felt more at ease with the appraisal if their evaluations had reached the ratee through a third person, such as a supervisor. In work settings, the use of an intermediary such as a HR specialist may reduce the discomfort with the rating aspect of the appraisal system, reducing rater leniency.

A third suggestion would be to inform raters about the purpose of the appraisal. In this study, peer ratings were incorporated into the ratees' final grade. Therefore, raters should have perceived this appraisal system as having evaluative purposes. Considering that McEvoy and Buller (1987) found that attitudes towards peer appraisal were more favourable when the evaluations were used for developmental purposes instead of having



an evaluative function. It may be the case that in order to reduce leniency and discomfort, peer ratings should be collected only for developmental reasons.

Finally, previous research has suggested that both leniency and discomfort with the appraisal are significantly reduced when raters are trained. Bernardin & Villanova (2005) provided Self-Efficacy Training to raters (SET-R) and discovered that after the training was completed, leniency was considerably reduced, as well as rater discomfort. Most of the students that participated in the present study had no previous experience with peer evaluations. Furthermore, students only received a brief explanation from their professors on how to use the appraisal system at the beginning of the semester, with no further training detailed or otherwise. Communicating the importance of accurate ratings and training raters in the appraisal process could help reduce the discomfort that raters feel with the process and therefore decrease rater leniency, making ratings more accurate and useful for every human resources strategy.

### Limitations

This research is not without limitations. The first issue is related to the sample since it was composed of undergraduate students. It is unknown whether the results obtained would be the same if the sample had been composed of full time employees. The main issue could be rooted on the characteristics that differentiate the working relationships and conditions of students and employees. The work relationship for students ends once the semester is over while for employees, the work relationship does not have a particular expiration date. It may be possible to find higher levels of

discomfort on an employee sample, since giving a bad rating to a peer could imply an ongoing source of social problems for the rater.

A second limitation deals the low response rate obtained for the measure dealing with rater's reaction to feedback. Since only 14 students completed this section of the peer evaluation system, it was impossible to reach any solid conclusions for hypotheses 4, 5a and 5b. Future studies need to address this issue in order to get a higher response rate that would allow hypotheses testing.

Third, the use of self report could imply that the results are tainted by social desirability. Several steps were taken in order to minimize this possibility. The students had to answer to separate questionnaires at different points in time. This time gap and the fact that students were not given any details on the hypotheses of this study, imply that there are no particular reasons to believe that the students would modify their responses in any particular way. Furthermore, social desirability was measured through the impression management scale and the results show no reason to suspect that the results could be altered by this variable.

A fourth limitation of this research corresponds to its design. Since this is a correlational study, it is impossible to determine causality or the direction of the relationship between variables. The hypotheses presented are based on the idea that individualism/collectivism is inherent to the person, rooted on education and upbringing. It is difficult to conceive that discomfort with the evaluation could cause individualism or collectivism. Yet, because of the correlational nature of this study, causality and directionality cannot be determined.

Finally, same source variance could be a threat to the conclusions reached by this study. To diminish this danger, two different questionnaires were filled by the students at two different points in time, therefore following the suggestion by Ostroff et al. (2002, pg. 366) that "...researchers can reduce response bias associated with common method variance by incorporating time delays between the measurement of independent and dependent variables." Furthermore, previous research has downplayed the importance of same source variance. Keeping and Levy (2000, pg 721) concluded that "...our study, along with these other studies, which examine the role of common method variance in separate research areas, arrive at a very similar conclusion: common method variance exists, but at low and usually inconsequential levels."

### Conclusions

Individual characteristics have a bigger impact on daily activities than we would like to believe. It is only natural, then, that these personal characteristics also have an effect on occupational activities. Therefore, it is reasonable to expect cultural background to also have an impact on commonly used human resources practices. Previous research has determined that cultural background in general, and individualism/collectivism in particular, has an impact on some human resources tools. For example, cultural background would have an affect on providing feedback. Ogawa and Welden (1972) found evidence to support the hypothesis that collectivistic societies provide less explicit feedback than individualistic societies, where non verbal communication is less relevant than in collectivistic cultures. Furthermore, there may be a difference between what

motivates individualistic and collectivistic workers. Trompenaars and Hampden-Turner (1998) suggested that individualistic workers will prefer tangible rewards (such as extra money and shares), while collectivistic workers will be motivated through intangible rewards (recognition and advancement possibilities, for example).

This research is the first step to identify that individualism/collectivism has an impact on rater discomfort and therefore on leniency. This implies that performance appraisals are directly affected by the cultural background of the rater, for ratings are altered and conclusions obtained are tainted by factors outside an employee's objective work performance.

Although this study has its limitations, it is a contribution to the rater discomfort literature and to the research on multicultural management. Future research could explore more in depth the causes of discomfort in order to determine with more precision which particular areas of the performance appraisal create discomfort for raters. Only then more specific recommendations could be made for managers that deal with multicultural workforces and that would like not only to improve the validity of the appraisal but also to make the evaluation process more agreeable for both raters and ratees. Performance evaluations are important and should be kept as a human resource management tool, but there is a lot of space for improvement for this tool and its implementation.

## References

Albright Funderburg, S. & Levy, P.E. (1997). The influence of individual and contextual variables on 360-degree feedback system attitudes. *Group & Organization Studies*. Vol. 22, Issue 2, pg. 210.

Atwater, L. & Brett, J. (2006). Feedback format: Does it influence manager's reactions to feedback? *Journal of Occupational and Organizational Psychology*. Vol. 79, Issue 4, pg. 517.

Atwater, L. & Brett, J. (2005). Antecedents and consequences of reactions to developmental 360° feedback. *Journal of Vocational Behaviour*. Vol. 66, Issue 3, pg. 532.

Austin, J.T. & Villanova, P. (1992). The criterion problem: 1917-1992. *Journal of Applied Psychology*, Vol 77, Issue 6, pg. 836.

Banks, C.G. & Murphy, K.R. (1985). Toward narrowing the research-practice gap in performance appraisal. *Personnel Psychology*. Vol. 38, Issue 2, pg. 335.

Baron, R.M. & Kenny, D.A. (1986). The moderator-mediator variables distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*. Vol 51, Issue 6, pg. 1173.

Bernardin, J.H. (1987). Development and validation of a forced choice scale to measure job related discomfort among customer service representatives. *Academy of Management Journal*. Vol 30, Issue 1, pg. 162.

Bernardin, J.H., Cooke, D.K. & Villanova, P. (2000). Conscientiousness and agreeableness as predictors of rating leniency. *Journal of Applied Psychology*. Vol 85, Issue 2, pg 232.

Bernardin, J.H. & Orban, J.A. (1990). Leniency effect as a function of rating format, purpose for appraisal and rater individual differences. *Journal of Business and Psychology*. Vol. 5, Issue 5, pg. 197.

Bernardin, J.H. & Villanova, P. (2005). Research streams in rater self-efficacy. *Group & Organization Management*. Vol. 30, Issue 1, pg. 61.

Bernthal, P., Sumlin, R., Davis, P. & Rogers, B. (1997). *Performance Management Practices Survey Report*. Pittsburgh, PA. Development Decisions International.

Bettenhausen, K.L. & Fedor, D.B. (1997). Peer and upward appraisals. A comparison of their benefits and problems. *Group & Organization Studies*. Vol. 22, Issue 2, pg. 236.

Brett, J.M., Tinsley, C.H., Janssens, M., Barsness, Z.I. & Lytle, A.L. (1997). New approaches to the study of culture in industrial/organizational psychology. *New Perspectives on International Industrial/Organizational Psychology*. Earley, P.C. & Erez, M. (Eds.) Pfeiffer.

Bretz, R.D., Milkovich, G.T. & Read, W. (1992). The current state of performance appraisal research and practice: Concerns, directions and implications. *Journal of Management*. Vol. 18, Issue 2, pg. 321.

Brochner, J., De Cremer, D., van den Bos, K. & Chen, Y.R. (2005). The influence of interdependent self-construal on procedural fairness effects. *Organizational Behaviour and Human Decision Processes*. Vol 96, pg. 155.

Cawley, B.D., Keeping, L.M. & Levy, P.E. (1998). Participation in the performance appraisal process and employee reactions: A meta-analytic review of field investigations. *Journal of Applied Psychology*. Vol. 83, Issue 4, pg. 615.

Chirkov, V., Ryan, R.M., Kim, Y., & Kaplan, U. (2003). Differentiating autonomy from individualism and independence: A self-determination theory perspective on internalization of cultural orientations and well-being. *Journal of Personality and Social Psychology*. Vol. 84, Issue 1, pg. 97.

Conway, J.M., Lombardo, K. & Sanders, K.C. (2001). A meta-analysis of incremental validity and nomological networks for subordinate and peer ratings. *Human Performance*. Vol. 14, Issue 4, pg. 267.

Costigan, R.D., Insinga, R.C., Kranas, G., Ilter S.S., Kurechov, V.A. & Berman, J.J. (2005). A study of Russian, Polish, and U.S. self- and peer raters: moderators of rating agreement. *International Journal of Commerce & Management*. Vol. 15, Issue 3/4, pg. 205.

Dipboye, R.L. & de Pontbriand, R. (1981). Correlates of employee reaction to performance appraisals and appraisal systems. *Journal of Applied Psychology*. Vol. 66, Issue 2, pg. 248.

Drexler, J.A., Beehr, T.A. & Stetz, T.A. (2001). Peer appraisals: Differentiation of individual performance on group tasks. *Human Resource Management*. Vol. 40, Issue 4, pg. 333.

Dooley, R. (2003). Four cultures, one company: Achieving corporate excellence through working cultural complexity. *Organization Development Journal*. Vol. 21, Issue 2, pg. 52.

Earley, C.P. & Stubblebine, P. (1989). Intercultural assessment of performance feedback. *Group & Organization Studies*. Vol. 14, Issue 2, pg. 161.

Fletcher, C. (1999). The implications of research on gender differences in self-assessment and 360-degree appraisal. *Human Resource Management*. Vol. 9, Issue 1, pg. 39.

Fletcher, C. & Perry, E.L. (2001). Chapter 7: Performance appraisal and feedback: A consideration of national culture and a review of contemporary research and future trends. *Handbook of Industrial, Work and Organizational Psychology: Volume 1: Personnel Psychology*. Sage Publications. London.

Fried, Y., Tieg, R.B., & Bellamy, A.R. (1992). Personal and interpersonal predictors of supervisors' avoidance of evaluating subordinates. *Journal of Applied Psychology*. Vol. 77, Issue 4, pg. 462.

Furnham, A. & Stringfield, P. (1998). Congruence in job performance ratings: A study of 360-degree feedback examining self, manager, peers and consultant ratings. *Human Relations*. Vol. 51, Issue 4, pg. 517.

Gillespie, T. (2005). Internationalizing 360-degree feedback: Are subordinate ratings comparable? *Journal of Business and Psychology*. Vol. 19, Issue 3, pg. 361.

Gilliland, S.W. & Langdon, J.C. (1998). Creating performance management systems that promote perceptions of fairness. *Performance Appraisal: State of the Art in Practice*. Smither, J.W. (Ed) Jossey-Bass, San Francisco, California.



Harris, M.M. (1994). Rater motivation in the performance appraisal context: a theoretical framework. *Journal of Management*. Vol. 20, Issue 4, pg. 737.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Sage Publications. Beverly Hills, California, USA.

Hofstede, G. (1983). The cultural relativity of organizational practices and theories. *Journal of International Business Studies*. Vol. 14, Issue 2, pg. 75.

Hofstede, G. (1984). The cultural relativity of the quality of life concept. *The Academy of Management Review*. Vol. 9, Issue 3, pg. 389.

Hofstede, G. (2002a). Dimensions do not exist: A reply to Brendan McSweeney. *Human Relations*. Vol. 55, Issue 11, pg. 1355.

Hofstede, G. (2002b). Cross-cultural differences in Central Europe. *Journal of Managerial Psychology*. Vol. 18, Issue 1/2, pg. 76.

Hofstede, G., Van Deusen, C.A., Mueller, C.B. & Charles, T.A. (2002). What goals do business leaders pursue? A study in fifteen countries. *Journal of International Business Studies*. Vol. 33, Issue 4, pg. 785.

Huffcutt, A.I. & Roth, P.L. (1998). Racial group differences in employment interview evaluations. *Journal of Applied Psychology*. Vol. 83, Issue 2, pg. 179

Kane, J.S., Bernardin H.J., Villanova, P., & Peyrefitte, J. (1995). Stability of rater leniency: Three studies. *Academy of Management Journal*. Vol. 38, Issue 4, pg. 1036.

Keeping L.M. & Levy, P.E. (2000). Performance appraisal reactions: Measurement, modeling, and method bias. *Journal of Applied Psychology*. Vol. 85, Issue 5, pg. 708.

Kingstrom, P.O. & Mainstone, L.E. (1985). An investigation of the rater-ratee acquaintance and rater bias. *Academy of Management Journal*. Vol 28, Issue 3, pg. 641.

Leslie, J., Gryskiewicz, N. & Dalton, M. (1998). Understanding cultural influences on the 360-degree feedback process. *Maximizing the value of 360-degree feedback: A process for successful individual and organizational development*. Tornow, W. and London, M. Jossey-Bass. San Francisco. California.

Levy, P.E. & Williams, J.R. (2004). The social context of performance appraisal: A review and framework for the future. *Journal of Management*. Vol. 30, Issue 6, pg. 881.

Lobel, S. (1990). Global leadership competencies: Managing to a different drumbeat. *Human Resource Management*. Vol. 29, Issue 1, pg. 39.

London, M. & Beatty, R.W. (1993). 360-degree feedback as a competitive advantage. *Human Resource Management*. Vol. 32, Issue 2-3, pg. 353.

London, M. & Smither, J.W. (1995). Can multi-source feedback change perception of goal accomplishment, self-evaluation and performance-related outcomes? Theory-based applications and directions for research. *Personnel Psychology*. Vol. 48, Issue 4, pg. 803.

Longenecker, C.O., Sims, H.P. & Gioia, D.A. (1987). Behind the mask: The politics of employee appraisal. *The Academy of Management Executive*. Vol. 1, Issue 3, pg. 183.

Luthans, K.W. & Farner, S. (2002). Expatriate development: The use of 360-degree feedback. *The Journal of Management Development*. Vol. 21, Issue 9/10, pg. 780.

Malos, S. (1998). Current legal issues in performance appraisal. In Smither, J.W. (Ed.) *Performance appraisal: State of the art and practice*. Jossey-Bass. San Francisco.

McEvoy, G.M & Buller, P.F. (1987). User acceptance of peer appraisals in an industrial setting. *Personnel Psychology*. Vol. 40, Issue 4, pg. 785.

McSweeney, B. (2002a). Hofstede's model of national cultural differences and their consequences: A triumph of faith – a failure of analysis. *Human Relations*. Vol. 55, Issue 1, pg. 89.

McSweeney, B. (2002b). The essentials of scholarship: A reply to Geert Hofstede. *Human Relations*. Vol. 55, Issue 11, pg. 1363.

Mendonca, M. & Kanungo, R.N. (1996). Impact of culture on performance management in developing countries. *International Journal of Manpower*. Vol. 17, Issue 4/5, pg. 65.

Milliman, J., Nason, S., Gallagher, E., Huo, P., Von Glinow, M.A. & Lowe, K.B. (1998). The impact of national cultural on human resource management practices: The case of performance appraisal. *Advances in International Comparative Management*. Vol. 13, pg. 157.

Milliman, J., Taylor, S. & Czaplewski (2002). Cross-cultural performance feedback in multinational enterprises: Opportunity for organizational learning. *Human Resource Planning*. Vol. 25, Issue 3, pg. 29.

Murphy, K.R. & Cleveland, J.N. (1991). *Performance appraisal: an organizational perspective*. Boston. Allyn & Bacon. Chapter 5: Obtaining information about performance.

Napier, N.K & Latham, G.P. (1986). Outcome expectancies of people who conduct performance appraisals. *Personnel Psychology*. Vol 39, Issue 4, pg. 827.

Ogawa, D.M. & Welden, T.A. (1972). Cross-cultural analysis of feedback behaviour within Japanese American and Caucasian American small groups. *The Journal of Communication*. Vol. 22, Issue 2, pg. 189.

Ostroff, C., Kinicki, A.J. & Clark, M.A. (2002). Substantive and operational issues of response bias across levels of analysis: An example of climate–satisfaction relationships. *Journal of Applied Psychology*. Vol. 87, Issue 2, pg. 355.

Oyserman, D., Coon, H.M. & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*. Vol 128, Issue 1, pg. 3.

Paulhus, D.L. (1991). Measurement and control of response bias. *Measures of Personality and Social Psychological Attitudes*. Robinson, J.P., Shaver, P. & Wrightsman, L.S. (Eds). San Diego, Academic Press.

Pearce, J.L. & Porter, L.W. (1986). Employees responses to formal performance appraisal feedback. *Journal of Applied Psychology*. Vol. 71, Issue 2, pg. 211.

Porter, S.R. (2004). Raising response rates: What works? *New Directions for Institutional Research*. No. 121, pg. 5.

Robert, C., Lee, W.C. & Chan, K.Y. (2006). An empirical analysis of measurement equivalence with the INDCOL measure of individualism and collectivism: implications for valid cross-cultural inference. *Personnel Psychology*. Vol. 59, pg. 65.

Russell, J.S. & Goode, D.L. (1988). An analysis of manager's reactions to their own performance appraisal feedback. *Journal of Applied Psychology*. Vol 73, Issue 1, pg. 63.

Sackett, P.R. & Wilk, S.L. (1994). Within-group norming and other forms of score adjustment in preemployment testing. *American Psychologist*. Vol. 49, Issue 11, pg. 929.

Seifert, C.F., Yukl, G. & McDonald, R.A. (2003). Effects of multisource feedback and feedback facilitator on the influence behaviour of managers toward subordinates. *Journal of Applied Psychology*. Vol. 88, Issue 3, pg. 561.

Shrauger, J., & Rosenberg, S. (1970). Self-esteem and the effects of success and failure feedback on performance. *Journal of Personality*. Vol. 38, Issue 3, pg. 404.

Singelis, T.M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*. Vol. 20, Issue 5, pg. 580.

Singelis, T., Triandis, H., Bhuvuk, D. & Gelfand, M. (1995). Horizontal and vertical dimensions of individualism and collectivism: A theoretical and measurement refinement. *Cross-Cultural Research*. Vol. 29, Issue 3, pg. 240.

Smith, W.J., Harrington, K.V., & Houghton, J.D. (2000). Predictors of performance appraisal discomfort. *Public Personnel Management*. Vol. 29, Issue 1, pg. 21.

Smith, R. E., & Sarason, I. G. (1975). Social anxiety and evaluation of negative interpersonal feedback. *Journal of Consulting and Clinical Psychology*. Vol. 43, Issue 3, pg. 429.

Smither, J.W., London, M. & Richmond, K.R. (2005). The relationship between leaders' personality and their reactions to and use of multisource feedback. *Group & Organization Management*. Vol. 30, Issue 2, pg. 181.

Smither, J.W., Wohlers, A.J. & London, M. (1995). A field study of reactions to normative versus individualized upward feedback. *Group and Organization Management*. Vol 20, Issue 1, pg. 61.

Sivadas, E., Bruvold, N.T. & Nelson, M.R. (2008). A reduced version of the horizontal and vertical individualism and collectivism scale: A four-country assessment. *Journal of Business Research*. Vol. 61, Issue 3, pg. 201.

Stull, J.B. (1988). Giving feedback to foreign-born employees. *Management Solutions*. Vol. 33, Issue 7, pg. 42.

Sully de Luque, M.F. & Sommer, S. (2000). The impact of culture on feedback seeking behaviour: An integrated model and propositions. *The Academy of Management Review*. Vol. 25, Issue 4, pg. 829.

Triandis, H.C. (1989). The self and social behaviour in differing cultural contexts. *Psychological Review*. Vol. 96, Issue 3, pg. 506.

Trompenaars, F. & Hampden-Turner, C. (1998). *Riding the Waves of Culture*. McGraw-Hill, Quebecor / Fairfield.

Tziner, A., Murphy, K.R. & Cleveland, J.N. (2005). Contextual and rater factors affecting rating behaviour. *Group & Organization Management*. Vol. 30, Issue 1, pg. 89.

Tziner, A. & Murphy, K.R.(1999). Additional evidence of attitudinal influences in performance appraisal. *Journal of Business and Psychology*. Vol. 13, Issue 3, pg. 407.

Villanova, P., Bernardin, H.J., Dahmus, S.A. & Sims, R.L. (1993). Rater leniency and performance appraisal discomfort. *Educational and psychological measurement*. Vol 53, Issue 3, pg. 789.

Waldman, D.A. & Bowen, D.E. (1998). The acceptability of 360 degree appraisals: A customer-supply relationship perspective. *Human Resource Management*. Vol. 37, Issue 2, pg. 117.

Westerman, J.W. & Rosse, J.G. (1997). Reducing the threat of rater nonparticipation in 360-degree feedback systems. An exploratory examination of antecedents to participation in upward ratings. *Group & Organization Studies*. Vol. 22, Issue 2, pg. 288.

Wexley, K.N. & Youtz, M.A. (1985). Rater beliefs about others: Their effects on rating errors and rater accuracy. *Journal of Occupational Psychology*. Vol. 58, Issue 4, pg. 265.

Whitley, B.E. (2001). *Principles of Research in Behavioral Science*. McGraw Hill, NY. Second edition.

## Appendix 1

### DPE

**When evaluating the performance of your group members or peers, how comfortable do you feel...**

**ND=No discomfort and HD=High discomfort**

Collecting information of your peers' performance to assign accurate ratings

ND    O    O    O    O    O    HD

Using and trusting your observations to assign ratings

ND    O    O    O    O    O    HD

Evaluating peer's performance independent of your personal like or dislike for that person

ND    O    O    O    O    O    HD

Assigning ratings that are accurate but that you know may disagree with your peer's expectations

ND    O    O    O    O    O    HD

Distributing points among your peers according to their performance

ND    O    O    O    O    O    HD

Providing written feedback or comments regarding peer's performance

ND    O    O    O    O    O    HD

Talking to a peer about the evaluation you gave him/her

ND    O    O    O    O    O    HD

Telling a peer how his/her performance can improve if he/she asks for your advice

ND    O    O    O    O    O    HD

In future courses, being in the same work group with a peer whose performance you evaluated as below average

ND    O    O    O    O    O    HD

Developing a friendship or social relationship with a peer whose performance you evaluated as below average

ND    O    O    O    O    O    HD

- ◆ Subscale Collecting Info: Items 1 & 2
- ◆ Subscale Rating: Items 3, 4, 5 & 6
- ◆ Subscale Post Rating: Items 7, 8, 9 & 10