

Preference for Appraisal Format: How Individual Differences Shape Perceptions of Ratings and
Comments

Andrew Watson

A Thesis

In

John Molson School of Business

Presented in Partial Fulfillment of the Requirements

For the Degree Master of Science (Administration) at

Concordia University

Montreal, Quebec, Canada

August, 2014

© Andrew Watson, 2014

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Andrew Watson

Entitled: Preference for Appraisal Format: How Individual Differences Shape
Perceptions of Rating and Comments

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Administration)

complies with the regulations of the University and meets the accepted standards with
respect to originality and quality.

Signed by the final examining committee:

_____ Chair
Satyaveer Chauhan

_____ Examiner
Dr. Devasheesh Bhave

_____ Examiner
Dr. Yu-Ping Chen

_____ Supervisor
Dr. Stephane Brutus

Approved by _____
Chair of Department or Graduate Program Director

Dean of Faculty

Date _____

ABSTRACT

Preference for Appraisal Format: How Individual Differences

Shape Perceptions of Ratings and Comments

Andrew Watson

The current study explores the differences between the two most commonly used appraisal formats, ratings and comments. Lending both from literature on performance appraisals, and individual differences, this study looks at an appraisee's perceptions of how useful each performance appraisal format is dependent on the appraisee's individual differences (Goal orientation, Tenure). A total of 188 university professors took part in this study. Participants completed an online questionnaire that dealt with perceptions of their past teacher evaluations, and their individual characteristics. Regression analysis was used to explore the relationship between different individual differences and the perceptions the participants had of how useful ratings and comments were. The findings conclude that overall university professors find both comment and ratings as useful, with comments being more useful than ratings. It was also found that as tenure increases, comments will decrease in perceived usefulness faster than ratings. Practical implications and future research avenues are discussed.

Acknowledgements

I would like to begin by thanking my supervisor Dr. Stéphane Brutus, the completion of this project would not have been possible without his time, patience, support, and guidance, throughout the entire process. I would also like to thank the members of my thesis committee, Dr. Devasheesh Bhave and Dr. Yu-Ping Chen for their insight and help. Last but not least I would like to thank my family for their support and encouragement.

Table of Contents

Introduction.....	1
Purpose of Performance Appraisal.....	2
Formats: Ratings and Comments.....	4
Social Context and Appraisee Reaction.....	7
Teacher Evaluations.....	8
Role of Personality in Organizations.....	10
Summary of Literature Review.....	13
Ratings versus Comments.....	13
Goal Orientation.....	15
Tenure.....	18
Valence.....	20
Methodology.....	22
Procedure.....	22
Sample.....	23
Measures.....	23
Data Preparation.....	25
Analytical Strategy and Results.....	28
Discussion.....	37
Ratings versus Comments.....	37
Goal Orientation.....	38
Tenure.....	40
Valence.....	41
Practical implications.....	42

Limitations.....	44
Future Research.....	46
Conclusion.....	48
References.....	49

List of Tables

Table 1: Measures Descriptives.....	27
Table 2: Correlations.....	27
Table 3: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of ratings.....	29
Table 4: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of comments.....	29
Table 5: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of ratings.....	30
Table 6: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of comments.....	31
Table 7: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of ratings.....	32
Table 8: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of comments.....	32
Table 9: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of ratings.....	33
Table 10: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of comments.....	34
Table 11: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of ratings.....	35
Table 12: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of comments.....	36

List of Appendices

Email Asking for Participation.....	57
Questionnaire.....	58
Consent Form.....	65
Follow Up Email.....	67

Preference for Appraisal Format: How Individual Differences Shape Perceptions of Ratings and Comments

Organizations have recognized the usefulness of performance appraisals and they are widely used in most organizations today. Research has found that the proper execution of a performance appraisal system can have a positive effect on a variety of employee attitudes and behaviors (Gilliand & Langdon, 1998; Dusterhoff, Cunningham, Macgregor, 2014; Jacobs, Belshak, Hartog, 2014). For example, perceived fairness of a performance appraisal system has been found to positively correlate with satisfaction with the performance appraisal session, organizational commitment, and the motivation to perform well (Steensma & Visser, 2007). Overall performance appraisals play a key role in the health of an organization as they can have a direct effect on attitudes and behaviors of employees. Therefore, it is important to understand the factors within these appraisal systems that will lead to positive employee reactions and to a healthy organization.

As Fletcher (2001) explains “performance appraisal has become a general heading for a variety of activities through which organizations seek to assess employees and develop their competence, enhance performance, and distribute rewards.” Overall, the key of a performance appraisal is to assess an employee’s performance. This process requires two participants, the rater, and the ratee. The rater is the person who is in charge of observing and assessing the ratee’s performance. On the other hand, the ratee is the person who performs and is assessed by the rater. All in all, if a performance appraisal process is done properly it can give the ratee a good understanding of their performance, their strengths, their weaknesses, and how they can improve.

The current project focuses on the different formats typically used in performance appraisals. These two formats are ratings and comments. Ratings are usually numerical representations of a ratee’s performance, while comments are written feedback that describes and gives suggestions on how to improve a ratee’s performance. Both ratings and comments are usually used together on most performance appraisals. However, as researchers have pointed out, there has been an over emphasis on ratings and not enough attention paid to comments, (Brutus, 2009). This means we have very little understanding of the role comments play in creating a good appraisal process. The purpose of this project is to explore and attempt to identify the

positive components of each ratings and comments, to understand the usefulness of each format, and to understand why both are usually used together on performance appraisals.

Another objective of the current project focuses on the reactions of the ratee's to the performance appraisal process. While a proper appraisal system can bear many rewards, research has identified that an absolute perfect appraisal system does not exist (Levy & Williams, 2004). The success of an appraisal system is dependent on many situational factors. Over the past fifteen years research has begun to recognize the importance of social context in the performance appraisal process (Levy et al, 2004). Among the topics of interest in this field of study is how a ratee's reaction to a given appraisal process can influence the effectiveness of this process. It has been argued that both structural and procedural components of a given appraisal system will have a direct effect on a ratee's performance rating, reactions (behavioral, cognitive, attitudinal), and perceptions of justice (Gilliand et al, 1998). For example, the ratee tends to find a given appraisal as more just when they are invited to participate in the appraisal process. Research in performance appraisal has moved away from just trying to create the most objective and accurate appraisal system, and has started to recognize important factors such as development, ownership, perceptions of being valued and so forth. Overall, research has moved into a direction of trying to figure out what are the components at plays that will lead to having a ratee have positive reactions to the appraisal process. These positive reactions are beneficial as they have been shown to be positively linked to organizational commitment, and an employee's motivation to perform well (Steensma et al. 2007). More specifically this project attempts to identify how individual differences or personality traits may interact with the appraisal process to create certain outcomes.

In summary, the purpose of this project is to clarify two issues. First to understand when each format, ratings, or comments is a useful tool to use, and second, how personality and individual differences influence the ratee's reaction to the appraisal process. This project takes a further step to understand the relationship between personality and individual differences, and how they might interact with the different appraisal formats.

Purpose of Performance Appraisal

One important factor for both the rater and ratee to keep in mind during the appraisal process is the goal or purpose of the appraisal. Historically, the purpose of performance appraisals was to rank employees (Austin et al. 1992). This is useful in situations where

management have to make decisions regarding selection, promotion or termination. In these situations, performance appraisals are used to compare employees to identify who is best for the job.

However as discussed by Ryan and Sackett (1992), using performance appraisal for development purposes has become increasingly popular, accounting for at least 40% of the performance appraisals used in organizations. The purpose of development is to enact some kind of behavioral intervention, diagnosing what someone is doing wrong, and trying to correct the behavior to a more favorable one. The ideas for developmental feedback have grown out of many different behavioral change theories.

Using performance appraisals for assessment purposes such as selection, promotion, or termination usually involves giving feedback that creates a picture of how an individual may have performed. On the other hand, using performance appraisals for developmental purposes goes beyond this; not only do developmental appraisals state how a ratee performed, it also gives information about how their performance stack up to the ideal performance, and offers help to better performance.

The conception and ideas behind performance appraisals used for developmental purposes are rooted in theories of behavioral change. Many theories exist and they tend to focus on how behavioral change has an effect on the individual, the environment, or both. Below, we will discuss a few of behavioral change theories, what role they play in a performance appraisal setting, and how personality or individual differences might interact with these theories and the resulting behavioral changes.

The first theory of behavioral change that influenced the creation of developmental feedback is the theory of Reasoned Action (Fishbein & Ajzen, 1975). This theory focuses on an individual's attitudes and beliefs. The main argument of this theory is that the intention to act is the best predictor of action or behavior, where intentions can be defined as "motivation factors that influence behavior" (Ajzen, 1991). The theory argues that intentions are directly influenced by an individual's attitudes about desired behaviors and their beliefs about outcomes. Overall, intention has been found to be the most important factor when it comes to behavioral change, (Godin & Kok, 1996). This theory argues that in order to develop an employee's skills, the rater needs to focus on creating positive attitudes in the ratee as well as providing favorable outcomes. As will be discuss later, personality can influence what a person deems as a favorable outcome.

For example some personality traits might influence an individual to deem performing better than other people as a favorable outcome; in turn this can lead people to favor certain feedback as they compare performance more easily. Raters need to recognize that not all outcomes are equal to everybody, and they need to take into account individual personalities. Moreover, the theory states that intentions can be influenced by individual subjective norms, such as, what they believe others believe they should do. Overall, behavioral change is centered on the individual, as individuals can differ greatly in many ways, providing feedback needs to account for these individual differences.

Another theory of behavioral change that influenced developmental feedback is the Control Theory. The main argument of this theory is that individuals change their behavior because they wish to close a gap between their current selves and their desired selves. Carver and Sheier (1982) describe this theory as a negative feedback loop where people are constantly reviewing their current self and receiving negative feedback because they do not match up to the ideal self. This negative feedback pushes an individual to change their behavior in order to try to obtain the desired state. Developmental performance appraisals can be used to state to the person where they are, where they should be, and offer ways to alter behavior to reach the ideal state. Once again, individual differences and personality traits can influence how useful a person finds feedback. People will have different ideal states dependent on many individual differences such as culture, gender, or age.

In summary, performance appraisal can be used for many different assessment reasons such as, selection, promotion, termination, and development. As people have individual differences and different personality traits, they may react in different ways to feedback; therefore we need to build a greater understanding of how individual differences influence reactions to feedback in order to create appraisals that will be effective.

Formats: Ratings and Comments

One of the most important structural elements of an appraisal system is the format. Scholars have begun to question how the role of format, mainly quantitative ratings or qualitative narrative comments, come into play with performance appraisal systems (Brutus, 2009). Ratings are calculated by using items measured using scales. For example, the rater will rate how well the ratee performed on a scale from one to five. Ratings are used as a normative measure in order to rank all the ratee's in order. Narrative comments, on the other hand, consist of written words.

These comments allow for more detailed feedback and can be used for developmental purposes. For narrative comments, the rater is asked to give comments on either specific performance measures, (which will result in specific feedback), or on the ratee's general performance, (which will result in general feedback). Some scholars have begun to argue that there has been too much attention given to quantitative approaches (Pearce & Porter, 1986; Brutus, 2009) which in turn has left little research that has looked at qualitative approaches. While both formats often appear together on performance appraisals, the performance appraisal research has mostly focused on ratings. This favoritism towards ratings has risen out of many factors.

The main reason for ratings being favored would simply be the history of performance appraisal; and its relations to the history of personnel psychology (Landy & Farr, 1980). Towards the beginning of the twentieth century a new economical order arose. This was characterized by a shift from an agricultural economy towards a manufacturing economy. This resulted in an entirely new set of skills that was required of workers. The problem that arose from this was that we did not know who was right for which job. Influenced by intelligence tests such as the Binet scale, personnel psychologists looked for ways to rank workers in order to create more order to this problem (Austin & Villanova, 1992). Therefore, performance appraisals were first used to rank employees, which was facilitated by ratings. While ratings are not the only way to rank employees, it was the preferred method for many reasons. The first reason for using ratings was measurement precision which can be easily translated into making administrative decisions (Brutus, 2009). Ratings have also been favored as they are efficient and can be completed relatively quickly; basically they translate into substantial economies for raters (Murphy & Cleveland, 1995).

Although research in performance appraisal has been dominated by quantitative formats, research has begun to explore the utility of qualitative formats. This move towards new explorations can be attributed to three factors. The first factor is that technological advances have made it easier to process qualitative information. This has had good implications for doing research surrounding qualitative data as it has made it faster and easier. Computer software can now be used to code and sort through qualitative data, and what would have taken weeks of manual sorting, can now be done in a matter of hours. The second reason for exploring qualitative formats is that researchers have begun to acknowledge the limits of ratings and have begun to search for alternatives. This factor is measurement based. Performance is usually

related to work context (Ferris, Munyon, Baslik, and Buckley, 2008; Judge and Ferris, 1993), and ratings fall short of capturing the complex nature of the work context. It is believed that qualitative appraisal formats would be better equipped to deal with work context as they can offer deeper information about evaluations (Ferris et al, 2008). Overall, the rising interest in qualitative formats in appraisal research is in line with the rising interest in organizational research that has been looking more and more at contextual approaches, (Cappelli & Scherer, 1991). The third reason for exploring qualitative formats is that evidence shows that qualitative information alters how recipients perceive performance information (Brutus, 2009). For Example, qualitative feedback such as comments can provide more information that will help a ratee interpret ratings, (Dalessio, 1998). Other research has shown that people are more likely to pay attention to comments than to ratings, (Fersti & Bruskiwicz, 2000). Overall, when comparing ratings and narrative comments it is argued that narrative comments convey a more personal focus and will lead to stronger reactions from recipients, will lead to less social comparison than ratings, and can have a substantial effect on behavior (Brutus, 2009).

The types of format (quantitative and qualitative) both have their own specific benefits. Very often both formats are used together; however research still remains scarce on qualitative formats. Research has shown that attributes of a performance appraisal system can influence an appraisee's perception of the appraisal process, meaning their perceived validity and acceptance of information given and their perceived utility of the process. For example, some have argued that in the presence of the feeling of unfairness in process and in equity in evaluations, any appraisal system will be doomed to failure (Cardy & Dobbins, 1994). Performance appraisal systems needs to be structured properly in order for them to be accepted by the appraisee and therefore meet the goals it has set out to do, whether it be for administrative or developmental purposes. This leads to the question: How do different performance appraisal formats (qualitative or quantitative) affect appraisee perceptions of validity and utility of the appraisal process? In line with Levy et al. (2004) the appraisal context is inherently a social context; we need to take into consideration people's individual differences. Therefore, it is argued that people's individual differences will affect whether or not people find ratings or narrative comments more valid and useful. Certain individual differences will lead to the preference of ratings, while other individual differences will lead to the preference of narrative comments.

Social Context and Appraisee Reactions

While a proper appraisal system can bear many rewards, research has identified that an absolute perfect appraisal system does not exist. The success of an appraisal system is dependent on many situational factors. Over the past fifteen years research has begun to recognize the importance of social context in the performance appraisal process (Levy & Williams, 2004). Among the topics of interest in this field of study is how a ratee's reaction to a given appraisal process can influence the effectiveness of this process. It has been argued that both structural and procedural components of a given appraisal system will have a direct effect on a ratee's performance rating, reactions (behavioral, cognitive, attitudinal), and perceptions of justice. For example, the ratee tends to find a given appraisal as more just when they are invited to participate in the appraisal process (Gilliland & Langdon, 1998). Research in performance appraisal has moved away from trying to create the most objective and accurate appraisal system, and has started to recognize important factors such as development, ownership, perceptions of being valued, and so forth. Overall, research has moved into a direction of trying to figure out what are the components at play that will lead to having a ratee have positive reactions to the appraisal process. These positive reactions are beneficial as they have been shown to be positively linked to organizational commitment, and an employee's motivation to perform well (Steensma et al. 2007).

In line with the shift towards the study of social context in performance appraisal research, many theories have been proposed to explain appraisee reactions to a performance appraisal. Pichler (2012) summarizes a few of these theories. These theoretical models all rest on the assumption that there are three different main components that will influence a ratee's reacting to an appraisal. These three components are the rater-ratee relationship, appraisal participation, and performance rating favourability. While all three models recognize the components as factors that influence appraisal reactions, the theories differ in how these components interact.

The first model, "an independent model of appraisal reactions," views the three components as all independent factors. This model is highly related to the procedural justice theory, where employee reactions are related to their participation in the appraisal process. Participation in the appraisal process usually results in more favorable reactions (Greenberg, 1990). The three main component mentioned above allow employees to have some control over

the process, and allows them to add their voice to the process which in turn increases perceptions of fairness in the process (Folger, 1977; Lind & Tyler, 1988).

The second model, “An instrumental, exchange oriented model of appraisal reactions,” view the three components as dependent, where the rater-ratee relationship influences both appraisal participation, and performance rating favourability. In this model, the ratee is not interested in creating a fair appraisal process, rather they use their relationship with the rater to try and influence the process and steer it towards giving more favorable outcomes to the ratee, which in turn leads to better appraisal reactions.

The third model, “relationship quality-driven model of the appraisal process,” is similar to model two as it views rater-rate relationship at the forefront and having influence over appraisal participation and performance rating favourability. However this model is not driven by trying to influence the process towards certain outcomes. This model is based in the group value model of procedural justice (Lind et al. 1988), where having one’s voice heard signifies a group members status. Basically, being viewed as a valued member of a team will make the ratee feel good and will in turn result in positive appraisal reactions.

Overall, the social context plays an important role in the performance appraisal process and how individuals react to the process. As Herold and Fedor (1998) discuss, there is an interaction between the context and the players within the context. While early appraisal research viewed the participants in the appraisal process as passive, more recent research has recognized that this is not necessarily the case. In other words, there has been a shift from studying just situational variables to studying individual variables. This recent research argues that the ratee is very active in the appraisal process. In many cases the purpose of the appraisal process is to give feedback to the ratee. The new paradigm suggests that the ratee has influence over what feedback they receive as they are in part responsible for creating their own feedback environment (information available about one’s performance). Individuals do this by the way they seek out, process, and respond to feedback.

Teacher Evaluations

Teacher evaluations are a very unique and controversial performance appraisal context. The basic idea behind these evaluations is that students grade professors on their teaching abilities. These evaluations are then used for two purposes, professors take this feedback and use it to make changes to their courses (Beran, Violato, Kline, & Frideres, 2005), or these

evaluations are used for administrative purposes by department heads, (Wrights & Jenkins-Guarnieri, 2011). Teacher evaluations are not the typical appraisal context that you will find in most organizations where a superior evaluates a subordinate. Rather it is more of a customer feedback context. Many organizations have stepped away from the traditional appraisal context and have started using the 360-degree feedback method where individuals receive feedback from many different sources, one of these being customers (Ghorpade, 2000). While customer feedback is used mostly for developmental purposes, this feedback is still considered in administrative decisions (Rogers, Rogers, & Metlay, 2002). The teacher evaluation context is good for this study as it focuses on the type of feedback individuals prefer, and not the feedback source. As well both ratings and comments are given in the teacher evaluation context.

Although this type of appraisal is used in most higher education institutions, the validity and utility of student evaluations of teaching are a constant source of conflict (Algozzine, Beattie, Flowers, Gretes, Howley, Mohanty, & Spooner, 2004). Attitudes about teaching evaluations range from total acceptance to outright dismissal (Nasser & Fresko, 2002). The outright dismissal of teaching evaluations is very surprising considering the evidence that supports them. Wright and Jenkins-Guarnieri (2011) conducted a thorough meta-analysis of the validity and susceptibility to bias of teacher evaluations. They drew the conclusion that teacher evaluations can be used as useful tools for both professors and administrators, as these appraisals were found to be related to student achievement and were found to be a valid measure of an instructors skills and teaching effectiveness.

If teaching evaluations are valid, why do some professors still dismiss them? Tuytens and Devos (2013) identified three main factors that influence whether or not a professor finds the teaching evaluations they receive as useful or not. The first is their supervisor. It was found that if a given professor had a supervisor who was a charismatic leader, the professor would perceive there to be more procedural justice in the appraisal process, and would be more likely to view teacher evaluations as useful. A charismatic leader motivates and encourages a professor more than non-charismatic leaders. This support builds a special relationship between the leader and professor which is built on trust. If a professor trusts a charismatic leader more, they will then trust that the evaluations the charismatic writes is more fair, accurate, and useful. The second factor was whether or not the professor took part in the appraisal process. If they did take part, they would be more likely to accept the teacher evaluations as useful and valid; taking part on

the appraisal process means taking part in the designing and set up of teacher evaluation procedures (Cawley, Keeping, and Levy, 1998). The third factor that has an effect on professor's perceptions of teacher evaluations is individual differences. It was found that individual characteristics such as job experience and willingness to collaborate influenced perceptions of teacher evaluations. Penny and Coe (2004) suggest that some professors can have difficulties making sense of teacher evaluations, and that a professor would view these appraisals as more useful if they had a consultation session with a supervisor to help make sense of the teacher evaluations.

Overall, teacher evaluations have been shown to be valid appraisal tools. The reasons that not all professors accept these appraisals are due to personal perceptions of the appraisal process. These personal perceptions can be related to perceptions of procedural justice, or due to personal characteristics. In order to increase the acceptance of teacher evaluations, research needs to further explore what can influence how a professor interacts with the teacher evaluations process. The purpose of this study will be to understand how professors perceive different appraisal formats, as well as how their individual differences might influence these perceptions.

Role of Personality in Organizations

In line with the area of research dealing with social context, scholars have recognized that there is value in looking at the role that personality and individual differences may play in human resource practices. A person's personality is the emotional, attitudinal, and behavior response patterns that the person engages in different circumstances. Individual differences on the other hand deal with characteristics of an individual, such as gender or age that may influence the individual's responses. Both personality and individual differences have been extensively researched, exploring how they may affect people's behaviors within an organization, their reactions to feedback, and their reactions towards different performance appraisal formats and structures.

Much of the management literature has explored how personality may affect many different organizational behaviors. For example, research has found that locus of control may influence an individual's job satisfaction and organizational commitment. Locus of control is a personality trait which refers "to the extent to which people believe them or external factors such as chance and powerful others are in control of the events that influence their lives" (Firth et al. 2004). Research has found that people with an internal locus of control, who feel that they are in

charge of their own lives, tend to have high levels of job satisfaction and organizational commitment (Srivastava, 2013; Chhabra, 2013). Similarly other research has explored how individual differences may influence different organizational variables. For example research has found that individual difference, such as education level and how many children a person has can influence an employee's readiness to deal with change or their organizational commitment. Those who have a higher education and more children were more likely to be ready to deal with change and have high organizational commitment compared to those with low education levels and less children, (Abrow & Abrishamkar, 2013). In sum, personality traits and individual differences have been shown to influence people's organizational behaviors.

It has also been argued that personality may have an effect on the way an individual interacts with a performance appraisal system. Ogunfowora, Bourdage and Lee (2009) looked at how different personality traits may have an effect on the weight and importance individuals give to different performance dimensions. They looked at how individual levels of extraversion, agreeableness, conscientiousness, and openness to experience would affect an individual's value of different performance dimensions. In their study Ogunfowora, et al. (2009) concluded that "rater personality does significantly influence the weight individual raters place on certain performance dimensions." For example, they found that people who scored high on the personality trait openness to experience tend to give more weight to adaptive performance than those low in openness to experience. Overall personality can influence the way that people interact with performance appraisal systems, influencing their focus of different performance dimensions, and what they find important.

Research has also explored how personality and individual differences interact with how people deal with feedback. For example, some research has found that an individual's self esteem will impact how they seek out feedback; people may avoid receiving feedback if they fear that it will damage their self-esteem (Ashford, 1989). Similarly, Vande Walle and Cummings (1997) found that goal orientation affects how an individual seeks out feedback. Overall, individuals who have learning goals tended to be more likely to seek performance feedback. Other research has looked at how people might respond to feedback. For example, research has found that older people are less likely than younger people to change responses when faced with negative feedback (McMurtie, Baxter, Obonsawin, & Hunter, 2012). As well, several researchers have found that the personality trait narcissism can influence how people

react to feedback. Those who scored high in the personality trait narcissism were found to show more aggression after receiving negative feedback, (Martinez, Zeichner, Reidy, & Miller, 2007; Barry, Chaplin, and Grafeman, 2006). Overall, research has shown that an individual's personality or individual differences can influence the way they interact with feedback, mainly the way they seek out, approach and avoid feedback, and the way they react to feedback. Therefore, it is important to continue forward exploring more ways how the individual influences the organization.

While research in the area of social contexts looks at how procedural and structural components influence the ratee, research in personality explores how people may react differently depending on their unique individual differences. Although the context might increase the likelihood that ratees will react in a given way, a person's individual preference may interact with this effect and alter it. That being said, it is possible that given different personality traits, a ratee may have a preference for different procedures and structures than another ratee may have. Different individual differences may influence the way that an individual seeks out, processes, or responds to feedback. Individuals high in public self consciousness desire more feedback than people who are low in public self consciousness. This will obviously have an effect on how a ratee seeks out feedback, and what type of feedback they receive (Levy, Albright, Cawley & Williams, 1995).

As we have seen above, the structure of a performance appraisal system is critical in its success. As far as social context goes, the structure of an appraisal system can influence the reactions of the ratee. However, it is important to keep in mind that not every ratee will react in the same way to a given structure. Individual differences may lead ratees to have a preference for one appraisal structure over another. This may influence how useful a ratee may find a given appraisal system. Much research has been dedicated to examining the effects of a ratee's view of a performance appraisal system. For example, Swiercz, Bryan, Eagle, Bizzotto, & Renn (2012) argued that the effectiveness of a performance appraisal system relies on an employee's reaction to the performance appraisal system. In their research they found that positive views of a performance appraisal system will lead to higher employee performance, organizational commitment, supervisory satisfaction, job satisfaction, and pay satisfaction. As Gilliland et al. (1998) argued the structure of a performance appraisal system will affect how fair an employee

may view the process. In sum, how a ratee feels about the process is important as it will lead to higher performance and job satisfaction.

Summary of the Literature Review

Overall, performance appraisals are essential to running a successful organization. There has been a vast amount of research that would support this statement, and insight into how performance appraisals work. While the literature provides good insight, many areas still need to be explored. The literature has moved its focus towards the social context, however, there is information lacking about how individual differences and personality may influence a ratee's reaction to the appraisal process. As well, there has been an over emphasis on ratings in the performance appraisal literature, and comments have not been researched enough.

The theoretical appraisal literature argues that the individual is the key player in performance appraisals. It is the individual's beliefs, attitudes, and behaviors that influence the reactions of a ratee. Moreover the research has shown that individual differences and personality traits can influence the way people interact with the performance appraisal process, for example people may seek out feedback in different ways dependent on their goal orientation.

Ratings and comments differ in many ways in their uses. Overall, ratings are more normative and provide information that is easy to compare to previous performance, or the performance of others. Comments can provide prescriptive feedback that gives information to the ratee about how they might change their behavior in order to increase their performance. While it is known that both formats differ, there has been an over emphasis on ratings in the research literature, and little is known about the reactions each format illicit from individuals being rated.

The purpose of this study will be to explore individual reactions to the different formats to try and determine if each format illicit different responses. Furthermore this study will look at the role of personality in the performance appraisal process. Overall it will be expected that not only will both formats differ in how useful the ratee's find them, Individual differences and personality traits will influence whether an individual finds one format as more useful over the other.

Ratings Versus Comments

Above it was discussed that ratings and comments are usually both delivered together on a performance appraisal. While they are two formats that are largely used together, they each offer

their own unique information as to how someone may have performed on a job. The main question we need to ask is what are the unique information each format delivers? While this paper is interested in individual differences and perceived usefulness of each format, we need to ask what if individual differences are not considered. If we compare the pros and cons of each format, can we predict that one format will be viewed as more useful than the other?

A rating is a form of feedback that is quantitative. A typical rating on a performance appraisal will be a number given to an individual that for example will tell the person how well they performed on a scale from one to five. The biggest benefit to this performance appraisal format is that it is normative, which means that a rating can be easily compared to other ratings. For example we can compare an individual's performance to their previous performance and determine if they have improved or gotten worse. Furthermore it also permits us to compare a person's rating to another individual's rating. This is ideal in situations where a performance appraisal is used for administrative purposes such as raises, promotions and bonuses. Since we can easily compare the performances between two individuals and determine who performed better.

Comments are a form of feedback that is qualitative. They are narrative words that describe how an individual performed, and often will give prescriptive advice on how the individual can mold their behavior in order to perform better next time. Comments have many benefits, for example they can be more specific and give personalized feedback to an individual. While comments are difficult to use for administrative purposes where one must compare multiple employees, comments are often used for developmental purposes. Since comments can provide clear feedback as to what an individual needs to change in order to perform better, they can be used in situation where an evaluator is attempting to develop and advance an employee.

As can be seen both ratings and comments have their own strengths. Ratings are normative and are usually useful when it comes to administrative purposes. On the other hand comments are more specific and give personalized information to individuals. Feys et al. (2011) discuss how there has been a lack of research on the topic of an appraisee's reaction to different appraisal formats, to their knowledge only one study has looked directly this topic. That study showed the appraisee's to have more positive reactions to numeric formats than to written text formats (Atwater & Brett, 2006). However, their paper did not go into details about how they operationalized each format. Other research has gone into more detail about the components of

each format and how they affect the appraisee. For example, it has been found that more specific feedback can have a positive effect on the appraisee's performance (Goodman & Wood, 2004). This demonstrates that more specific feedback can be utilized by the appraisee in order to increase performance. It can then be assumed that more specific feedback would be viewed as more useful. Since ratings fail to take context into account, it is predicted that comments are viewed as more useful information to the appraisee as comments provide more specific information such as the what, where, and when of job performance.

Hypothesis 1: Professors will perceive receiving comments as more useful than receiving ratings.

Goal Orientation

As Cron, Slocum, Vande Walle, and Fu (2005) discuss goal orientation has become a highly researched theory which looks to explain how an individual's personal goal preference in achievement situations impact many outcomes. Individuals can differ in the type of goal they set for themselves; they can either have learning goals or performance goals. Those who adopt a learning goal view their abilities as malleable and seek to learn from each situation, incrementally improving their ability. Individuals who adopt a learning goal view success as the result of effort put into a task. On the other hand those who choose performance goals view ability as fixed, innate, and difficult to modify. They also tend to view success as the result of having high ability. As Potosky (2010) explains some research has even split performance goals into two different categories. The first is labelled Approach Performance Goals, where individuals attempt to outperform others. The second is Labelled Avoidance Performance Goals, where individuals attempt to avoid looking incompetent. A learning goal orientation has been shown to be positively related to performance outcomes (e.g., Phillips & Gully, 1997; Button, Mathieu, & Zajac, 1996; Martocchio & Hertenstein, 2003). There appears to be an agreement that the level in goal orientation an individual has, whether it be learning goal or performance goal is a relatively stable trait (Dweck, 1986). This means that individuals who have a certain level of a given goal orientation is likely to have the same level over different situations. Although early research saw goal orientation as dichotomous, more recent research argues that people can actually have varying levels of both goal orientations.

In performance management settings, the importance of goal orientation is recognized. As Seijts and Latham (2005) discuss goal setting is often used in the work place to motivate

employees. However, it can be used in very different ways depending on an individual's goal orientation. For those with a performance goal orientation, managers can give concrete objectives, For example, they may be told to increase sales by ten percent. In contrast, a learning goal orientation takes the focus off of the goal. In many cases a learning goal orientation will be considered during the hiring process, where managers will hire individuals based on aptitude rather than skills they possess. From there these employees are given mentors and developed. Therefore, they very often end up performing better than those with performance goals. Overall, managers recognize that goal orientation is an important individual difference which can impact outcomes. However, further research needs to be done in order to explore how an individual's goal orientation may interact within a performance appraisal process.

The type of goal orientation that an individual falls under also influence the type of feedback that they seeks out (Park, Schmidt, Scheu, & DeShon, 2007). The literature on feedback identifies three main types of feedback in which people can seek out. The first is normative feedback; this allows people to use social comparison in order to estimate their own ability compared to others (Butler, 2000). The second type of feedback is called diagnostic feedback. This provides corrective information by identifying what a person has done wrong and providing suggestions as how to correct whatever faulty behavior an individual has engaged in (Moreno & Mayer, 2005). The third type of feedback in which people seek out is called self-enhancing feedback. This is when individuals seek out feedback that they know will provide positive information to enhance their self esteem (Swann, Pelham, & Krull, 1989). Overall it has been argued that people who have high learning orientation perceive high value in seeking out feedback. They use feedback as a tool to gauge and increase their competencies (Vande Walle et al. 1997). On the other hand, people with performance goals seek feedback to know how well they are doing compared to others as they are preoccupied with competing and outperforming others (Park et al. 2007). These people see ability as fixed, and therefore do not look at feedback as a tool for developing one's ability (Parks et al. 2007).

People differ in their goal orientations, and these differences are relatively stable traits. As Chen, Gully, Whiteman, and Kilcullen (2000) explain, trait-like individual differences such as goal orientation and other personality characteristics are not specific to a certain task or situation and are stable over time. In contrast, state-like individual differences are specific to certain tasks and tend to be more malleable over time. The level of the different types of goal orientation an

individual has will influence the type of feedback that this individual will seek out. As it has been found those with high learning goal orientation will seek out feedback that will help them increase their competencies (Vande Walle, 1997). This will rely on feedback that gives prescriptive information on how to improve performance. On the other hand, individuals with a high level performance goal orientation seek out feedback that allows them to compare themselves to others, allowing them to gauge how they performed relative to others (Parks et al. 2007). This will rely on feedback that is normative, and allows for ranking of individuals. The type of performance appraisal format will influence what kind of feedback can be attained. It is predicted that given certain goal orientations, individuals will have a preference for one format over another as this format will give the individual the feedback that they require to meet their needs.

According to Parks et al. (2007) one of the types of feedback that can be sought out is normative feedback, and this type of feedback is usually desired by those with a performance goal orientation. Normative feedback is directly related to ratings, which is an appraisal format designed to give normative feedback. On the other hand comments are less likely to provide normative information that can easily allow for social comparison. Therefore it is predicted that people with a performance goal orientation will find ratings as more a more useful format than comments as those with a performance goal orientation seek out feedback information that is normative (Parks et al. 2007).

Another important type of feedback information that people seek out is diagnostic feedback, which provides specific information about peoples behaviors and how they can correct behaviors in order to increase performance (Parks et al. 2007). This type of feedback is typically sought out by people with a learning goal orientation, (Vande Walle, 1997). Diagnostic feedback is directly related to comments, which is an appraisal format that is designed to provide specific information about an appraisee's performance and give diagnostic feedback about how to improve (Brutus, 2009). Ratings on the other hand do not provide diagnostic information, and therefore would be seen as less useful for people with a learning goal orientation. It is predicted then, that people with a learning goal orientation would find comments as a more useful format than ratings, as comments provide the diagnostic feedback they seek out.

Hypothesis 2a: Performance goal orientation will be more positively related to the perceived usefulness of ratings than to the perceived usefulness of narrative comments.

Hypothesis 2b: Learning goal orientation will be more positively related to the perceived usefulness of narrative comments than to the perceived usefulness of ratings.

Tenure

Tenure is an individual difference characteristic which refers to how long someone has been working at a job. People with low tenure are relatively new to an organization, while people with high tenure have been working at an organization for a long period of time. Ferris and King (1992) argue that the age and tenure of an individual working in an organization are very important individual differences that need to be explored and better understood. Historically there have been legal issues dealing with age and tenure, where organizational decisions such as selection, promotions, or terminations have been directly linked to age and tenure. Research dealing with tenure in organizations brings up some interesting questions. For example, Miller (1980) found that 65% of contracts reviewed considered tenure while making promotion decisions, however also found that there is little research that supports a link between tenure and performance. Research looking at tenure and performance has had mixed results, and no clear link between the two variables has been found (Gordon & Johnson 1982; Jacobs, Hofmann, & Kriska 1990) Overall tenure is an individual difference which is often considered in the performance appraisal process, especially in situations that warrant selection, promotion, or termination. However there is no clear link between tenure and performance. We need to further explore the impact of tenure on the performance appraisal process. One purpose of this study is to explore the ratee's perception of the performance appraisal process, and test to see if an individual's tenure plays a role in how they perceive the process.

As mentioned above, people being evaluated are active in creating their feedback environment. People's individual differences may influence the type of information a person will receive as feedback, as they seek out, process and react in different ways to feedback. The purpose of this research is to determine whether people seek out ratings or comments to receive the most useful feedback for their own use. Levy et al. (1995) do a good job of explaining what feedback seeking is, and explain that there are three main motives for feedback seeking. The first motive is to reduce uncertainty. The argument for this motive is that if people are unable to assess themselves on how well they are doing on a job, or they are uncertain about how they are doing, they will then seek feedback from others to help reduce this uncertainty. The second motive is to protect one's ego. The argument for this motive is that people will seek feedback

when they know it will be good and boost their self esteem; the motive is to feel good about one self. The third motive is for impression management purposes. In this motive individuals will seek out feedback when it will make them look good in others eyes, however will avoid it if the feedback will make them look worse in other people's eyes.

Feedback seeking explains the motives that people might have for seeking out feedback, however not everyone will be influenced by every motive equally. Individual difference may influence which motives might influence and individual. One individual difference that has been shown to influence feedback seeking is tenure. Tenure is the number of years someone has been working at a job. Morrison (1993) offers some insight into how tenure is related to feedback seeking. She explains that people just starting at a job have high uncertainty of what they should be doing, since they do not hold the information needed to master a given job. Therefore people with low tenure are preoccupied with both trying to fit in and mastering their job. Seeking out feedback will ameliorate these problems by offering information of how they compare to others (normative feedback) in the organization as well as prescriptive suggestions (narrative feedback). As Morrison (1993) goes on to explain, as people stay within an organization longer they will be less concerned with fitting in, focusing more on just mastering their job. Therefore normative information becomes less important while narrative information remains to be relatively important. As time progresses and a person remains at a job for a long time this person will seek out feedback much less. This is because people who have been at a job have come to a point where they feel that they have mastered the job. The literature suggests that people who have been at a job a long time seek feedback less as it will undermine their status as the senior employees at a company. Therefore people with high tenure become motivated by ego protection and impression management and therefore do not seek out as much feedback.

As has it had been discussed, there are multiple motives for seeking feedback. These motives can be influenced by tenure. Individuals will have different motives for seeking feedback depending on how long they have worked at a job. As different motives may require different information, different motives may require ratings while others narrative comments. Therefore how long someone has worked at a job may be linked to their preference for either ratings or narrative comments.

It has been argued that individuals who are new to a job need both normative information to fit in as well as narrative information on how to master their job. Normative information can

be provided through ratings. On the other hand narrative units can provide prescriptive information which will give information to an employee about how to shape their behavior in order to master a job.

It has also been argued that as tenure increases, an employee will become more comfortable with their environment, and will have less of a need to fit in. Therefore the use of normative information will become less important. This normative information comes from ratings. However these employees still need to master their job and will rely on narrative units to give them information on how to properly perform their job.

Lastly it has been argued that people with high tenure may avoid seeking feedback. An individual with high tenure may feel that they should have mastered their job. They would expect normative ratings to be as high as possible, low scores may hurt their ego, therefore they might reject the information to protect their ego. As well, since they feel they have mastered their job they no longer seek prescriptive feedback that they would receive from narrative ratings. Therefore they might reject any narrative feedback that they receive.

Overall the relationship between feedback seeking and tenure is that feedback seeking will become less important the longer someone has worked at a job. In other words low tenure can be linked to perceiving both ratings and narrative comments as useful. On the other hand high tenure can be linked to perceiving both ratings and narrative comments as not being useful. More specifically the perceived usefulness will decrease faster for ratings than for narrative comments.

Hypothesis 3a: As tenure increases the perceived usefulness of both ratings and narrative comments will decrease for professors.

Hypothesis 3b: As tenure increases the perceived usefulness of ratings will decrease faster than the perceived usefulness of narrative comments for professors.

Valence

One important variable that is discussed in the performance appraisal literature is valence. Valence refers to the degree that information in a given appraisal is positive or negative. Both ratings and narrative units can be positive or negative. As Feys, Anseel, and Wille (2011) discuss whether a rating or comment is positive or negative can be defined very simply. First of all it is important to note that valence is something subjective, and is defined by the person who is receiving the feedback, and how they interpret it. Basically individuals develop standards as to what an appropriate performance level is. These standards are built from company standards,

personal abilities, and personal values. The standard is basically the goal of how well the person feels they should perform. When it comes to being evaluated, people who reach or surpass their standards of performance will view the evaluation as positive. On the other hand those who fall short of their standards will be disappointed with their evaluation, and will view the evaluation as negative. Essentially valence of a performance appraisal is subjective and based on personal expectations. When a person meets their personal standards, they will perceive the feedback as positive, however when the person does not meet their personal standards they will view the feedback as negative.

In turn the ratee's perception as to how positive or negative their evaluation was has a direct effect on their reactions to the performance appraisal process. More specifically, it will have an effect on whether or not the ratee views the information as useful or not. For example, as outlined by Feys et al. (2011) the valence of an appraisal influences the reaction of the appraisee. Positive feedback is usually accepted by the feedback recipient, and negative feedback is often rejected. When a ratee receives negative feedback, they often attribute it to an external factor out of their own control. This ties in directly with the vast literature that deals with fairness in the appraisal process. The ratee needs to view the process as fair, with one of the largest determinants of whether or not they will find the process fair being if the evaluation was based on relevant information (Gilliand et al. 1998). If the ratee is attributing negative feedback to external factors out of their control they will not find the process as fair and negative reactions will be the result. For example the perception of unfair appraisals have been linked to low levels of job satisfaction and the intention to quit, (Poon, 2004).

Research suggests that different appraisal formats will be preferred dependent on certain individual differences. The literature on valence indicates that even if someone prefers a given format, they may view the feedback as the result of external factors, and not attribute it to their own performance if the feedback is negative. Therefore they perceive this feedback as useless, since they view the information as inaccurate, and attribute poor appraisals to an external uncontrollable factor. Therefore the implications of valence on the proposed relationships need to be explored.

Overall it would be suggested that the proposed relationships would be stronger when a person has received a positive appraisal because they are more likely to accept this information as accurate. In contrast, the proposed relationships would be weaker when a person has received

a negative appraisal due to the fact they are less likely to accept this information as accurate.

Hypothesis 4a: Valence will moderate the positive relationship between performance goal orientation and perceived usefulness of ratings where the relationship will be stronger when ratings are positive.

Hypothesis 4b: Valence will moderate the positive relationship between learning goal orientation and perceived usefulness of comments where the relationships will be stronger when comments are positive.

Hypothesis 4c: Valence will moderate the relationship between tenure and perceived usefulness where the relationship will be stronger when ratings are positive than when the ratings are negative.

Hypothesis 4d: Valence will moderate the relationship between tenure and perceived usefulness where the relationship will be stronger when comments are positive than when the comments are negative.

Research Methodology

Procedure

The sample for this research project was based on professors at Concordia University. The first step of data collection process was to obtain the email addresses of the participants, which were collected from the Concordia website. The data collection period took place in 2013 between September 3rd and September 20th. Professors were sent an email on September 3rd, giving them a brief summary of the research project, and asking them to participate in the study (see appendix A). In the email, participants were provided with a hyperlink to an online questionnaire (see appendix B) that was designed using the Qualtrics software. Upon clicking this link, the participants were brought to a page that contained a consent form (see appendix C), which informed them about the purpose of the study, and their rights as participants. Participants gave their consent to participate by proceeding to the following web page. A follow up email (see appendix D) was sent on September 12th reminding the participants about the study and asking them again to participate. The participants of the study were asked to answer questions that dealt with their perception of the evaluations they receive from students at the end of teaching a course. Participants were assured in both the initial email and the consent form that their responses would be completely anonymous. The data was exported from the Qualtrics software into SPSS to be analyzed. The data collection protocol was examined and approved by

the Ethics Review Committee at Concordia University (certification number #30001220).

Sample

One-thousand-two-hundred-and forty-one faculty members from the university were asked to participate in this study on a voluntary basis. Participants were chosen to participate if they were from Concordia University, and if they had taught for at least one semester

Of the 1241 faculty members asked to participate in the study, 188 of them returned completed surveys, for a response rate of 15.15%. Incomplete surveys were removed from the sample.

As for the 188 participants who returned surveys 122 (64.89%) were from the Faculty of Arts and Science, 11 (5.85%) were from the Faculty of Engineering and Computer Science, 20 (10.63%) were from the Faculty of Fine Arts, and 35 (18.62%) were from the John Molson School of Business. When asked how long they had been teaching at Concordia University answers ranged from 1 to 47 years ($M=13.85$, $SD=11.38$). Professors were also asked on average how many students were in their classes, answers ranged from 10 to 180 students ($M=49.28$, $SD=32.23$). Another question asked was about the average amount of comments a professor had receive on course evaluations from students, answers ranged from 1 to 150 comments ($M=19.25$, $SD=18.25$)

Measures

Independent Variables

Tenure was measured using a single item. Participants were asked the question “how many years have you been employed at Concordia University?” and were instructed to enter the number of years that they have been employed at Concordia.

Goal orientation was measured using a scale that consisted of 16-items. This scale was taken from Button, Mathieu, and Zajac (1996). As Vandewalle (1997) discusses, compared to several other goal orientation scales, this one had the “strongest presentation of validation evidence.” As well this scale was chosen as it deals with goal orientation in an organizational setting rather than a grade school or high school setting that most of the past goal orientation research has looked at.

The goal orientation measure was split into two different scales of 8 items each. The first scale measured a person’s level of performance goal orientation. Participants were asked to indicate on a 5-point Likert type scale, how much they agree or disagree with the 8 items in the

scale. Possible answers ranged from 1 “Strongly disagree” to 5 “Strongly agree.” Participants with a score closer to 5 were said to have a high performance goal orientation, and participants with a score closer to 1 were said to have a low performance goal orientation. Sample items include “I prefer to do things I can do well rather than things that I do poorly,” and “I’m happiest at work when I perform tasks on which I know that I won’t make any errors.” The internal consistency was found to be good with a Cronbach’s alpha of 0.77

The second 8-item scale measured an individual’s level of learning goal orientation. Participants were asked to indicate on a 5-point Likert type scale, how much they agree or disagree with the 8 items in the scale. Possible answers ranged from 1 “Strongly disagree” to 5 “Strongly agree.” Participants with a score closer to 5 were said to have a high learning goal orientation, and participants with a score closer to 1 were said to have a low learning goal orientation. Sample items include “The opportunity to extend the range of my abilities is important to me,” and “when I have difficulty solving a problem, I enjoy trying different approaches to see which one will work.” The internal consistency was found to be good with a Cronbach’s alpha of 0.85.

Dependent Variables

Perceived usefulness of ratings was measured using a scale consisting of 3 items which was adapted from Subramanian (2007). As the original study by Subramanian (2007) dealt with the perceived usefulness of voice mail, the scale was adapted to talk about ratings by replacing the word “voice mail” by “ratings”. Participants were asked to indicate on a 5-point Likert type scale, how much they agree or disagree with the 3 items in the scale. Possible answers ranged from 1 “Strongly disagree” to 5 “Strongly agree.” A higher score indicated that participants found ratings as more useful. Sample items include “ratings provide information that increases my productivity,” and “ratings provide information that increases my effectiveness on the job.” The internal consistency was found to be good with a Cronbach’s alpha of 0.89.

Perceived usefulness of comments was measured using the same scale as above (Subramanian, 2007). However the word rating was replaced with comments. Sample items include “comments provide information that increases my productivity,” and “comments provide information that increases my effectiveness on the job.” The internal consistency was found to be good with a Cronbach’s alpha of 0.87.

Moderator Variables

Valence of ratings was measured using a single item. For this item participants were asked “thinking back to your past teaching evaluations please indicate how positive or negative the overall rating score you received was.” Possible answers ranged from 1 “very negative” to 5 “very positive.”

Valence of comments was measured using a single item. For this item participants were asked “thinking back to your past teaching evaluations please indicate how positive or negative the overall comments were.” Possible answers ranged from 1 “very negative” to 5 “very positive.”

Control Variables

additional measures were collected to be used as control variables. The first measure was class size, which asked participants to state the average number of students they have per class. It was important to control for class size as it can have an effect on performance appraisals given by students. For example in small class sizes a professor may interact more frequently with each student, which can influence a student’s response, (Westerlund, 2008).

The second control variable was faculty. Participants were asked to state which faculty they were a part of at Concordia University. Possible answers consisted of the Faculty of Arts and Science, the Faculty of Engineering and Computer Science, the Faculty of Fine Arts, and the John Molson School of Business. This was coded using three different dummy variables, with Faculty of Art and Science as the reference variable. This means that for each of the three dummy variables if 0 was selected that meant that the professor was from the Faculty of Arts and Science, however if 1 was selected, then the professors was in the faculty the dummy variable represented. It was important to include faculty as a control variable as different faculties often require different teaching methods and require faculty members take on different roles. This can influence how professors interact with students, and in turn the performance appraisals that the students write (McMurtrie, 2013).

Data Preparation

The first step of the data preparation was to manage missing and incorrect data. Any survey that was missing more than fifty percent of responses was removed. As well, one question asked participants to state which type of class they were going to focus on for their responses to a certain section of the survey. This question was put in the survey to ensure that participants answer questions in a consistent way, as focusing on different types of classes may result in

varying responses. If a participant did not answer the type of class question, their responses were removed from the data. In order to check for outliers, all scales were turned into z-scores.

Outliers were identified as z scores that were 3.29 standard deviation scores above or below the mean. One outlier was identified for the performance goal scale, therefore this outlier was removed. Before data management techniques were employed there were 248 surveys, however after these methods were employed 188 completed surveys remained.

The next step in data preparation was to examine each scale. This was done by correcting for any items that had been reverse coded. Next the internal consistency was tested by calculating the Cronbach's alpha of each scale. The scales were also tested to see if the Cronbach's alpha would be better if items were removed from the scales. All of the scales were found to have good internal consistency, and no significant benefits were found if any item was removed from any scale. Following this, a test variable was created by calculating the mean of all the items in each scale. For example the performance goal orientation scale had 8 items, an individual's performance goal orientation score was calculated by finding the mean for their responses across all items. The last step in managing the scales was to test for skewness and kurtosis. This was calculated using SPSS. Hair, Black, Babin, Anderson (2009) offer a quick rule for testing skewness or kurtosis. The basic rule is that an absolute skewness or kurtosis score below 1 indicated acceptable levels of skewness and kurtosis. Using this rule it was found that all scales had acceptable levels of kurtosis, and all scales with the exception of tenure has acceptable levels of skewness. In order to correct for the skewness in tenure the \log_{10} transformation command was used in SPSS.

Because the analytic strategy requires regression analysis, the next logical data preparation step was to test the regression assumptions. In order to test these assumptions, a method described by Kleinbaum, Kupper, Nizam, and Muller, (2008) was employed. This method consisted of creating both a scatter plot matrix and residual plots of each regression and comparing them to patterns Kleinbaum et al. (2008) show as being correct or incorrect. Based on this method, none of the regression assumptions were found to have been violated.

Below are two tables. Table 1 gives a brief overview of the preliminary analyses described above, while table 2 gives the correlations between the different variables. It is interesting to note that both usefulness of ratings and usefulness of comments are both positively correlated with valence.

Table 1: Measure Descriptives

	N	# Items	Mean	Sd	Min	Max	Skewness	Kurtosis	Cronbach's alpha
Tenure	176	1	13.85	11.3	1	47	1.12	.81	N/A
Ratings	187	3	3.20	.96	1	5	-.37	-.27	.89
Comments	186	3	3.57	.88	1	5	-.79	.72	.87
Valence Ratings	188	1	4.31	.65	2	5	-.65	.52	N/A
Valence Comments	188	1	4.17	.78	2	5	-.85	.59	N/A
PGO	184	8	3.50	.52	1.75	4.88	-.27	.74	.77
LGO	182	8	4.16	.46	2.75	5	-.02	-.28	.85

PGO = Performance goal orientation

LGO= Learning goal orientation

Table 2: Correlations

	Tenure	Valence Rating	Valence Comment	PGO	LGO	Ratings	Comments
Tenure		-.08	-.08	-.09	-.08	.01	-.14
Valence Rating	-.08		.59**	.07	.10	.36**	.19**
Valence Comment	-.08	.59**		.00	.12	.23**	.35**
PGO	-.09	.07	.00		.122	.09	.01
LGO	-.08	.10	.12	.122		.11	.15*
Ratings	.01	.36**	.23**	.09	.11		.51**
Comments	-.14	.19**	.35	.12	.15*	.50**	

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

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

PGO = Performance goal orientation

LGO= Learning goal orientation

Analytical Strategy and Results

Hypothesis 1 states that people will perceive receiving comments as more useful than receiving ratings. In order to test this hypothesis a paired sample t test was used. The two variables, perceived usefulness of ratings, and perceived usefulness of comments were entered into SPSS and were tested using a paired sample t test to see if the means of the two variables were different.

Hypothesis 1 was supported. The paired sample t test showed that there was a significant difference in the scores for perceived usefulness of ratings ($M = 3.19$, $SD = .96$) and perceived usefulness of comments ($M = 3.58$, $SD = .87$), $t(184) = -5.87$, $p = .05$ (one-tailed), $d = .86$. This suggests that on average, university professors in this study perceive comments to be more useful than ratings.

Hypothesis 2a states that a performance goal orientation will be more positively related to the perceived usefulness of ratings than to the perceived usefulness of narrative comments. In order to test this hypothesis two regression models must be compared, the first model being how performance goal orientation relates to perceived usefulness of ratings and the second model being how performance goal orientation relates to perceived usefulness of narrative comments. To test model 1, a hierarchical regression was run using SPSS in which the control variables, faculty, and average number of students were entered into the first block of independent variables, and performance goal orientation was entered into the second block of independent variables. The dependent variable was the perceived usefulness of ratings. This regression was run using the enter method. For model 2, an identical regression was run except the dependent variable was the perceived usefulness of comments. The beta coefficients of each model were then extracted from the SPSS output and compared using the equation $T = \frac{b_1 - b_2}{\sqrt{\frac{df_1(SEb_1^2) + df_2(SEb_2^2)}{df_1 + df_2}}}$

which as discussed by Paternoster, Brame, Mazerolle, and Piquero (1998) is a widely accepted equation to compare two coefficients.

Hypothesis 2a was not supported. The regression for model 1 showed that a performance goal orientation was not significantly related to the perceived usefulness of ratings, $\beta = .19$, $R^2 = .03$, *n.s.* (see table 3). The regression for model 2 showed that a performance goal orientation was not significantly related to the perceived usefulness of comments, $\beta = .21$, $R^2 = .03$, *n.s.* (see

table 4). As well beta coefficient 1 ($b1 = .18$, $SEb1 = 0.14$) and beta coefficient 2 ($b2 = .21$, $SEb2 = 0.13$) were not found to be statistically different from one another, $t(180) = 0.16$, n.s.

Table3: Hypothesis 2a: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of ratings (N=181)

Variables	B	SE B	β
Step 1			
Average Number of Students	.00	.00	.06
Faculty of Engineering	-.25	.30	-.06
Faculty of Fine Arts	.31	.25	.09
John Molson School of Business	-.18	.19	-.07
Step 2			
Average Number of Students	.00	.00	.06
Faculty of Engineering	-.24	.30	-.06
Faculty of Fine Arts	.34	.25	.1
John Molson School of Business	.19	.19	.08
Performance Goal Orientation	.19	.14	.10

Note: $R^2 = .02$ for step 1. $\Delta R^2 = .01$ for step 2

Table 4: Hypothesis 2a: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of comments (N=180)

Variables	B	SE B	β
Step 1			
Average Number of Students	.00	.00	.06
Faculty of Engineering	-.44	.28	-.12
Faculty of Fine Arts	.06	.22	.02
John Molson School of Business	-.14	.17	-.06
Step 2			
Average Number of Students	.00	.00	.00
Faculty of Engineering	-.42	.28	-.12
Faculty of Fine Arts	.10	.22	.03
John Molson School of Business	-.13	.17	-.05
Performance Goal Orientation	.21	.13	.12

Note: $R^2 = .02$ for step 1. $\Delta R^2 = .01$ for step 2

Hypothesis 2b states that a learning goal orientation will be more positively related to the perceived usefulness of narrative comments than to the perceived usefulness of ratings. In order to test this hypothesis two regression models must be compared, the first model being how a learning goal orientation relates to perceived usefulness of ratings and the second model being how a learning goal orientation relates to perceived usefulness of narrative comments. To test model one a hierarchical regression was run using SPSS where the control variables, tenure, and faculty, was entered into the first block of independent variables, and learning goal orientation

was entered into the second block of independent variables. The dependent variable was the perceived usefulness of ratings. This regression was run using the enter method. For model two an identical regression was run except the dependant variable was the perceived usefulness of comments. The beta coefficients of each model were then compared using the same equation as hypothesis 1.

Hypothesis 2b was not supported. The regression for model 1 showed that a learning goal orientation was not significantly related to the perceived usefulness of ratings, $\beta = .18$, $R^2 = .02$, *n.s.* (see table 5). The regression for model 2 showed that a learning goal orientation was not significantly related to the perceived usefulness of comments, $\beta = .25$, $R^2 = .05$, *n.s.* (see table 6). As well beta coefficient 1 ($b1 = .18$, $SEb1 = 0.16$) and beta coefficient 2 ($b2 = .25$, $SEb2 = 0.14$) were not found to be statistically different from one another, $t(167) = 0.49$, *n.s.*

Table 5: Hypothesis 2b: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of ratings (N=168)

Variables	B	SE B	β
Step 1			
Tenure	-.15	.19	-.06
Faculty of Engineering	-.06	.34	-.01
Faculty of Fine Arts	.23	.25	.07
John Molson School of Business	.20	.19	.08
Step 2			
Tenure	-.11	.19	-.05
Faculty of Engineering	-.08	.34	-.02
Faculty of Fine Arts	.18	.25	.06
John Molson School of Business	.20	.19	-.08
Learning Goal Orientation	.18	.17	.09

Note: $R^2 = .01$ for step 1. $\Delta R^2 = .01$ for step 2.

Table 6: Hypothesis 2b: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of comments (N=167)

Variables	B	SE B	β
Step 1			
Tenure	-.24	.16	-.11
Faculty of Engineering	-.42	.29	-.11
Faculty of Fine Arts	.14	.21	.05
John Molson School of Business	-.15	.17	-.07
Step 2			
Tenure	-.19	.16	-.09
Faculty of Engineering	-.44	.29	-.12
Faculty of Fine Arts	.08	.21	.03
John Molson School of Business	-.14	.17	-.07
Learning Goal Orientation	.25	.14	.14

Note: $R^2 = .03$ for step 1. $\Delta R^2 = .02$ for step 2.

Hypothesis 3a states that as tenure increases the perceived usefulness of both ratings and narrative comments will decrease. In order to test this hypothesis two regression models must be run, the first model being how tenure relates to perceived usefulness of ratings and the second model being how tenure relates to perceived usefulness of narrative comments. To test model one, a regression was run using SPSS where the independent variable was tenure, and the dependent variable was the perceived usefulness of ratings. This regression was run using the enter method. For model two an identical regression was run except the dependant variable was the perceived usefulness of comments.

Hypothesis 3a was partially supported. The regression for model 1 showed that tenure was not significantly related to the perceived usefulness of ratings, $\beta = -.08$, $R^2 = .00$, n.s. (see table 7). However, the regression for model two also showed that tenure was significantly related to the perceived usefulness of narrative comments, $\beta = -.34$, $R^2 = .03$, $p = .05$ (see table 8). This means that as tenure increases people view comments as less useful.

Table 7: Hypothesis 3a: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of ratings (N=174)

Variables	B	SE B	β
Tenure	-.08	.18	-.04

Note: $R^2 = .00$

Table 8: Hypothesis 3a: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of comments (N=173)

Variables	B	SE B	β
Tenure	-.34	.16	-.16*

Note: $R^2 = .03$ Note: * = $p = .05$.

Hypothesis 3b states that as tenure increases, the perceived usefulness of ratings will decrease faster than the perceived usefulness of narrative comments. This hypothesis was tested using the same Z score formula used for hypotheses 2a and 2b. Hypothesis 3b was not supported. However the results show that the opposite of the proposed relationship was in fact supported. Meaning that the perceived usefulness of narrative comments ($b_2 = -.34$, $SEb_2 = .16$) decreases faster than the perceived usefulness of ratings ($b_1 = -.08$, $SEb_1 = .18$) where $t(173) = -1.49$, $p = .01$ (one-tailed), $d = .23$. This suggests that as tenure increases, people will tend to find comments as a less useful form

Hypothesis 4a states that valence will moderate the positive relationship between performance goal orientation and perceived usefulness of ratings where the relationship will be stronger when ratings are positive. The first step to test this hypothesis was to centralize the independent variables around the mean so that they had similar values. This was done by turning both valence of ratings and performance goal orientations values into Z scores. Then the new variables were multiplied together to create the interaction variable. A hierarchical regression was then run using SPSS where the control variables, tenure, faculty, and average number of students were entered into the first block of independent variables, performance goal orientation, and valence of ratings were entered into the second block of independent variables, and the interaction term was entered into the third block of independent variables. The dependent variable was the perceived usefulness of ratings. This regression was run using the enter method. The R^2 of the block 3 was then compared to the R^2 of block 2 using a partial f test to test the significance of valence when added to the model.

Hypothesis 4a was not supported. The regression analysis showed that there is a significant

relationship between performance goal orientation, valence, and the perceived usefulness of ratings, $R^2 = .13$, $p < .01$. (see table 9). However the partial F test demonstrated that the interaction term does not significantly add to the model $F(2,165) = .39$, n.s.

Table 9: Hypothesis 4a: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of ratings (N=169)

Variables	B	SE B	β
Step 1			
Tenure	-.16	.18	-.07
Average Number of Students	.00	.00	.07
Faculty of Engineering	-.07	.34	-.02
Faculty of Fine Arts	.29	.26	.09
John Molson School of Business	-.18	.19	.08
Step 2			
Tenure	-.06	.17	-.02
Average Number of Students	.00	.00	.08
Faculty of Engineering	-.16	.321	-.04
Faculty of Fine Arts	.13	.25	.04
John Molson School of Business	.18	.18	-.07
Performance Goal Orientation	.14	.14	.08
Valence of Ratings	.51	.11	.34*
Step 3			
Tenure	-.06	.17	-.02
Average Number of Students	.00	.00	.08
Faculty of Engineering	-.15	.32	-.03
Faculty of Fine Arts	.15	.25	.05
John Molson School of Business	.20	.18	.08
Performance Goal Orientation	.13	.14	.07
Valence of Rating	.51	.11	.33*
Performance Goal Orientation X Valence	.05	.09	.05

Note: $R^2 = .02$ for step 1. $\Delta R^2 = .11$ for step 2. $\Delta R^2 = .00$ for step 3. Note: * = $p = .05$.

Hypothesis 4b states that valence will moderate the positive relationship between learning goal orientation and perceived usefulness of narrative comments where the relationship will be stronger when comments are positive. The first step to test this hypothesis was to centralize the independent variables around the mean so that they had similar values. This was done by turning both valence of comments and learning goal orientation values into Z scores. Then the new variables were multiplied together to create the interaction variable. A hierarchical regression was then run using SPSS where the control variables faculty, average number of students, and tenure were entered into the first block of independent variables, learning goal orientation, and valence of comments were entered into the second block, and the interaction term was entered

into the third block of independent variables. The dependent variable was the perceived usefulness of comments. This regression was run using the enter method. The R^2 of the second block was then compared to the R^2 of the third block using a partial f test to test the significance of valence when added to the model.

Hypothesis 4b was not supported. The regression analysis showed that there is a significant relationship between performance goal orientation, valence, and the perceived usefulness of ratings, $R^2 = .14$, $p < .01$ (see table 10). However the partial F test demonstrated that the interaction term does not significantly add to the model $F(2,163) = .84$, n.s.

Table 10: Hypothesis 4b: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of comments (N=179)

Variables	B	SE B	β
Step 1			
Faculty of Engineering	-.46	.27	-.13
Faculty of Fine Arts	.09	.21	.03
John Molson School of Business	-.12	.17	.05
Step 2			
Faculty of Engineering	-.39	.26	-.11
Faculty of Fine Arts	.02	.20	.00
John Molson School of Business	-.13	.16	-.06
Learning Goal Orientation	.20	.13	.10
Valence of Comments	.36	.08	.32*
Step 3			
Faculty of Engineering	-.44	.26	-.12
Faculty of Fine Arts	.03	.20	.01
John Molson School of Business	-.13	.16	-.06
Learning Goal Orientation	.18	.14	.10
Valence of Comments	.36	.08	.32*
Learning Goal Orientation X Valence	.06	.07	.06

Note: $R^2 = .02$ for step 1. $\Delta R^2 = .12$ for step 2. $\Delta R^2 = .00$ for step 3. Note: * = $p = .05$.

Hypothesis 4c states that valence of ratings will moderate the relationship between tenure and perceived usefulness of ratings where the relationship will be stronger when ratings are positive. The first step to test this hypothesis was to centralize the independent variables around the mean so that they had similar values. This was done by turning both valence of ratings and tenure values into Z scores. Then the new variables were multiplied together to create the interaction variable. A hierarchical regression was then run using SPSS where the control variables, faculty, and average number of students was entered into the first block of independent variables, tenure and valence of ratings were entered into the second block, and the interaction

term were entered into the third block of independent variables. The dependent variable was the perceived usefulness of ratings. This regression was run using the enter method. The R^2 of the third block was then compared to the R^2 of the second block using a partial f test to test the significance of valence when added to the model.

Hypothesis 4c was not supported. The regression analysis showed that there is a significant relationship between tenure, valence, and the perceived usefulness of ratings, $R^2 = .15$, $p < .01$. (see table 11). However the partial F test demonstrated that valence does not significantly add to the model $F(2,170) = 0$, n.s.

Table 11: Hypothesis 4c: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of ratings (N=174)

Variables	B	SE B	β
Step 1			
Average Number of Students	.00	.00	.99
Faculty of Engineering	-.08	.34	-.02
Faculty of Fine Arts	.30	.26	.09
John Molson School of Business	.17	.19	.07
Step 2			
Average Number of Students	.00	.00	.10
Faculty of Engineering	-.19	.32	-.04
Faculty of Fine Arts	.07	.24	.02
John Molson School of Business	.15	.18	.06
Tenure	-.04	.17	-.02
Valence of Ratings	.57	.11	.37*
Step 3			
Average Number of Students	.00	.00	.10
Faculty of Engineering	-.19	.32	-.04
Faculty of Fine Arts	.07	.25	.02
John Molson School of Business	.15	.18	.06
Tenure	-.04	.17	-.02
Valence of Rating	.56	.11	.37*
Tenure X Valence	.00	.08	.00

Note: $R^2 = .02$ for step 1. $\Delta R^2 = .13$ for step 2. $\Delta R^2 = .00$ for step 3. Note: * = $p = .05$.

Hypothesis 4d states that valence of comments will moderate the relationship between tenure and perceived usefulness of comments where the relationship will be stronger when comments are positive. The first step to test this hypothesis was to centralize the independent variables around the mean so that they had similar values. This was done by turning both valence of comments and tenure values into Z scores. Then the new variables were multiplied together to create the interaction variable. A hierarchical regression was then run using SPSS where the

control variables, faculty, and average number of students was entered into the first block of independent variables, tenure and valence of comments were entered into the second block, and the interaction term was entered into the third block of independent variables. The dependent variable was the perceived usefulness of comments. This regression was run using the enter method. The R^2 of the third block was then compared to the R^2 of the second block using a partial f test to test the significance of valence when added to the model.

Hypothesis 3d was not supported. The regression analysis showed that there is a significant relationship between tenure, valence, and the perceived usefulness of ratings, $R^2 = .15, p < .01$. (see table 12). However the partial F test demonstrated that the interaction variable does not significantly add to the model $F(2,169)=0$, n.s.

Table 12: Hypothesis 4d: Summary of Hierarchical Regression Analysis for Variables Predicting Perceived usefulness of comments (N=173)

Variables	B	SE B	β
Step 1			
Average Number of Students	.00	.00	.03
Faculty of Engineering	-.39	.30	-.10
Faculty of Fine Arts	.14	.22	.05
John Molson School of Business	-.19	.17	-.09
Step 2			
Average Number of Students	.00	.00	-.11
Faculty of Engineering	-.32	.28	-.08
Faculty of Fine Arts	.11	.21	.04
John Molson School of Business	-.15	.16	-.07
Tenure	-.27	.15	-.13
Valence of Ratings	.35	.08	.32*
Step 3			
Average Number of Students	.00	.00	-.01
Faculty of Engineering	-.32	.29	-.08
Faculty of Fine Arts	.11	.21	.04
John Molson School of Business	-.15	.16	-.07
Tenure	-.27	.15	-.13
Valence of Rating	.35	.09	.32*
Tenure X Valence	.00	.06	.00

Note: $R^2 = .02$ for step 1. $\Delta R^2 = .13$ for step 2. $\Delta R^2 = .00$ for step 3. Note: * = $p = .05$.

Discussion

The purpose of this study was to explore the difference between different appraisal formats, more specifically the relationship between personality and individual differences with a ratee's perception of how useful they find different appraisal formats to be. This study looked at the personality trait of goal orientation, as well as the individual difference of tenure. The appraisal formats that were examined were ratings and comments. It was argued that different personality traits or individual differences would lead a ratee to find one format as more useful over another. Finally valence was explored as a moderator for the relationships mentioned above.

Ratings Versus Comments

This study found a statistically significant difference in how useful the ratee finds each performance appraisal format. This means that it was found that on average a ratee finds comments as a more useful form of feedback than ratings. While both formats were considered as useful feedback, comments appear to be the favored format.

Penny and Coe (2004) discuss how one of the major reason a professor might find teaching evaluations unhelpful is that they can be difficult to interpret. Ratings can provide information about the level of performance, but are not always useful in understanding how to change behavior in order to reach a better level of performance. Comments on the other hand, might be viewed as a more useful appraisal format because they provide information about how behavior can be changed in order to reach desired performance. Comments can provide diagnostic feedback that is specific in identifying an individual's strengths and weaknesses, and providing prescriptive information about how they might go about improving.

As discussed in the introduction, ratings are most often used for administrative decisions, while comments are used for developmental purposes. While institutions of higher education use appraisals for both administrative and developmental purposes, traditionally, professors have supported the use of teacher evaluations as a development tool but not for administrative purposes, (Denson, Loveday, & Dalton, 2010). This means that professors most likely put more importance on teacher evaluations being used as a developmental tool. In turn, since comments are the appraisal format most related to development, professors find comments as a more useful format over ratings.

In more general terms beyond the university classroom, comments have many benefits. Dalessio (1998) argues that comments actually compliment ratings by providing more information that helps individuals interpret ratings, meaning comments go above and beyond ratings to provide pertinent feedback. There also exists evidence that a ratee tends to pay more attention to words than numbers, (Fersti, & Bruskiwicz, 2000). This could be due to the fact that although people often have access to many kinds of information, they usually prefer more natural forms of communications (Mintzberg, 1973), which would be words more so than numbers. Overall comments appear to be favored, people pay more attention to them, they can provide information beyond ratings, and people tend to prefer comments as they are a more natural form of communication.

Brutus (2009) details many characteristics that make comments a strong form of feedback including the domain coverage of comments, the specificity of comments, suggestions for developments provided by comments, and the amount of narrative units. While ratings are designed to capture performance in a domain accurately (Bretz, Milkovitch, & Read, 1992), comments can capture performance in a broader and richer context. Comments also differ in their level of specificity, which means that comments can provide either a little information, or a large amount of information. Ratings on the other hand provide only specific information. This makes comments as a more useful format as it can be adapted to different contexts that may need more or less feedback information. Comments is also a strong appraisal format because it can provide specific information like ratings, but can also go further to provide suggestions as to how an individual can change behavior in order to improve performance. Furthermore, while ratings provide one number for a specific performance domain, comments are relatively unbound, and can provide multiple narrative units per performance domain. Overall comments might be seen as a more useful appraisal format over ratings because comments provide richer, more specific, and a greater amount of information than ratings can.

Goal Orientation

This study found no significant relationship between goal orientation and the perceived usefulness of the different appraisal formats. As well each of the two goal orientations were not found to be more positively related to one format over another. This means that a performance goal orientation was not statistically linked to the perceived usefulness of ratings or comments.

As well a performance goal orientation was not more positively linked the perceived usefulness of ratings than the perceived usefulness of comments as was predicted. Similarly a performance goal orientation was not statistically linked to the perceived usefulness of ratings or comments. As well a performance goal orientation was not more positively linked the perceived usefulness of comments than the perceived usefulness of ratings as was predicted.

Contrary to what was predicted performance goal orientation was not linked to the perceived usefulness of ratings. As the literature shows, individuals high in performance goal orientation look to outperform their colleagues. Ratings provide a simple way for people with a performance goal orientation to compare their performance to another person. As far as teaching evaluations go, the performance appraisals are kept relatively secure and secret; in turn this might make it difficult for professors to compare their performance to one of their colleagues. In turn people with a performance goal orientation are not able to compare their ratings scores to others. While ratings are viewed as a relatively useful form of feedback, the reason for this is not because people have a performance goal orientation, we know this because in this case the evaluations of others were kept secret.

Other possible reasons for why a performance goal orientation is not linked to the perceived usefulness of ratings may exist. For example alternate views of a performance goal orientation exist. As Potosky (2010) suggests a performance goal orientation can be split into two sub categories. The first is approach goal orientation which is the traditional performance goal orientation trait used in this project where people seek to outperform others. The second is avoidance performance goal orientation where people attempt to avoid looking incompetent often by avoiding taking part in achievement situations all together. It is possible that individuals with an avoidance performance goal orientation will not seek out any kind of feedback as it may harm their self esteem, therefore they would not find ratings as useful. Over all since both types of performance goal orientations were grouped together, it is possible that one cancelled the other out.

Another finding of this study was that performance goal orientation was not more positively related to ratings than comments, and a learning goal orientation was not more positively related to comments than ratings. There might be a very good explanation for this finding. As this research project took the stance that goal orientation is a relatively stable trait with both orientations at each end of a continuum, recent research has built alternate views. One

alternate view states that a performance goal orientation and learning goal orientation are not mutually exclusive, and are dependent on situational variables, (Farr, Hoffman, & Rigenbach, 1993). Research argues that two types of achievement climate or situations exist. “In a task involving climate emphasis is placed on personal effort and working with others to develop and enhance skills, which may foster self referenced forms of ability such as striving for personal mastery. On the other hand, in an ego involving climate individuals are rewarded and evaluated based on their abilities compared to others, which may promote judgment of competence based on normative comparison,” (Boyce, Gano-Overway, & Campbell, 2009). In a teacher evaluation context, professors strive both to learn how to improve teaching techniques, as well to outperform colleagues so that administrative decisions can be made in their favor. In this context both comments and ratings would be highly valuable feedback for professors.

Tenure

This study did find that as tenure increases the perceived usefulness of appraisal formats decrease at different rates. It was found that as tenure increases, the perceived useful of comments decreases faster than the perceived usefulness of ratings. This supports the overall argument of this study that individual differences will influence the perceived usefulness of different performance appraisal formats. This finding however was the opposite of what was hypothesized. There may be a very good explanation for this. People new to a job need to develop their skills. As comments can provide specific prescriptive feedback as to how an individual may change their behavior in order to perform their job better, people with low tenure will find comments as useful. However the longer the person has been at a job, the less they will have to learn how to perform their job as they already have the knowledge. Therefore comments become perceived as a less useful appraisal format. Ratings on the other hand are primarily used for administrative purposes such as promotions and raises. Basically ratings provide a measure of performance that can be easily compared to other performances. As people’s tenure and work experience increase they will expect promotions and raises as they are continually bettering their skills. Ratings provide these people with a concrete way to gauge their performance, and can be used as evidence as to why they deserve a promotion or raise. Ratings will be viewed as more useful than comments for a longer time because ratings have more to do with compensation than actually developing ones skills.

Tenure was not linked to the perceived usefulness of appraisal formats. Intuitively one would expect that people who are new to a job would find feedback as more useful as they need to learn how to perform their jobs correctly. People with more experience will have more knowledge of their job, and therefore rely on feedback less to understand their job role. On the other hand there may be some very good explanations as to why no significant relationship was found between tenure and perceived usefulness of appraisal formats. For example, this study assumed that a person's job is stagnant. This means that a person who has been at a company for one year will be expected to perform the exact same tasks as someone who has been at the same company for twenty years. In actuality, as an individual's tenure increases at a company, they will be given promotions, their job roles will become diverse and they will have to learn how to perform new job tasks. In this case people who have been at a company longer may have to learn more complicated tasks, and therefore more rely on feedback even more than people who have not been at a company as long and may have simpler job tasks to learn.

Valence

Although it was found that valence did not moderate any of the proposed relationships between individual differences and perceived usefulness of appraisal formats, it was positively correlated with the perceived usefulness of both comments and ratings. There are many possible explanations for this finding. First of all it is possible that the finding supports previous research that argues that when people receive negative feedback, they are likely to attribute the feedback to an external factor, blaming some external force for why they were not able to perform up to par. However when individuals receive positive feedback, they are likely to attribute the feedback to internal factors, for example claiming they performed well because they worked hard. When an appraisal is attributed to an external factor, it is seen as not accurate feedback, therefore not useful. However when feedback is attributed to internal factors the feedback is perceived as accurate, and therefore it is perceived as useful.

The correlation between valence and perceived usefulness of appraisal format can also be due to measurement bias. In this study valence was measured using a self report measure instead of the actual valence of their appraisal. This is a problem because participants may not report accurately the past valence due to faulty memory or wanting to inflate their past scores.

Based on the established literature it was predicted that valence would moderate the relationship between individual differences and perceived usefulness of appraisal formats, where

more positive ratings or comments would increase the proposed relationships. This was because people are more likely to react in a positive way to favorable feedback, and react in a negative way to less favorable feedback, (Feys et al. 2011) This study found that valence did not significantly moderate any of the proposed relationships. This means that the valence did not increase the positive relationship between perceived usefulness of ratings and performance goal orientation. Valence did not increase the positive relationship between perceived usefulness of comments and learning goal orientation. Valence did not increase the negative relationship between tenure and the perceived usefulness of ratings. As well, valence did not increase the negative relationship between tenure and the perceived usefulness of comments.

Practical Implications

Although the proposed relationships between individual differences and perceived usefulness of appraisal format were not found to be as proposed, the results do offer interesting insights as to how appraisal formats differ. The findings support the argument that both ratings and comments are perceived as different formats, with different uses. As well there is support that individual difference can influence how useful an individual perceives each appraisal format to be.

One of the biggest issues regarding teacher evaluations is professors' attitudes towards these appraisals, ranging from acceptance to total dismissal, (Nasser, & Fresko, 2002). If the professors do not view the appraisal as useful information, there is no point in trying to give feedback at all. This issue raises two questions. First of all what are the aspects of the appraisal process that professors like or dislike, and second, what differs between the professors who view teacher evaluations as useful and those professors who don't.

The main goal of this study was to explore people's reactions to different appraisal formats, mainly ratings and comments. This study found that both ratings and comments were rated as useful by professors, with comments being viewed as a more useful format than ratings. When answering the first question stated above, it would appear that appraisal format is an aspect of the appraisal process that has an effect on professor's perceptions of teacher evaluations. This is evident in the fact that perceived usefulness fluctuates depending on the type of appraisal format. Since both appraisal formats are perceived as useful information by professors, they should both be included together on teacher evaluations. This will give professors the most amount of useful information possible, aiding them to develop their teaching

abilities.

The second goal of this study was to look at how individual differences might affect an appraisee's reaction to different appraisal formats. Past research has found that individual characteristics can have an effect on how useful an appraisee finds the appraisal process in general, (Tuytens & Devos, 2013). This study found that as tenure increases, the perceived usefulness of comments will decrease faster than the perceived usefulness of ratings. This shows that not only can individual differences affect the perceived usefulness of appraisal feedback, but it can also affect the perceived usefulness of the type of feedback. Overall if we want professors to view teacher evaluations as useful information, and to use the feedback provided to develop teaching skills, the appraisal process needs to be structured in a way that takes into account format, and individual differences. For example it would make most sense to give comments to professors with low tenure as they will view the information as useful; however, comments would be less effective for professors with high tenure.

Beyond just looking at the individual difference of tenure, those appraising performance should keep in mind that other individual differences may influence how the appraisee perceives the appraisal process. For example, Ogunfowora, et al. (2009) concluded that personality traits can influence what performance dimensions an individual puts weight on. Although there is still little research on which individual differences influence perceptions of appraisal formats, there are still ways to indirectly account for them. In the case of teacher evaluations, professors can be asked to take part in the creation of the tool that will be used to appraise their teaching abilities. Taking part in this process has been shown to illicit positive reactions from professors (Tuytens et al, 2013). This will allow professors to help build an appraisal tool around their personal preferences for feedback, which in turn will take into account their individual differences.

This study also provides good news for those who use performance appraisals for developmental purposes based on the theories of behavioral change. The theory of reasoned action (Fishbein & Ajzen, 1975) states that in order to motivate an individual to change behavior, it is important to create positive attitudes in the appraisee and to offer desirable outcomes. This study found that overall an appraisee finds comments as a useful format. Comments are the format most used for developmental purposes. Therefore just using comments in a developmental context will bring about positive attitudes in the appraisee as they have positive attitudes towards receiving comments. Comments may inherently be a motivating force to

change behavior.

Atwater et al (2006) found that individuals prefer ratings over comments, while this study found that comments are preferred over ratings. This strengthens arguments that context is an important factor that can influence appraisal outcomes (Levy et al, 2004). Depending on the context individuals may prefer one format over another. Considering this information, those delivering an appraisal need to consider the context in which they are giving an appraisal. For example, a more complex job may require more specific feedback which can be best given through comments; this appraisal should be more heavily weighted with comments. On the other hand if an appraisal is being used for administrative purposes there should be a heavier weight placed on ratings.

Overall Performance appraisals are complex tools that need to be tailored to context. Practitioners creating these tools need to keep many factors in mind. First of all, the formats are different and can each give different types of feedback. Second, individual differences influence what feedback an individual finds useful. Individual differences should be considered during the creation of the appraisal tool, and the appraisee should be included in the process in order to insure their preferences for feedback are included. Third of all, different formats may be perceived as more useful in different contexts, when deciding on what format to focus on; the context needs to be considered.

Limitations

As mentioned above this study failed to find support for the proposed relationships. Perhaps one explanation for this failed support can be explained by the limitations of this study. First of all this study relied on self report measures. This poses accuracy issues with the data collected, as it asks participants to indicate perceptions, as well as remembering things such as the valence of previous appraisals. Perceptions are subjective and can be inaccurate due to personal biases of the participants (Dalal, 2013). As well, trying to remember past events can be inaccurate as memories are often faulty (Bryman, Bell, Mills, & Yue, 2011). Professors were also asked to leave any additional comments at the end of the survey they were given. A few repeated comments point to some limitations of this study. For example some participants felt that the questionnaire was repetitive to a degree, and this would influence participants to answer inaccurately as they got bored or annoyed with the questions.

The university setting is also very unique. For this reason the findings of this study cannot

be generalized to general business professionals. In the university setting professors have to take on multiple roles, teachers, administrators, and researchers. Universities differ in what they place emphasis on, and for many the most important role of a professor is research. This might have an impact when taking into account the purpose of the appraisal. Furthermore the job characteristics model proposed by Hackman and Lawler (1971) argues that characteristics of a job will have an effect on employees' cognitions and behaviors. One of these characteristics is job scope. It has been argued that individuals with a larger job scope will have higher levels of motivation, job satisfaction, and performance, (Hackman & Oldham, 1976). Professors have a large job scope, taking on a variety of roles; this in turn has a direct effect on their performance. Being responsible for taking on many different roles requires broader feedback that can give information to a professor regarding all the different tasks they have to perform, how well they performed them, and how they might improve. Professors would therefore find having both ratings and comments as useful as the formats give broader feedback that can cover a broad job scope. On the other hand, jobs that have a small job scope would require less information for the appraisee to understand how they performed and how to improve.

Autonomy, which "refers to the control the worker enjoys with respect to choosing among operations, ordering the operations, and selecting a work space," (Dodd & Ganster, 1996), has also been positively linked to job performance, (Hackman et al, 1976). Professors have a high level of autonomy at their job. Research has found that individuals who perceive their job as having low autonomy tend to define their jobs in a very narrow sense. Individuals who perceive their job as having high autonomy, such as professors view their job role in a more flexible way, and often perform tasks beyond their formal job descriptions, (Bizzie & Soda, 2011). If professors view their job as having high levels of autonomy and perform tasks beyond their job description, it could be possible that professors will feel performance appraisals will only capture performance related to their formal job descriptions, and not account for the tasks they perform beyond their formal job description. This could impact how much a professor feels an appraisal accurately captures performance of all the tasks they take part in. In another organization where individuals have low autonomy, and have a narrower job description, it could be easier to capture performance on all tasks they perform, and therefore give a more accurate performance appraisal.

Professors tend to find teacher appraisals as a useful tool to develop their teaching skills,

but not for administrative purposes, (Denson, et al. 2010). Administrative decisions might be more influenced by the administrative and research roles professors take on. This could have had an impact on how useful professors rated ratings on teacher evaluations, as they are most commonly associated with administrative decisions. Unfortunately this study did not take this into account and was unable to fully understand the importance of professors' teacher evaluations on their overall appraisal.

Future Research

While this study gave some insight into the perceived usefulness of ratings and comments, it would be useful to do more research in this area to better understand the two appraisal formats. Some suggestions for future research are discussed below.

According to Tuytens and Devos (2013) there are three factors that will influence a professor perceptions of usefulness of teacher evaluations, having a charismatic leader, taking part in the appraisal process, and individual characteristics. Similarly Pichler (2012) argued that the rater-ratee relationship, appraisal participation, and performance rating favourability impact reactions to the appraisal process. While the current study examined individual characteristic, and rating favourability, and how they impact perceived usefulness of ratings and comments, this study did not touch on the other proposed antecedents of a ratee's reaction to the appraisal process. Fundamentally by looking at both theories, having a charismatic leader, taking part in the process, and the rater-ratee relationship are all related to procedural justice. Future research should look at how ratings and comments impact procedural justice.

This study found some support that individual differences can influence how professors perceive ratings and comments. Future research would benefit from looking at different individual differences and personality traits that might influence individuals' perceptions of different appraisal formats. One of these traits is extraversion, which can be defined as individuals who are sociable, gregarious, assertive and active (Bell & Arthur, 2008). This personality trait has been found to be positively related to feedback seeking, (Wanberg, & Kammeyer-Mueller 2000), and has also been found to be positively related to feedback acceptance, (Bell et al, 2008). Future research could take these finding a step further, and try and determine if extroverted find one appraisal format as more useful than another.

An individual's level of emotional stability can also influence their reaction to feedback. Individuals who have low emotional stability can view feedback as more of a personal attack

than constructive criticism, and therefore have negative reactions to the feedback, (Bell et al, 2008). Future research can look at how emotional stability affects perceptions of appraisal formats. If individuals with low emotional stability view feedback as a personal attack, these feelings might be stronger when receiving personalized comments.

Self esteem has also been shown to influence individual reactions to feedback. Individuals with high self esteem tend improve performance more than individuals with low self esteem after receiving positive feedback. As well people with low self esteem tend to have a greater decline in performance after negative feedback than do individuals with high self esteem (Brockner, Derr, & Laing, 1987). Future research should seek to determine if the level of self esteem will cause individuals to react differently to different appraisal formats.

Feys et al. (2011) discuss how there have been very few studies that have looked at the difference between ratings and comments. The current study found that overall comments were perceived as more useful than ratings. While this might be the case, it gives little information as to why comments are found as more useful. Future research should look at the characteristics of comments to see which characteristics are most important in different contexts. Brutus (2009) outlines five different characteristics of comments which include, amount, valence, domain coverage, specify, and provision of suggestions. This study could easily be replicated with replacing the usefulness of comments with each of the five characteristics of comments in order to determine which characteristic leads an appraisee to view a given performance appraisal as useful.

One of the most important research areas in organizational research and one of the reasons for studying comments is context, (Cappelli & Scherer, 1991). The current study used the university context, and examined professor's perceptions of teacher evaluations. While the current study found comments to be perceived as more useful than ratings, other studies have found that ratings were perceived to be more useful than comments, (Atwater & Brett, 2006). It would be useful for future research to look at different organizational contexts and try to figure out in which organizational contexts comments are perceived as a useful appraisal format and which in which contexts ratings are perceived as a more useful appraisal format.

Conclusion

The purpose of this study was to examine the difference between appraisal formats, ratings and comments. Furthermore this study looked to identify individual differences that might influence individuals perceiving one format as more useful than another. While largely the hypotheses of this study were unsupported, some support was found that appraisal formats differ, and that individual differences can influence how useful individuals find each appraisal format. By improving upon the methodology and design of this study, future researchers can build upon this study to further examine the differences between ratings and comments. Since past research has focused on ratings and largely ignored comments, while both formats often appear together on performance appraisals, it is important for researchers to explore the role that comments play in the appraisal process. Furthermore it is important to examine what role the two formats play together, exploring the strengths and weaknesses of each format, in order to create a strong performance appraisal process.

References

- Abrow, H.A., & Abrishamkar, M.M. (2013). Individual differences as a moderator of the effect of organisational commitment on readiness for change: A study of employee in the higher education sector in Iraq. *International Journal of Management*, 30, 294-309.
- Ajzen I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211.
- Algozzine, B., Beattie, J., Bray, M., Flowers, C., Gretes, J., Howley, L., Mohanty, G., et al. (2004). Student evaluation of college teaching: A practice in search of principles. *College Teaching*, 52, 134-141.
- Ashford, S.J. (1986). Feedback-seeking in individual adaptation: A resource perspective. *Academy of Management Journal*, 29, 465-487.
- Ashford, S.J. (1989). Self assessment in organizations: A literature review and integrative model. *Research in Organizational Behavior*, 11, 133-174.
- Atwater, L.E., & Brett, J.F. (2005). Antecedents and consequences of reactions to developmental 360 degree feedback. *Journal of Vocational Behavior*, 66, 532-548.
- Austin, J.T. & Villanova, P. (1992). The criterion problem: 1917-1992. *Journal of Applied Psychology*, 77, 836-874.
- Bailey, C.T. (1983). *The measurement of job performance*. Aldertshot, England: Gower Press.
- Barry, C.T., Chaplin, W.F., & Grafeman, S.J. (2006). Agression following performance feedback: The influences of narcissism, feedback valence, and comparative standard. *Personality and Individual Differences*, 41, 177-187.
- Bell, S.T., & Arthur, W. (2008). Feedback acceptance in developmental assessment centers: The role of feedback message, participants personality, and affective response to the feedback session. *Journal of Organizational Behavior*, 29, 681-703.
- Beran, T., Violato, C., Kline, D., & Frideres, J. (2005). The utility of student ratings of instruction of students, and administrators: A consequential validity study. *Canadian Journal of Higher Education*, 35, 49-70.
- Bizzie, L., & Soda, G. (2011). The paradox of authentic selves and chameleons: Self-monitoring, Perceived job autonomy, and contextual performance. *British Journal of Management*, 22, 324-339.
- Boyce, B.A., Gano-Overway, L.A., & Campbell, A.L. (2009). Perceived motivational climate's

- influence on goal orientations, perceived competence, and practice strategies across the athletic season. *Journal of Applied Sport Psychology*, 21, 381-394.
- Bretz, R., Milkovitch, G.T., & Read, W. (1992). The current state of performance appraisal research and practice. *Journal of Management*, 18, 321-352.
- Brockner, J., Derr, W.R., & Laing, W.N. (1987). Self-esteem and reactions to negative feedback: Toward greater generalizability. *Journal of Research in Personality*, 21, 318-333.
- Brutus, S. (2009). Words versus numbers: A theoretical exploration of giving and receiving narrative comments in performance appraisal. *Human Resource Management Review*, 20, 144-157.
- Bryman, A., Bell, E., Mills, A.J., & Yue, A.R. (2011). *Business research methods*. Don Mills, Ontario: Oxford University Press.
- Butler, R. (2000). Making judgements about ability: The role of implicit theories of ability in moderating inferences from temporal and social comparison information. *Journal of Personality & Social Psychology*, 78, 965-978.
- Button, S. B., Mathieu, J. E., & Zajac, D. M. (1996). Goal orientation in organizational research: A conceptual and empirical foundation. *Organizational Behavior and Human Decision Processes*, 67, 26-48.
- Cardy, R. L., & Dobbins, G. H. (1994). *Performance appraisal: Alternative perspectives*. Cincinnati, OH: South-Western Publishing.
- Cappelli, P., & Scherer, P.D. (1991). The missing role of context in OB: The need for meso-level approach. *Research in Organizational Behavior*, 13, 55-110.
- 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*, 83, 615-633.
- Chen, G., Gully, S.M., Whiteman, J., & Kilcullen, R.N. (2000). Examination of relationships among trait-like individual differences, and learning performance. *Journal of Applied Psychology*, 85, 835-847.
- Chhabra, B. (2013). Locus of control as a moderator in the relationship between job satisfaction and organizational commitment: A study of Indian IT professionals. *Organizations and Markets in Emerging Economies*, 4, 25-41.
- Cron, W.L., Slocum J.W., Vande Walle, D. & Fu, F.Q. (2005). The role of goal orientation

- on negative emotions and goal setting when initial performance falls short of one's Performance goals. *Human Performance*, 18, 55-80.
- Dalal, D.K. (2013). Dealing with deliberate distortions: Methods to reduce bias in self-report measures of sensitive constructs. *Dissertation Abstracts International*, 73.
- Dalessio, A.T. (1998). Using multi-source feedback for employee development and personnel decisions. In J.W. Smither (Ed.) *Performance appraisal: State of the art in practice* (pp. 278-330). San Francisco, CA: Jossey-Bass.
- Denson, N., Loveday, T., & Dalton, H. (2010). Student evaluation of courses: What predicts satisfaction? *Higher Education Research and Development*, 29, 339-356.
- Diekmann, K.A., Barsness, Z.I., & Sondak, H. (2004). Uncertainty, fairness perceptions, and job satisfaction: A field study. *Social justice research*, 17, 237-255.
- Dodd, N.G., & Ganster, D.C. (1996). The interactive effects of variety, autonomy, and feedback on attitudes and performance. *Journal of organizational behavior*, 17, 329-347.
- Dusterhoff, C., Cunningham, J.J., & MacGregor, J. (2014). The effects of performance ratings, leader-member exchange, perceived utility, and organizational justice on performance appraisal satisfaction: Applying moral judgement perspective. *Journal of Business Ethics*, 119, 265-273.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040-1048.
- Farr, J.L., Hofmann, D.A., & Ringerbach, K.L. (1993). Goal orientation and action control theory: Implications for industrial organizational psychology. *International Review of Industrial Organizational Psychology*, 8, 193-232.
- Ferris, G.R., and King, T. (1992). The politics of age discrimination in organizations. *Journal of Business Ethics*, 11, 341-350.
- Ferris, G.R., Munyon, T.P., Basik, K., & Buckley, M.R. (2008). The performance evaluation context: Social, emotional, cognitive, political, and relationship components. *Human Resource Management review*, 18, 146-163.
- Fersti, K.L., & Bruskiwicz, K.T. (2000). Self-other agreement and cognitive reactions to multirater feedback. Paper presented at 15th Annual Conference of the Society of Industrial and Organizational Psychology, New Orleans, LA.

- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.
- Feys, M., Anseel, F. & Wille, B. (2011). Improving feedback reports: The role of procedural information and information specificity. *Academy of Management Learning and Education*, 10, 661-681.
- Firth, L., Mellor, D.J., Morre, K.A., & Loquet, C. (2004). How can managers reduce employee intention to quit? *Journal Managerial Psychology*, 19, 170-187.
- Fishbein, M. & Ajzen, I. (1975). Belief, attitude, intention, and behavior: An introduction to theory and research. Reading, MA: Addison-Wesley.
- Fletcher, C. (2001). Performance appraisal and management: The developing research agenda. *Journal of Occupational and Organisational Psychology*, 74, 472-487.
- Folger, R. (1977). Distributive and procedural justice: Combined impact of "Voice" and improvement on experienced inequity. *Journal of Personality and Social Psychology*, 35, 108-119.
- Ghorpade, J. (2000). Managing five paradoxes of 360-degree feedback. *Academy of Management Executive*, 14, 140-150.
- Gilliland, S.W. & Langdon, J.C. (1998). Creating performance management systems that promote perceptions of fairness. In , Smither, J.W. (Ed.) Performance appraisal: State of the art and practice. Jossey-Bass. San Francisco.
- Godin, G., & Kok, G. (1996). The theory of planned behavior: A review of its applications to health related behaviors. *American Journal of Health Promotion*, 11, 87-98.
- Goodman, J.S., & Wood, R.E. (2004). Feedback specificity, learning opportunities, and learning. *Journal of Applied Psychology*, 89, 809-821.
- Greenberg, J. (1990). Organizational justice, yesterday, today, and tomorrow. *Journal of Management*, 16, 399-432.
- Gordon, M.E., & Johnson, W.A. (1982). Seniority: A review of its legal and scientific standing. *Personnel Psychology*, 35, 255-280.
- Hackman, J.R., & Lawler, E.E. (1971). Employee reactions to job characteristics. *Journal of Applied Psychology Monograph*, 55, 259-286.
- Hackman, J.R., & Oldman, G.R. (1976). Motivation through the design of work: test of a theory. *Organizational Behavior and Human Performance*, 16, 250-279.

- Hair, J.F., Black, W.C., Babin, B.J., & Anderson, R.E. (2009). *Multivariate data analysis* 7th edition. Upper Saddle River, N.J: Prentice Hall
- Heslin, P.A., & Latham, G.P. (2005). The effect of implicit person theory on performance appraisal. *Journal of Applied Psychology*, 5, 842-856.
- Herold, D.M., & Fedor, D.B. (1998). Individuals' interactions with feedback environment: The role of domain specific individual differences. In K.M. Rowland & G.R. Ferris (Eds.), *Research in personnel and human resource management*, (pp. 215-254). Greenwich CT: Jai Press.
- Jacobs, G., Belschak, F.D., & Hartog, D.N.D. (2014). (Un)Ethical behavior and performance appraisal: The role of affect, support, and organizational justice. *Journal of Business Ethics*, 121, 63-76.
- Jacobs, R., Hofmann, D.A., & Kriska, S.D. (1990). Performance and seniority. *Human Performance*, 3, 107-121.
- Judge, T.A., & Ferris, G.R. (1993). Social context of job performance evaluation decisions. *Academy of Management Journal*, 36, 80-105.
- Kleinbaum, D.G., Kupper, L.L., Nizam, A., & Muller, K.E. (2008). *Applied regression analysis and other multivariable methods*, 4th edition. Belmont, CA: Thomson-Brooks/Cole.
- Martinez, M.A., Zeichner, A., Reidy, D.E., & Miller, J.D. (2007). Narcissism and displaced aggression: effects of positive, negative, and delayed feedback. *Personality and Individual Differences*, 44, 140-149.
- Martocchio, J. J., & Hertenstein, E. J. (2003). Learning orientation and goal orientation context: Relationships with cognitive and affective learning outcomes. *Human Resource Development Quarterly*, 14, 413–434.
- McMurtie, H., Baxter, J.S., Obonsawin, M.C., and Hunter, S.C. (2012). Consistent witness responses: The effects of age and negative feedback. *Personality and Individual Differences*, 53, 958-962.
- McMurtrie, B. (2013). Data support changes of gender bias in political science. *Chronicles of Higher Education*, 60, A17.
- Miller, R.U. (1980). Hospitals. In G.G. Somers (Ed.), *Collective bargaining: Contemporary American experience* (pp. 1-48). Madison, WI: Industrial relations research association.
- Mintzberg, H. (1973). *The nature of managerial work*. New York: NY: Harper & Row.

- Moreno, R., & Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology, 97*, 117–128.
- Morrison, E.W. (1993). Newcomer information seeking: Exploring types, modes, sources, and outcomes. *Academy of Management Journal, 36*, 173-183.
- Murphy, K.R. & Cleveland, J.N. (1995). Understanding performance appraisal: Social, organizational, and goal based perspectives. Thousand Oaks, CA: Sage.
- Nasser, F. & Fresko, B. (2002). Faculty views of student evaluation of college teaching. *Assessment and Evaluation in Higher Education, 27*, 187-198.
- Landy, F.H. & Farr, J. (1980). Performance rating. *Psychological Bulletin, 87*, 172-207.
- Levy, P.E., Albright, M.D., Cawley, B.D., & Williams, J.R. (1995). Situational and individual determinants of feedback seeking: A closer look at the process. *Organizational Behavior and Human Decision Processes, 62*, 23-37.
- Levy, P., & Williams, J.R. (2004) The social context of performance appraisal: A review and agenda for the future. *Journal of Management, 30*, 881-995.
- Lind, E. A., & Tyler, T.R. (1988). The social psychology of procedural justice, New York, NY: Plenum.
- Ogunfowora, B., Bourdage, J., & Lee, K. (2010). Rater personality and performance dimension weighting in making overall performance judgment. *Journal of Business Psychology, 25*, 465-476.
- Park, G., Schmidt, A.M., Scheu, C., & DeShon, R.P. (2007). A process model of goal orientation and feedback seeking. *Human Performance, 20*, 119-145.
- Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology, 36*, 859-866.
- Pearce, J.L., & Porter, L.W. (1986). Employee responses to format performance appraisal Feedback. *Journal of Applied Psychology, 71*, 211-218.
- Penny, A.R., & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research, 74*, 215-253.
- Phillips, J. M., & Gully, S. M. (1997). Role of goal orientation, ability, need for achievement and locus of control in the self-efficacy and goal-setting process. *Journal of Applied Psychology, 82*, 792–802.

- Pichler, S. (2012). The social context of performance appraisal and appraisal reactions: A meta-analysis. *Human Resource Management, 51*, 709-732.
- Poon, J.M.L. (2004). Effects of performance appraisal politics on job satisfaction and turnover intention. *Personnel Review, 33*, 322-334.
- Potosky, R. (2010). Goal orientation, learning self efficacy, and climate perceptions in a post acquisition corporate context. *Human Resource Quarterly, 21*, 273-289.
- Rogers, E., Rogers, C.W., & Metlay, W. (2002). Improving the payoff from 360-degree feedback. *Human Resource Planning, 25*, 44-54.
- Rosenberg, M. (1965). *Society and the adolescent self image*. Princeton, NJ: Princeton University Press.
- Ryan, A.M., & Sackett, P.R. (1992). Relationships between graduate training, professional affiliation, and individual psychological assessment practices for personnel decisions. *Personnel Psychology, 45*, 363-387.
- Seijts, G.H., & Latham, G.P. (2005). Learning versus performance goals: When should each be used? *Academy of Management Executive, 19*, 124-131.
- Sheier, M.F., & Carver, C.S. (1982). Self-directed attention and the comparison of the self with standards. *Journal of Experimental Social Psychology, 19*, 205-222.
- Swiercz, P.M., Bryan, N.B., Eagle, B.W., Bizzotto, V., & Renn, R.W. (2012). Predicting employee attitudes and performance from perceptions of performance appraisal fairness. *The Business Renaissance Quarterly, 7*, 25-46.
- Smither, J.W., Wohler, A.J., & London, M. (1995). A field study of reactions to normative versus individualized upward feedback. *Group and Organization Management, 89*, 61-89.
- Srivastava, S. (2013). Job satisfaction and organizational commitment relationship: Effect of personality variables. *Vision: The Journal of Business Perspective, 17*, 159-167.
- Steensma, H., & Visser, E. (2007). Procedural justice and supervisors' personal power bases: Effects on employees' perceptions of performance appraisal sessions, commitment, and motivation. *Collective Negotiations, 31*, 101-118.
- Subramanian, G.H. (1994). A replication of perceived usefulness and ease of use measurement. *Decision Sciences, 25*, 863-874.

- Swann, W. B., Jr., Pelham, B. W., & Krull, D. S. (1989). Agreeable fancy or disagreeable truth? How people reconcile their self-enhancement and self-verification needs. *Journal of Personality and Social Psychology*, *57*, 782–791.
- Thorndike, R.L. (1949). *Personnel selection*. New York, NY: Wiley.
- Tuckey, M., Brewer, N. & Williamson, P. (2002). The influence of motives and goal orientation on feedback seeking. *Journal of Occupational and Organizational Psychology*, *75*, 195-216.
- Tuytens, M. & Devos, G. (2013). How to activate teachers through teacher evaluations. School Effectiveness and School Improvement. Retrieved April 10th, 2014, from <http://hdl.handle.net/1854/LU-4206946>
- Vandewalle, D., & Cummings, L. L. (1997). A test of the influence of goal orientation on the feedback seeking process. *Journal of Applied Psychology*, *82*, 390–400.
- Vandewalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and psychological measurement*, *57*, 995-1015.
- Wanberg, C.R., & Kammeyer-Mueller, J.D. (2000). Predictors and outcomes of proactivity in the socialization process. *Journal of Applied Psychology*, *85*, 373-385.
- Wayne, S.J & Liden R.C. (1995). Effects of impression management on performance ratings: A longitudinal study. *The academy of Management Journal*, *38*, 232-260.
- Westerlund, J. (2008). Class size and student evaluations in Sweden. *Education Economics*, *16*, 19-28.
- Wright, S.L., & Jenkins-Guarnieri, M.A. (2012). Student evaluations of teaching: Combining the meta-analyses and demonstrating further evidence for effective use. *Assessment and Evaluation in Higher Education*, *37*, 683-699.

Appendix A
Email Asking For Participation

Dear Faculty member,

I am a student in the Master of Science in Administration (MScA) program of John Molson School of Business and I am seeking your voluntary participation for a research project. Under the supervision of Professor Stéphane Brutus I am conducting a study that focuses on your format preference (ie. Ratings vs. comments) of teaching evaluations. The survey is short and should take about 5-10 minutes. All of the information that you provide by filling out the survey is anonymous and no identifying information is collected.

Thank you in advance,

Andrew Watson

MSc Candidate

Appendix B
Questionnaire

This section of the questionnaire focuses on your experience being employed at Concordia University.

How many years have you been employed at Concordia University? (Please enter a number. For example for thirty years, type 30)

Which faculty at Concordia University are you a member of?

- Faculty of arts and science
- Faculty of engineering and computer science
- Faculty of Fine arts
- John Molson school of business

For this section I am interested in your general opinion on the teaching evaluations that you receive from students at the end of each semester. Please indicate how much you agree or disagree with the following statements.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Ratings provide information that helps me make adjustments to my communication skills with students.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ratings provide information that enhances my effectiveness as a professor.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ratings provide information that makes it easier for me to do my job.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Comments provide information that helps me make adjustments to my communication skills with students.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Comments provide information that enhances my effectiveness as a professor.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Comments provide information that makes it easier for me to do my job.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

This section focuses on your past teaching evaluations. Please answer the following questions thinking about the AVERAGE teaching evaluation you have received in the past. As you may teach many different types of classes, (undergraduate, graduate, lab, etc) please focus ONLY on the one type of class. This could be the class that you find most important, the class you have taught the most, or the class you have most recently taught.

Please indicate what type of class you will be focusing on.

On average how many students do you have per class? (Please enter a number. For example for thirty students, type 30)

On average how many comments do you receive per class? (Please enter a number. For example for thirty comments, type 30)

	Very Negative	Negative	Neutral	Positive	Very Positive
Thinking back to your past teaching evaluations please indicate how positive or negative the overall rating score you received was.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Thinking back to your past teaching evaluations please indicate how positive or negative the overall comments were.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

This section of the questionnaire focuses on how you react to and seek out feedback. Please indicate how much you agree or disagree with the following statements.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
It is important for me to obtain useful information about my performance.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Receiving feedback about my performance helps me to improve my skills.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would like to obtain more information to let me know how I am performing.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I would like to receive more useful information about my performance	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I'm not really concerned whether or not I receive useful information about my performance.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Feedback is not really useful to help me improve my performance.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Obtaining useful feedback information is not very important to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I don't really require more feedback to let me know how I am performing.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

This section of the questionnaire focuses on how you react to and seek out feedback. Please indicate how much you agree or disagree with the following statements.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
If I received negative feedback I would have a negative attitude towards myself, so I try to avoid criticism.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Negative feedback doesn't really lower my self-worth, so I don't go out of my way to avoid it.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Receiving negative feedback wouldn't really change the way I feel about myself.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It's hard to feel good about myself when I receive negative feedback.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I don't really worry about getting negative feedback because I still feel I am a person of worth.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I try to avoid negative feedback because it makes me feel bad about myself.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I worry about receiving feedback that is likely to be negative because it hurts to be criticized.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Negative feedback doesn't really worry me because I still have a positive attitude towards myself.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

This section of the questionnaire focuses on the goals that you have while performing your job. Please indicate how much you agree or disagree with the following statements.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I prefer to do things I can do well rather than things that I do poorly.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I'm happiest at work when I perform tasks on which I know that I won't make any errors.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The things I enjoy the most are the things I do best.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The opinions others have about how well I can do certain things are important to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel smart when I do something without making any mistakes.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I like to be fairly confident that I can successfully perform a task before I attempt it.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I like to work on tasks that I have done well on in the past.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel smart when I can do something better than most other people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The opportunity to do challenging work is important to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I fail to complete a difficult task, I plan to try harder the next time I work on it.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

I prefer to work on tasks that force me to learn new things.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The opportunity to learn new things is important to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I do my best when I'm working on a fairly difficult task.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I try hard to improve on my past performance.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The opportunity to extend the range of my abilities is important to me.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
When I have difficulty solving a problem, I enjoy trying different approaches to see which one will work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you have any additional comments please write them in the space provided below.

Appendix C
Consent Form

CONSENT TO PARTICIPATE IN PREFERENCE FOR APPRAISAL FORMAT.

I understand that I have been asked to participate in a research project being conducted by Andrew Watson of the management department of Concordia University (*A_wat@jmsb.concordia.ca*, (438) 883-6593) under the supervision of Stéphane Brutus, PhD of the management department of Concordia University (*Brutus@jmsb.concordia.ca*).

A. PURPOSE

I have been informed that the purpose of the research is to study a professor's perception of their teacher evaluations. More specifically the study will be examining how useful a professor finds receiving both numbers and words as feedback on their teaching evaluations.

B. PROCEDURES

I understand that I will be asked to fill out a questionnaire on line; as well I understand that the questionnaire will take about 5 to 10 minutes to fill out. I understand that my identity will be anonymous.

C. RISKS AND BENEFITS

I understand that there will be no risks by filling out the following questionnaire.

D. CONDITIONS OF PARTICIPATION

- I understand that I am free to withdraw my consent and discontinue my participation at anytime without negative consequences.
- I understand that my participation in this study is *confidential and anonymous*
- I understand that the data from this study may be published.

I HAVE CAREFULLY STUDIED THE ABOVE AND UNDERSTAND THIS AGREEMENT.
BY COMPLETING THE FOLLOWING QUESTIONNAIRE I FREELY CONSENT AND
VOLUNTARILY AGREE TO PARTICIPATE IN THIS STUDY.

If at any time you have questions about the proposed research, please contact the study's
Principal Investigator

*Andrew Watson,
Management department, Concordia University*

(438) 883 6593

A_wat@Jmsb.Concordia.ca

Faculty supervisor

Management department, Concordia University

Brutus@jmsb.concordia.ca

If at any time you have questions about your rights as a research participant, please contact the Research Ethics and Compliance Advisor, Concordia University, 514.848.2424 ex. 7481

ethics@alcor.concordia.ca

Appendix D
Follow Up Email

Dear faculty member,

I would like to thank all of those who have taken the time to respond to my last email, it is greatly appreciated.

I still need a few more respondents to have a large enough sample size for my thesis, and it would be greatly appreciated if you would take 5-10 minutes to fill out my survey if you have not already done so.

The survey can be found

here: http://jmsbconcordia.qualtrics.com/SE/?SID=SV_eh8cyAHaKSO2UPr

Thank you,

Andrew Watson
MSc Candidate